

Bernard of Clairvaux and Nicholas of Montiéramey: Tracing the Secretarial Trail with Computational Stylistics

By Jeroen De Gussem

Although the past few decades of medieval studies have witnessed some renewed interest in the collaborative process by which medieval Latin prose was composed, such interest has nevertheless remained all too scant, and only few solutions have been offered to cope with the difficulties that rise in cases of dubious authorship. Whereas it has been rightly acknowledged that scribes, notaries, and secretaries should be regarded not merely as instrumental in the literary process, but as active participants in the composition process who have left a considerable impact on the image and style of the *dictator*¹ and on the materialization and dissemination of the text,² this acknowledgement has nevertheless been accompanied by difficulties

This article is a result of the research project “Collaborative Authorship in Twelfth-Century Latin Literature: A Stylometric Approach to Gender, Synergy and Authority,” funded by the Ghent University Special Research Fund (BOF). Its execution rests on a close collaboration between the Henri Pirenne Institute for Medieval Studies at Ghent University, the CLiPS Computational Linguistics Group at the University of Antwerp, and the Centre Traditio Litterarum Occidentium division for computer-assisted research into Latin language and literature housed in the Corpus Christianorum Library and Knowledge Centre of Brepols Publishers in Turnhout (Belgium). I am much indebted to the wisdom and continuous and patient guidance of Jeroen Deplouge, Wim Verbaal, and Mike Kestemont, who—each in their respective fields of expertise (medieval cultural history, Latin medieval literature, and computational stylistics)—have tremendously inspired and challenged me in writing this piece. Their voices inevitably resound from this text, so much so that I cannot solely take credit for the whole. I also warmly thank my colleagues from the Latin and History Department in Ghent who have gone through the trouble of reading my preliminary drafts. In particular, Dinah Wouters, Micol Long, and Theo Lap have my sincerest gratitude for personally sending me their valuable feedback. In conclusion, my gratitude goes out to Paul De Jongh, Bart Janssens, Jeroen Lauwers, and Luc Jocqué of Brepols for their commitment to this project.

¹ A more general introduction to twelfth-century notaries, especially in respect to epistolography, can be found in Giles Constable, “Dictators and Diplomats in the Eleventh and Twelfth Centuries: Medieval Epistolography and the Birth of Modern Bureaucracy,” *Dumbarton Oaks Papers* 46 (1992): 37–46, at 38, where he stresses that “[notaries] took on a new importance in the eleventh and twelfth centuries, when they formed a distinct group of recognizable personalities whose activities extended beyond the scriptorium.” Lynn Staley Johnson, albeit focusing mainly on late medieval female authors, has drawn attention to how “scribes not only left their marks upon the manuscripts they copied, they also functioned as interpreters, editing and consequently altering the meaning of texts. Writers, however, did not simply employ scribes as copyists; they elaborated upon the figurative language associated with the book as a symbol and incorporated scribes into their texts as tropes,” in “The Trope of the Scribe and the Question of Literary Authority in the Works of Julian of Norwich and Margery Kempe,” *Speculum* 66 (1991): 820.

² See Bernard Cerquiglini, *In Praise of the Variant: A Critical History of Philology*, trans. Betsy Wing (Baltimore, 1999); Stephen G. Nichols, “Introduction: Philology in a Manuscript Culture,” *Speculum*

Speculum 92/S1 (October 2017). © 2017 by the Medieval Academy of America. All rights reserved.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0), which permits non-commercial reuse of the work with attribution.

For commercial use, contact journalpermissions@press.uchicago.edu.

DOI: 10.1086/694188, 0038-7134/2017/92S1-0008\$10.00.

in defining the exact extent and sphere of influence of such secretarial mediation. This challenge is especially marked in the oeuvre of Bernard of Clairvaux (1090–1153). By 1145, the abbot's acclaim as the icon and figurehead of the Cistercian movement had brought along such a considerable administrative workload that the assistance of a group of secretaries—which, it could be argued, amounted to a kind of chancery³—was indispensable. These secretaries acted as Bernard's stand-ins and spared him the time and effort it would have cost if he had had to take up the quill himself at every single occasion.⁴ The *reportatio*, as it was called, entailed that the contents of Bernard's letters or sermons be engraved on wax tablets in a tachygraphic fashion. The cues, keywords, and biblical references that Bernard had spoken aloud provided a framework that captured the gist of his diction.⁵ Afterwards, the scribe reconstructed what he had heard as a text on parchment, which could pass for Bernard of Clairvaux's in its literary allure. Among these amanuenses, Nicholas of Montiéramey († 1176/78) was a focal figure and a highly skilled imitator of his master's writing style. The influence of Nicholas's mediation on several particular texts within Bernard's corpus, and more generally on his entire oeuvre, has been subject to much debate. This article revisits the authorship of a selection of texts from Bernard's corpus. A detailed listing of the texts under scrutiny can be consulted in the Appendix of Tables (Tables 3–11). Generally, the corpus comprises Nicholas of Montiéramey's letters and sermons⁶ and Bernard of Clairvaux's letter corpus

65 (1990): 1–10; Eric H. Reiter, "The Reader as Author of the User-Produced Manuscript: Reading and Writing Popular Latin Theology in the Late Middle Ages," *Viator* 27 (1996): 151–69.

³ Some could argue that the word "chancery" is inappropriately used of Bernard's scriptorium, as it was not primarily a formal or institutional body of administration charged with the composition and dispatch of official documents.

⁴ The workings of Clairvaux's scriptorium are extensively investigated in Peter Rassow, "Die Kanzlei St. Bernhards von Clairvaux," *Studien und Mitteilungen zur Geschichte des Benediktiner-Ordens* 34 (1913); and Jean Leclercq, "Saint Bernard et ses secrétaires," in Leclercq, *Recueil d'études sur saint Bernard et ses écrits* (hereafter *Recueil d'études*), 4 vols. (Rome, 1962–87), 1:3–25. Constable also commented on the difficulty of the redaction process: "Aside from a few outlines dictated by Bernard or based on sermons he gave, most of the surviving texts are later compositions drawn up by either himself or his secretaries, and they bear little resemblance to what he actually preached, if they were ever delivered," in Giles Constable, "The Language of Preaching in the Twelfth Century," *Viator* 25 (1994): 131–52, at 134.

⁵ Stenography, or shorthand, systems had been forgotten by the twelfth century, making place for tachygraphy, a rapid form of writing; see Malcolm Beckwith Parkes, "Tachygraphy in the Middle Ages: Writing Techniques Employed for *Reportationes* of Lectures and Sermons," in *Scribes, Scripts and Readers: Studies in the Communication, Presentation and Dissemination of Medieval Texts* (London, 1991), 19–33.

⁶ Nicholas of Montiéramey's letters can be found under PL 196:1593a–1651d. The sermons have been identified by Jean Leclercq in "Les collections de sermons de Nicolas de Clairvaux," *Recueil d'études*, 1:52–54. They are collected among those of Peter Damian in PL 144, more specifically "Sermo in nativitate S. Ioannis Baptistae" (627), "Sermo in natali apostolorum Petri et Pauli" (649), "Sermo in natali S. Benedicti de evangelio" (548), "Sermo in festivitate S. Mariae Magdalenae" (660), "Sermo in festivitate S. Petri ad vincula" (646), "Sermo in assumptione B. Mariae" (717), "Sermo in nativitate B. Mariae" (736), "Sermo in exaltatione S. crucis" (761), "Sermo in festivitate angelorum" (794), "Sermo in dedicatione ecclesiae" (897), "Sermo in festivitate S. Victoris" (732), "Sermo in festivitate omnium sanctorum" (811), "Sermo in festivitate S. Martini" (815), "Sermo in festivitate S. Andreae" (828), "Sermo in festivitate B. Nicholai" (835), "Sermo in festivitate B. Mariae" (557), "Sermo in vigilia nativitatis" (839), "Sermo in nativitate Domini" (847), and "Sermo in festivitate B. Stephani" (853).

(*Corpus epistolarum*), *Sermones de diversis* (hereafter *De diversis*), and *Sermones super Cantica canticorum*.⁷

As a means of determining the extent of Nicholas's stylistic presence in the aforementioned works, this article advocates computational stylistics (or stylometry), which is a method that detects stable and recurring patterns of writing style in texts that have been reduced to a range of marked style features. These features are consequently quantified to numerical data in order to gain objective and measurable ground for distinguishing among works of varying authorship. Computational stylistics claims to offer a scientific and objective "distant reading"⁸ of literature, as opposed to human expert-based methods, which are often liable to intersubjectivity.⁹ A stylometric approach to the case of Bernard of Clairvaux and Nicholas of Montiéramey will prove rewarding on two levels. On the first level, we will determine the authorship of the aforementioned texts, some of which have been subject to long-standing dispute. This will address an elementary concern within Bernardian studies, which Gillian Rosemary Evans has aptly put forward as such: "A turn of phrase can bear only so heavy a load of interpretation as it exactly reflects the author's thought."¹⁰ On a second level, however, this study should also raise awareness of how the (mis)attributions of disputed medieval texts, as they have occurred in earlier studies, are often guided by personal intuitions or theoretical convictions of medieval authorship.¹¹ This implies that, as we begin to challenge the established

⁷ These works are edited in the *Sancti Bernardi opera* (hereafter SBO), ed. Jean Leclercq et al., 8 vols. (Rome, 1957–77): the *Corpus epistolarum* (vols. 7–8), *Sermones de diversis* (vol. 6), and *Sermones super Cantica canticorum* (vols. 1–2).

⁸ The term "distant reading" was coined by Franco Moretti, "Conjectures on World Literature," *New Left Review* 1 (2000): 54–68, and was further developed in his monograph *Graphs, Maps, Trees: Abstract Models for Literary History* (Brooklyn, 2005).

⁹ It was Frederick Mosteller and David L. Wallace's influential study on the disputed authorship of the eighteenth-century *Federalist Papers* in the early 1960s that would launch statistical approaches as tools by which to objectively determine authorship, currently known as nontraditional authorship attribution: see Frederick Mosteller and David L. Wallace, *Applied Bayesian and Classical Inference: The Case of The Federalist Papers* (New York, 1964). For four excellent state-of-the-art surveys on the history of nontraditional authorship attribution and the current debate within the field, see Patrick Juola, "Authorship Attribution," *Foundations and Trends in Information Retrieval* 1 (2008): 233–334; Efstathios Stamatatos, "A Survey of Modern Authorship Attribution Methods," *Journal of the Association for Information Science and Technology* 60 (2009): 538–56; Moshe Koppel, Jonathan Schler, and Shlomo Argamon, "Computational Methods in Authorship Attribution," *Journal of the Association for Information Science and Technology* 60 (2009): 9–26; and Walter Daelemans, "Explanation in Computational Stylometry," in *Computational Linguistics and Intelligent Text Processing* 7817, ed. Alexander Gelbukh (Berlin, 2013), 451–62.

¹⁰ Gillian Rosemary Evans, *Bernard of Clairvaux*, Great Medieval Thinkers (New York, 2000), 20.

¹¹ The bibliography of scholarship on medieval authorship is extensive and cannot be listed here in full. The following titles should nevertheless point any reader who is interested in the theory of medieval authorship in the right direction. A major work of reference is by Alastair J. Minnis, *Medieval Theory of Authorship: Scholastic Literary Attitudes in the Later Middle Ages* (Philadelphia, 1984), whose approach to medieval authorship is largely based on study of the prologues of commentaries and exegetical works especially from the later Middle Ages (the scholastic period). In this later period he describes how the schools defined for themselves a framework of literary theory from the newly translated Aristotelian logic, through which they could approach the biblical texts and patristic *auctores* more literally, and therefore more literarily, as "a new type of exegesis emerged, in which the focus had shifted from the divine *auctor* to the human *auctor* of Scripture," 5. Some other indispensable publications on medieval au-

authorship of some of these texts, we also may offer metahistorical reflections on how earlier, more intuitive scholarly approaches have both enriched and yet predetermined our current understanding of medieval authorship. The case of Bernard and Nicholas is particularly suitable to demonstrate how computational stylistics provides new answers to old questions, and raises new questions about old problems.

NICHOLAS OF MONTIÉRAMEY

The daily routines and workings of Clairvaux's chancery are rather poorly documented. We rarely know any of the scribes by name, and for those whom we do—a select group of six—only three give us a faint clue of their specific tasks and responsibilities.¹² Nicholas began serving Bernard as an emissary around 1138–41, carrying letters concerning Abelard's heresy to Rome. At this time he was still chaplain of Hato, bishop of Troyes, and Peter the Venerable's friend and secretary, but he must already have been collaborating with Bernard from 1140 onwards.¹³ He

thorship that are of particular relevance here are Michel Zink, *La subjectivité littéraire au siècle de saint Louis* (Paris, 1985); Stephen C. Jaeger, "Charismatic Body, Charismatic Text," *Exemplaria* 9 (1997), 117–37; Michel Zimmermann, ed., *Auctor et auctoritas: Invention et conformisme dans l'écriture médiévale; Actes du colloque de Saint-Quentin-en-Yvelines (14–16 juin 1999)* (Paris, 2001); Edith Wenzel, "Der Text als Realie? Auf der Suche nach dem Text und seinem Autor," in *Text als Realie: Internationaler Kongress Krems und der Donau 3. bis 6. Oktober 2000*, Österreichische Akademie der Wissenschaften, Philosophischen-historische Klasse, Sitzungsberichte 704 (Vienna, 2003), 81–95; Jeroen Deploige, "Anonymat et paternité littéraire dans l'hagiographie des Pays-Bas Méridionaux (ca. 920–ca. 1320): Autour du discours sur l'original et la 'copie' hagiographique au Moyen Âge," in *Scribere sanctorum gesta: Recueil d'études d'hagiographie médiévale offert à Guy Philippart*, ed. Étienne Renard, Michel Trigalet, Xavier Hermand, and Paul Bertrand (Turnhout, 2005), 77–107; Jan M. Ziolkowski, "Cultures of Authority in the Long Twelfth Century," *Journal of English and German Philology* 108 (2009), 421–48; Stephen Partridge and Erik Kwakkel, *Author, Reader, Book: Medieval Authorship in Theory and Practice* (Toronto, 2012).

¹² We know that Bernard's earliest secretary was William of Rievaulx. He must have been active from 1120 until 1132, before traveling to northern England to establish the monastery of Rievaulx in the diocese of York, a daughter house for Clairvaux, to become its first abbot: see Rassow, "Die Kanzlei St. Bernhards von Clairvaux," 5. William's intimate bond with Bernard must have established a solid base upon which Clairvaux and Rievaulx were able to cooperate, communicate, and exchange recruits: see Brian Patrick McGuire, "Introduction," in *A Companion to Bernard of Clairvaux*, ed. McGuire, Brill's Companions to the Christian Tradition 25 (Leiden, 1992), 8. Three other names that have come down to us are Balduin of Pisa, Gerard of Peronne, and Raynaud of Foigny, but none of these seems to have had much significance: see Evans, *Bernard of Clairvaux*, 20. A more important personality was Geoffrey of Auxerre, who was a former student of Peter Abelard and allegedly denounced the Parisian schools in favor of the monastery after having witnessed Bernard's genius and eloquence in preaching: "continuo tres ex illis compuncti sunt et conversi ab inanibus studiis ad verae sapientiae cultum, abrenuntiantes saeculo et Dei famulo adhaerentes," in Geoffrey of Auxerre, *Liber quartus sancti Bernardi abbatis Clarae-vallensis vita* (PL 185:327). He entered Clairvaux in 1140 and became the abbot's secretary in 1145, a time when the administrative obligations in Clairvaux reached their peak and an official chancery had been established. He would become abbot of Clairvaux himself in 1163 but had to abdicate his leadership after two years, presumably as a consequence of an internal dispute over the papal schism between Alexander III and Victor IV: see Ferruccio Gastaldelli, "Introduzione," in Goffredo di Auxerre, *Super Apocalypsim*, ed. Gastaldelli, *Temi e Testi* 17 (Rome, 1970), 14–15.

¹³ There is scholarly debate over when exactly Nicholas initiated his collaboration with Bernard, but recent research tends to agree that it must have been earlier than his accession in 1145–46. See Constable, "Dictators and Diplomats," 43–44: "Nicholas was at Clairvaux probably from the early 1140s

Speculum 92/S1 (October 2017)

would officially become a monk at Clairvaux around the end of 1145. His literary qualities, likely to have been acquired through his education in the Benedictine abbey of Montiéramey,¹⁴ enabled him to enter the scriptorium immediately and officially become Bernard's closest secretary. He appears to have been responsible for supervising the workings of the chancery,¹⁵ and he may have been the monastery's librarian.¹⁶ But their friendship knew an abrupt and painful ending in the final years of Bernard's life, around 1151–52, when Nicholas must have severely breached his master's trust. In a letter to Pope Eugene III, we find Bernard disconcerted over the fact that letters had been sent out under his name and seal by "false brethren" without his permission.¹⁷ Later, Bernard would identify Nicholas as the culprit among these brethren,¹⁸ although the exact reasons why the latter deserved this accusation are nowhere explicitly disclosed.¹⁹ In any case one can assume from his correspondence and his own words²⁰ that Nicholas's talent as a writer and his "versatility" ingratiated him with the greatest men of his time.²¹ Equally so, Nicholas appears to have had—perhaps through this flamboyance and self-confidence—a talent for making enemies as well.²²

to 1152 and assisted Bernard with his sermons as well as his letters, but he continued to visit Cluny and to serve Peter the Venerable, one of whose letters, we have seen, he presented to Bernard orally"; and Anne-Marie Turcan-Verkerk, "L'introduction de l'ars dictaminis en France: Nicholas de Montiéramey, un professionnel du dictamen entre 1140 et 1158," in *Le dictamen dans tous ses états: perspectives de recherche sur la théorie et la pratique de l'ars dictaminis (XIe–XVe siècles)*, ed. Benoît Grévin and Turcan-Verkerk (Turnhout, 2015), 70: "Ami de Pierre le Venerable, [Nicolas] avait déjà servi les intérêts de Bernard en portant au pape, en 1140–1141, des lettres concernant Abelard—à la rédaction desquelles il avait peut-être déjà participé, comme le suggère le manuscrit Phillipps 1732. . . . Trois billets de recommandation envoyés par Bernard à Innocent II entre 1138 et 1143 semblent le concerner [Epp. 434–36], et montrent que s'il servait Hatton, il le faisait en obéissant à Bernard."

¹⁴ "Nicolas fit ses études à l'abbaye bénédictine de Montiéramey, près de Troyes en Champagne. On parle souvent de lui comme d'un *Magister*": see John Benton, "Nicolas de Clairvaux," in *Dictionnaire de Spiritualité*, 17 vols. (Paris, 1982), 11:256–59.

¹⁵ Leclercq, "Lettres de S. Bernard: Histoire où littérature?," *Recueil d'études*, 4:148.

¹⁶ Giles Constable, "Nicholas of Montiéramey and Peter the Venerable," in Constable, *The Letters of Peter the Venerable*, 2 vols. (London, 1967), 2:321.

¹⁷ "Periclitati sumus in falsis fratribus," Bernard, Ep. 284, SBO 8:198–99.

¹⁸ Ep. 298, likewise addressed to Eugene III (SBO 8:214).

¹⁹ Constable, *The Letters of Peter the Venerable*, 2:327.

²⁰ Nicholas, Ep. 56 (PL 196:1652), "Ab ineunte aetate mea placui magnis et summis principibus hujus mundi."

²¹ The word is Jean Mabillon's in PL 183:26: "Vir fuit ingenii facilis, versatilis, facile in aliorum affectus influens."

²² An example can be found in Nicholas's dispute with Peter of Celle. The two "were at odds over a substantive matter, a theological point about how to treat the attributes of God, and Abbot Peter took offense that Nicholas, who should have possessed the power to triumph with his own (verbal) arms, had used against him the authority and words of great philosophers:" see John Van Engen, "Letters, Schools, and Written Culture in the Eleventh and Twelfth Centuries," in *Dialektik und Rhetorik im früheren und hohen Mittelalter: Rezeption, Überlieferung und gesellschaftliche Wirkung antiker Gelehrsamkeit vornehmlich im 9. und 12. Jahrhundert*, ed. Johannes Fried (Munich, 1997), 27:109; and Julian Haseldine, "Peter of Celle and Nicholas of Clairvaux's Debate on the Nature of the Body, the Soul, and God," in *The Letters of Peter of Celle*, ed. Haseldine (Oxford, 2001), 706–11. The "versatile" aspect of Nicholas's personality, as noted by Mabillon (above, n. 21), therefore also seems to display itself on the level of language. Nicholas was accused of "inverting words and their meaning," a

Speculum 92/S1 (October 2017)

The scandal at Clairvaux and the breach of Bernard's trust has for a long time upheld the portrayal of Nicholas as a disreputable Judas by Bernard's side, an analogy for which Bernard himself was responsible.²³ Conversely and simultaneously, Bernard's status as a saint continued to grow during the intense process of canonization and idealization following his death.²⁴ These respective caricatural depictions, in which Nicholas was deplored as the mistrusted secretary and Bernard praised as the saint who had become victim of textual theft, show through on an academic level as well. Dom Jean Leclercq, one of the most prominent Bernard scholars of the twentieth century, was as relentless as Bernard in accusing Nicholas of deceit, shamelessness, and plagiarism.²⁵ Nicholas's most striking example of seeming textual theft presents itself in his letter to Henry the Liberal, count of Champagne, to whom he humbly offered his services as a secretary shortly after his expulsion from Clairvaux. Accompanying the letter we find nineteen sermons originally attributed to Peter Damian,²⁶ nine sermons attributed to Bernard of Clairvaux,²⁷ and seventy-four short commentaries to the Psalms that are ascribed to Hugh of St. Victor.²⁸ In the letter, Nicholas asserts that these writings are "of my invention, of my style, aside from what I have taken from others in a few places."²⁹ We know this assertion to be true of the nineteen sermons also found among those of Peter Damian, which have been identified by Leclercq as stemming from Nicholas. Bernard's and Hugh's writings, on the other hand, appear to have been copied almost literally, not merely rearranged or paraphrased "in a few places" (*paucis in locis*), as Nicholas seems to suggest. Most of the nine sermons can be found in Bernard's *De diversis*. It is striking that, months

characteristic that Peter interpreted as equal to a falsification of language: "verba quoque et sensus verborum praesumis quandoque invertere": see Peter of Celle, Ep. 66 (PL 202:512).

²³ Bernard literally made the analogy with Judas, which he significantly did not make often in his letters: see Brian Patrick McGuire, "Loyalty and Betrayal in Bernard of Clairvaux," in *Loyalty in the Middle Ages: Ideal and Practice of a Cross-Social Value*, ed. Jörg Sonntag and Coralie Zermatten (Turnhout, 2015), 317–18.

²⁴ As it was first initiated by his biographer William of St. Thierry († 1148). He wrote the *Vita prima Bernardi*, a biography with a hagiographical, panegyric slant. He shares the authorship of the entire *Vita* with Bernard's secretary Geoffrey of Auxerre and the Benedictine abbot Arnaud de Bonneval. Geoffrey was a strong advocate for Bernard's canonization, in which William's texts played a fundamental role.

²⁵ Jean Leclercq, "Les collections de sermons de Nicolas de Clairvaux," *Recueil d'études*, 1:56–58. Leclercq cannot but express his dislike for Nicholas in phrases such as "cet homme sans caractère, mais lettré, doué de mémoire, habile à manier les fiches, prompt à entrer 'dans le personnage' d'un autre, aurait pû être pour S. Bernard un parfait secrétaire, si seulement il avait été honnête," or "or la suite du recueil prouve qu'il était sans scrupules en ce domaine comme en d'autres," or "ainsi les témoignages les plus formels de Nicolas lui-même sont trompeurs, car il ment"; or, in "Deux épîtres de Saint Bernard," *Recueil d'études*, 2:317, "On sait combien cet esprit peu original aime se citer lui-même, reprendre, en les modifiant à peine, des expressions qu'il a déjà employées en d'autres écrits."

²⁶ See n. 6.

²⁷ Bernard, *Sermones de diversis*, 6, 7, 21, 62, 83, 100, and 104 (SBO 6/1).

²⁸ On the sermons, see Leclercq, "Les collections de sermons de Nicolas de Clairvaux," 57. Hugh's commentaries or chapters, the *Adnotationes elucidatoriae in quosdam Psalmos David*—the second book of the *Miscellanea*—are collected under PL 177:589.

²⁹ "Meo sensu inventos, meo stylo dictatos, nisi quod paucis in locis de sensibus alienis accipi" (my translation), Nicholas of Montiéramey, in the prefatory letter in MS Harley 3073, ed. Leclercq, *Recueil d'études*, 1:49–50.

Speculum 92/S1 (October 2017)

after his banishment from Clairvaux, Nicholas seemingly betrays his former abbot again with what appears to be a willful appropriation of Bernard's texts.

Such incriminating evidence contributed to his reputation as a plagiarist, this reputation in its turn provoking prejudicial conclusions in other attribution issues. Henri M. Rochais, for instance, in a codicological approach to the question of determining the disputed authorship of three other lengthy sermons in Bernard's *De diversis* corpus (*De diversis* 40, 41, and 42),³⁰ pointed out these sermons' close similarities to two of Nicholas's works and to chapter 100 of Hugh of St. Victor's sixth book of *Miscellanea* (the well-known writer somehow seems to be involved again);³¹ yet stated with confidence that Nicholas stole the texts from Bernard under false pretenses. This hypothesis Rochais sees corroborated in "the secretary's unscrupulous personality."³² At the same time Rochais casts aside Jean Mabillon's belief that the literary style of these sermons hardly seems that of Bernard as an all-too-subjective and unscientific argument.³³ To our view, Rochais' own subjective mistake was that—despite being fully aware of Bernard's collaboration with his secretaries—he treated codicological unity as identical to stylistic or authorial unity: "Cette tradition manuscrite ne donne donc aucun motif de doute sur l'authenticité bernardine des trois sermons étudiés, et, au contraire, elle constitue une telle probabilité en faveur de cette authenticité, qu'il faudrait des arguments incontestables pour dénier à Bernard leur composition."³⁴

Leclercq's and Rochais' attributions still stand in their editions, widely used today, although medievalists have seriously contested their highly subjective and speculative approach towards authorship attribution and their prejudiced view of Nicholas of Montiéramey's alleged deceitfulness and falsification, as we will show below. Moreover, the temptation for scholars to draw lines between imitation and plagiarism in order to categorize writings and collate them in attributed editions, valuable as it is, can also be rather anachronistic or even unbecoming in a medieval context. A fundamental rationale of the New Philology, in the wake of poststructuralist approaches to authorship and texts, is that in a medieval culture there is no place for the idea of an original author, a logic leading to the conclusion that Leclercq's and Rochais' quest for such an author only takes us further from the truth.³⁵ Medieval lit-

³⁰ Bernard, "Sermo 40: De viis vitae quae sunt confessio et oboedientia"; "Sermo 41: De via oboedientiae;" and "Sermo 42: De quinque negotiationibus, et quique regionibus," SBO 6/1:234–61.

³¹ The specific text referred to is Hugh of St. Victor, "De septem gradibus confessionis," PL 177: 856–58; see Henri M. Rochais, "Saint Bernard est-il l'auteur des sermons 40, 41 et 42 *De diversis*?" *Revue Bénédictine* 72 (1962): 324–45, at 326. There is a lack of clarity as to how exactly Hugh of St. Victor's *Miscellanea* was constituted—whether the collection was assembled by Hugh himself or whether it is a compilation assembled from his writings by others.

³² "Le caractère de ce secrétaire peu scrupuleux rend assez vraisemblable l'hypothèse d'un nouveau plagiat de Nicolas aux dépens de son ancien abbé," Rochais, *ibid.*

³³ "Dom J. Leclercq a dit justement ce qu'il faut penser de cette sorte d'argument trop subjectif pour avoir, à lui seul, une valeur réellement probante," Rochais, *ibid.*, 325. Jean Mabillon's argument for attributing the sermons to Nicholas can be found in a note to PL 183:647–48: "Hic sermo sequens in editione Lugdunensi anni 1514, in qua primum prodire, extra classem genuinorum Bernardi sermonum locati sunt; nec stylum ejus plene assequi videntur."

³⁴ Rochais, "Saint Bernard est-il l'auteur des sermons 40, 41 et 42?" 330.

³⁵ Nichols and Cerquiglini, two prominent figures of the New Philological approach, have already been cited above (see n. 2). The heritage of poststructuralists such as Roland Barthes, "The Death

erature depended strongly on its oral component (*dictare* had supplanted *scribere* to designate the act of authorship);³⁶ it “circulated through networks” and formed part of a “shared culture [that was] characterized by a knowledge of the same erudite language as well as a common foundation of texts and memories.”³⁷ Giles Constable touched on the core of the issue by noting that in the Middle Ages “there are infinite shadings between correction, revision, imitation, and falsification and, in works of art, between repair, restoration, reproduction, and copying.”³⁸ In this light, Nicholas’s appropriation of some of his former master’s works in his letter to Henry the Liberal is rather the continuation of a dialogue,³⁹ not a spiteful act of revenge. Stephen Jaeger has similarly argued that Nicholas indulges in the kind of *imitatio* that would have made little distinction between “honest” and dishonest intentions.⁴⁰ Like any distinguished writer of his time, Nicholas carefully applied for a new position by showcasing his complete immersion in a prevalent literary network.⁴¹

The juxtaposition of Leclercq’s and Rochais’ historical positivism with the more recent New Philology lays bare the dilemma that has arisen in medieval text studies. Although both practices have contributed immensely to the field, neither of the two stances is entirely satisfactory, leaving most scholars to agree to a compromise in cases of doubtful authorship. The first, rather positivist, approach acknowledges that the act of textual appropriation is suitable and possible. It presupposes that personal authorship is a retrievable aspect of the text, whose idealized state can be reconstructed from a hierarchical stemma. The disadvantage of this approach is that it

of the Author,” *Aspen Magazine* 5–6 (1967); and Michel Foucault, *Bulletin de la Société française de philosophie* (1969): 73–104; or in *Dits et écrits*, vol. 1, 1954–1988, ed. Daniel Defert, François Ewald, and Jacques Lagrange (Paris, 2001), 817–49, is apparent in this New Philological approach towards medieval authorship. Also see Virginie Greene, “What Happened to Medievalists after the Death of the Author?,” in *The Medieval Author in Medieval French Literature*, ed. Greene (New York, 2006): 205–27.

³⁶ Paul Saenger, “Silent Reading: Its Impact on Late Medieval Script and Society,” *Viator* 13 (1982): 380.

³⁷ Pascale Bourgain, “The Circulation of Texts in Manuscript Culture,” in *The Medieval Manuscript Book: Cultural Approaches*, ed. Michael Johnston and Michael Van Dussen (Cambridge, UK, 2015), 150. Also see Rebecca Moore Howard, *Standing in the Shadow of Giants: Plagiarists, Authors, Collaborators* (Stamford, CT, 1999), 65: “In the Middle Ages, mimesis was the means of establishing one’s authority, as well as being an expression of humility. The notion of the individual author, autonomous, original, and proprietary, played only the smallest role in this economy of authorship. With those textual values so much in decline, plagiarism was hardly an issue.”

³⁸ Giles Constable, “Forgery and Plagiarism in the Middle Ages,” in Constable, *Culture and Spirituality in Medieval Europe* (Aldershot, 1996), 1–41, at 3.

³⁹ Moreover, the assertion that Bernard never heard of Nicholas again after he left Clairvaux is far from certain: see Constable, *The Letters of Peter the Venerable*, 2:330.

⁴⁰ “Here then is a case in which a ‘skilled student of the *ars dictaminis*’ with alleged inclinations to forgery imitated a near-contemporary model, and we can assume that there would have been little difference between the ‘honest’ and dishonest imitation of Bernard’s style,” in Stephen Jaeger, “The Prologue to the *Historia calamitatum* and the ‘Authenticity Question,’” *Euphorion* 74 (1980): 13.

⁴¹ After all, Nicholas was applying for a position as Henry’s new secretary, and a familiarity with the greats of the twelfth century would have been one of the prerequisites. Leclercq’s assertion that Henry the Liberal must not have noticed Nicholas’s blatant plagiarism because he was a layperson unfamiliar with clerical texts seems unlikely: “Il était moins facile à un laïc comme Henri le Libéral qu’à un clerc de déceler le plagiat,” in Leclercq, *Recueil d’études*, 1:57. Henry’s recognition of the extent to which Nicholas’s compositions were indebted to other authorities would have been the whole point and is likely not to have been conceived of as problematic.

Speculum 92/S1 (October 2017)

reduces authorship debates to anachronistic, binary classification problems, whose conclusions often lack substantial evidence and set into motion more unsubstantiated debate. Therefore it both originates in bias and establishes a *circulus in probando* in which only additional factual evidence can provide closure. The second and more recent approach, on the other hand, comes to terms with the impossibility of closure through its recognition that—in a medieval context—knowing a text’s authorship was subordinate either, on the one hand, to acknowledging its implied authority (for example, the authority of a writer’s predecessors, such as the church fathers or God); or, on the other, to acknowledging the authority of its unique, materialized appearance: “There are as many texts . . . as there are scribal redactions; and there are as many authors of medieval texts as there were scribes composing new works in the act of writing them down in their [manuscripts].”⁴² By embracing the variance, this approach evades the impasse. Yet one is wary of where this might lead. Lena Wahlgren-Smith, who is preparing a critical edition of Nicholas of Montiéramey’s letters, has quite rightfully expressed her concern regarding a “wholesale adoption” of the New Philological approach, which “assumes that all medieval literature, in all languages, all genres, and all periods, operates in the same way.”⁴³ Such an approach is counterintuitive to those medieval attestations where value is attached to titled authority, where there is an outspoken preference for unviolated text, or where personal literary style is cultivated.⁴⁴ Bernard’s denunciation of Nicholas for sending out texts without his consent serves as a firsthand example. Correspondingly, Nicholas’s bold statement that Bernard’s texts are in fact his own—“meo sensu inventos, meo stylo dictatos”⁴⁵—also suggests that an explicit appropriation of texts was not unknown in the twelfth century. It is important not to underestimate the degree to which the Middle Ages was a “charismatic culture” in which texts were regarded in relation to “the body and the physical presence [that were] the mediators of cultural values.”⁴⁶ From this perspective, Leclercq and Rochais had justifiable reasons to care about the interdependence of text and physical author (or performer). Constable has referred to a “rising tide of concern” over textual theft in the late twelfth and thirteenth century, possibly instigated by rapidly changing approaches to “literary individuality.”⁴⁷ In the midst of such an impasse, historians and

⁴² Bernadette A. Masters, “The Distribution, Destruction and Dislocation of Authority in Medieval Literature and Its Modern Derivatives,” *Romanic Review* 82 (1991): 270–285, at 278.

⁴³ Lena Wahlgren-Smith, “Editing a Medieval Text: The Case of Nicholas of Clairvaux,” in *Challenging the Boundaries of Medieval History: The Legacy of Timothy Reuter*, ed. Patricia Skinner, *Studies in the Early Middle Ages* 22 (Turnhout, 2009), 173–83. For the upcoming critical edition, see Wahlgren-Smith, *The Letter Collections of Nicholas of Clairvaux* (Oxford, forthcoming).

⁴⁴ Minnis, *Medieval Theory of Authorship*, 10–11.

⁴⁵ Nicholas of Montiéramey, in the prefatory letter in MS Harley 3073, ed. Leclercq, *Recueil d’études*, 1:49–50.

⁴⁶ Jaeger, “Charismatic Body, Charismatic Text,” 122.

⁴⁷ Constable, “Forgery and Plagiarism,” 18, 32. He also noted that Nicholas “may have inspired the Cistercian legislation of 1157 defining the punishments for the falsifiers of charters and seals,” 19. However, the idea of literary individuality in the twelfth-century Renaissance that Constable here addresses is a subject of immense debate: see Caroline Walker Bynum, “Did the Twelfth Century Discover the Individual?,” *Journal of Ecclesiastical History* 31 (1980): 1–17; Colin Morris, *The Discovery of the Individual, 1050–1200* (New York, 1972); and Aron Iakovlevič Gurevič and Katharine Judelson, *The Origins of European Individualism* (Oxford, 1995).

philologists should hope to find more solid ground on which to establish the authorship of dubious texts. Authorship attribution—a seemingly trivial question concerning who wrote which text—forms a vital stepping-stone towards a more scientifically responsible understanding of how the medieval author perceived the proportional relationship between one’s personal authority over a text and one’s personal contribution to its composition.

COMPUTATIONAL STYLISTICS

Instead of treating the medieval text as collective and impersonal, the methodology of computational stylistics traces the linguistic traits of a text that originate from a highly individual stylome, a set of features that betray personal writing style (note here that mapping out individual stylistics should not eliminate the possibility of collective authorship).⁴⁸ Moreover, its ambitions—whether realistic or not—may extend to extracting information concerning the author’s sex or age⁴⁹ or to measuring a text’s genre or degree of “literariness.”⁵⁰ Although experiments and debates as to which textual features best capture stylistic difference are still ongoing, many state-of-the-art studies employ function words, which still prove to be the most robust discriminators for writing styles. Function words are usually short and insignificant words that pass unnoticed—such as pronouns, auxiliary verbs, articles, conjunctions, and particles—whose main advantages are their frequent occurrence, their less conscious use by authors, and their content- or genre-independent character. Their benefit and success for the study of stylometry in Latin prose have been convincingly demonstrated before,⁵¹ although the methodology still raises acute

⁴⁸ This is the so-called human stylome hypothesis: see Hans van Halteren et al., “New Machine Learning Methods Demonstrate the Existence of a Human Stylome,” *Journal of Quantitative Linguistics* 12 (2005): 65–77.

⁴⁹ For author-profiling studies, see Shlomo Argamon et al., “Automatically Profiling the Author of an Anonymous Text,” *Communications of the Association for Computing Machinery (ACM)* 52 (2009): 119–23.

⁵⁰ Karina van Dalen-Oskam, “A Literary Rat Race,” in *Digital Humanities 2016: Conference Abstracts* (Kraków, 2016), 388–90.

⁵¹ See Mike Kestemont, Sara Moens, and Jeroen Deploige, “Collaborative Authorship in the Twelfth Century: A Stylometric Study of Hildegard of Bingen and Guibert of Gembloux,” *Digital Scholarship in the Humanities* 30 (2013): 199–224; and Jeroen Deploige and Sara Moens, “Visiones Hildegardis a Guiberto Gemblacensi exaratae,” in *Hildegardis Bingensis Opera minora*, vol. 2, ed. Jeroen Deploige et al., CCCM 226A (Turnhout, 2016), 153–61; or the numerous investigations on the authorship of the *Scriptores Historiae Augustae*, of which the latest is Mike Kestemont and Justin A. Stover, “The Authorship of the ‘Historia Augusta’: Two New Computational Studies,” *Bulletin of the Institute of Classical Studies of the University of London* 59 (2016): 140–57. Also see Penelope J. Gurney and Lyman W. Gurney, “Authorship Attribution of the *Scriptores Historiae Augustae*,” *Literary and Linguistic Computing* 13 (1998): 119–31; Richard S. Forsyth, David I. Holmes, and Emily K. Tse, “Cicero, Sigonio, and Burrows: Investigating the Authenticity of the *Consolatio*,” *Literary and Linguistic Computing* 14 (1999): 1–26; Fiona J. Tweedie, David I. Holmes, and Thomas N. Corns, “The Provenance of *De doctrina Christiana*, attributed to John Milton: A Statistical Investigation,” *Literary and Linguistic Computing* 13 (1998): 77–87; and Earl Jeffrey Richards, David Joseph Wrisley, and Liliane Dulac, “The Different Styles of Christine de Pizan: An Initial Stylometric Analysis,” *Le Moyen Français* 78–79 (2016): 187–206.

questions, which keep stylometrists on the lookout for alternatives.⁵² Before returning to our case study of Nicholas of Montiéramey and his alleged falsification, we will expound on the availability of the data, on the preprocessing steps involved, and on the statistical technicalities of computational stylistics.

For our subsequent analysis, we relied upon the digitized texts of Bernard of Clairvaux's *Corpus epistolarum*, *Sermones de diversis*, and *Sermones super Cantica canticorum* as they appear in the state-of-the-art scholarly edition of the *Sancti Bernardi Opera* by Leclercq et al. included in the online Brepols Library of Latin Texts.⁵³ For Nicholas of Montiéramey's letters we are provisionally still reliant on the digitally available *Patrologia Latina*.⁵⁴ All text data are available in an online GitHub repository for experimental replication, yet in a camouflaged form so that the copyright protection on the original text editions is respected. Only the texts' function words were retained in their original form, whereas all content-loaded words were filtered out and replaced by dummy words.⁵⁵ Since Leclercq's editions and the *Patrologia Latina* make use of different orthographical conventions, and since Latin is a synthetic language with a high degree of inflection, Bernard and Nicholas's texts required some preprocessing for the sake of data alignment and feature culling. The result of the latter is that texts are more easily mined for information: thus, the lexemes are lemmatized (which means that a specific instance of the word is referred to its headword) and a text's words (tokens) are classified according to grammatical categories (parts of speech). For this purpose we applied the Pandora lemmatizer tagger on the texts, a piece of software developed to achieve specifically this.⁵⁶

Token	Lemma	PoS-tag (<i>simplex</i>)
harum	hic	PRO
imo	immo	ADV

An advantage of Pandora's design is that it normalizes orthographical variants to a classicized headword if necessary. Both *imo* and *immo* were categorized under

⁵² Mike Kestemont, "Function Words in Authorship Attribution: From Black Magic to Theory?," in *Proceedings of the 3rd Workshop on Computational Linguistics for Literature (CLL) at the 14th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, ed. Anna Feldman, Anna Kazantseva, and Stan Szpakowicz (Gothenburg, 2014), 59–66.

⁵³ See n. 7 for details on the edition of Bernard of Clairvaux's texts by Leclercq et al. The digitized text files of these editions have been generously provided for our experiments by our project partner, Brepols Publishers. For Brepols's online Library of Latin Texts, see <http://www.brepols.net>.

⁵⁴ PL 196:1651a–1654b. The original edition of Nicholas's letters was first published in 1610 by Jean Picard, who was believed to have had access to the original manuscripts. The text would later be republished in Lyon in 1677 and ultimately reprinted by Jacques-Paul Migne in the *Patrologia Latina*: see Luanne Meagher, "The Letters of Nicolas of Clairvaux," in *Heaven on Earth*, *Studies in Medieval Cistercian History* 9, *Cistercian Studies Series* 68, ed. Ellen Rozanne Elder (Kalamazoo, 1983), 128. The forthcoming critical edition is edited by Wahlgren-Smith: see n. 43.

⁵⁵ See <https://github.com/jedgusse/bernard>.

⁵⁶ Pandora was developed by Mike Kestemont and the author of this article. See Mike Kestemont and Jeroen De Gussem, "Integrated Sequence Tagging for Medieval Latin Using Deep Representation Learning," *Journal of Data Mining and Digital Humanities*. Special Issue on Computer-Aided Processing of Intertextuality in Ancient Languages, ed. Marco Büchler and Laurence Mellerin (2017), <https://arxiv.org/pdf/1603.01597v2.pdf>.

immo and identified accordingly as one and the same word.⁵⁷ The part-of-speech-tag (PoS) displayed in the third column in the above diagram allowed us to restrict the culling of the most frequent words to those word categories that make up the collection of function words: conjunctions (CON), prepositions (AP), pronouns (PRO), and adverbs (ADV). This likewise filtered out some noise caused by ambiguities or homonyms like *secundum*, which can be either a preposition or the accusative case of the adjective *secundus*. Afterwards, some lemmata in the list that did not qualify as style markers were culled and filtered out, such as *tu*, *tuus*, *vos*, and *vester*. *Vos* and *vester* betray a vernacular influence in Bernard's unpublished letters as formal, polite forms similar to the French *vous*. Bernard was known to adapt his sermons to his audience,⁵⁸ and in a literary or classicizing text he would maintain *tu* and *tuus* when addressing his correspondent. Therefore these pronouns were not regarded as suitable features for stylistic difference but as content-dependent, conscious authorial choices linked to register.⁵⁹ Aside from lemmatization, smaller interventions were undertaken, such as separating the enclitics *-que* and *-ve* from the token in order to be recognized as a feature. Once procedures of this sort were carried out in full, we arrived at a list of the 150 most frequent function words (MFFW) of the corpus examined in our experiment.⁶⁰ Tables 1 and 2 of the most frequent function words correspond to the two experiments (and their two respective corpora) described in this article and are listed in the appendix.

The corpora under scrutiny were subsequently segmented into parts with a fixed size, in other words, text samples. Sampling yields the advantage of “effectively [assessing] the internal stylistic coherence of works,”⁶¹ as it also allows for a more fine-grained comparative analysis with segments from external works. The sample sizes, however, can differ depending on the requirements of the experiment. As will become apparent in the appendices and the figures, Bernard's letters were segmented into 3,000-word samples, whereas his sermons were segmented into 1,500-word samples. The decrease of sample size in the second experiment was necessary due to the fact that it treats shorter texts. It should be noted that whereas 3,000-word

⁵⁷ Other pairs include *tanquam* and *tamquam*, *quoties* and *quotiens*, *numquid* and *numquid*, *quamquam* and *quamquam*, *nunquam* and *numquam*, etc.

⁵⁸ Constable, “The Language of Preaching in the Twelfth Century,” 131–52.

⁵⁹ See Leclercq, “Notes sur la tradition des épîtres de S. Bernard,” *Recueil d'études*, 3:317. It should be noted that our decision to disregard these features, or in other words to succumb to manual feature selection, implies a degree of supervision and subjectivity. We do not really see this as a problem, firstly since we have supplied evidence that including these features could only distort the results for obvious historical reasons, and secondly since this is in line with our approach that was already strongly determined by how we set limitations to the culling of function words (only prepositions, conjunctions, adverbs, and pronouns were taken into consideration). The omission of too-characteristic corpus features considerably improves precision and historical validity. David L. Hoover has demonstrated this for personal pronouns as well: see Hoover, “Testing Burrows's Delta,” *Literary and Linguistic Computing* 19 (2004): 453–75.

⁶⁰ On the culling of the most frequent words, see the pivotal work of John F. Burrows, “‘Delta’: A Measure of Stylistic Difference and a Guide to Likely Authorship,” *Literary and Linguistic Computing* 17 (2002): 267–87. Its workings have been considerably elucidated (and its formula simplified) by the publication of Shlomo Argamon, “Interpreting Burrows's Delta: Geometric and Probabilistic Foundations,” *Literary and Linguistic Computing* 23 (2008): 131–47.

⁶¹ Maciej Eder, Jan Rybicki, and Mike Kestemont, “Stylometry with R: A Package for Computational Text Analysis,” *R Journal* 16 (2016): 107–21, at 111.

samples correspond to a state-of-the-art norm, 1,500-word samples run the risk of increased imprecision, a consideration that should nuance any interpretation of the results.⁶² Once we divided our corpus, each of the text samples needed to be translated to a format that is also readable to computers, namely document vectors. A text sample is represented as an array of tallies for each of the one hundred and fifty function words on the checklist, as follows:

	et	in	qui	...
Sample 1	118	38	71	...
Sample 2	82	47	90	...
Sample n

These raw counts were TF-IDF normalized, a procedure that divides the function word frequencies by the number of text samples that respective function word appears in. As a consequence, less common function words received a higher weight, which prevents them from sinking away (and losing statistical significance) in between very common function words.⁶³ Once the data was preprocessed and regulated, two statistical techniques were applied to visualize its dynamics.

The first is k Nearest Neighbors (hereafter k -NN); the second is principal component analysis (hereafter PCA). Their respective results will prove to be similar in a general sense, yet crucially different in the details. We argue that such an additional statistical validation provides for a more accurate, nuanced interpretation and a better intuition of the data. In Figs. 1 and 3, the k -NN networks, we first calculated the five closest text samples to each text sample by applying k -NN on the frequency vectors.⁶⁴ Accordingly, for each text the five most similar, or closest, texts were

⁶² There is debate over the adequate sample length to capture a stylistic signal. The risk has been addressed for twelfth-century Latin in Kestemont, Moens, and Deploige, "Collaborative Authorship in the Twelfth Century," 210–11. Also see Maciej Eder, "Does Size Matter? Authorship Attribution, Small Samples, Big Problem," *Literary and Linguistic Computing* 30 (2015): 167–82; and Kim Luyckx and Walter Daelemans, "The Effect of Author Set Size and Data Size in Authorship Attribution," *Literary and Linguistic Computing* 26 (2011): 35–55.

⁶³ The TF-IDF vectorizer is therefore a normalization procedure that penalizes more frequent function words in favor of rare function words: see Christopher D. Manning et al., *Introduction to Information Retrieval* (New York, 2008), 117–33. The TF-IDF vectorizer is not an obvious choice for feature extraction in authorship-attribution studies, since it readjusts Burrows's Delta presupposition (see above) that the high-ranked most frequent words generate better distinctions between authors. We argue that observing less common function words can yield interesting authorial preferences, which might go unnoticed in the standard Delta approach. "In many ways, this model can be contrasted with the assumption that low-frequency items are bad predictors of authorial style. Nevertheless, a few studies suggest that it might be useful. Arguably, this model captures the intuition that if a highly rare feature is present in two documents, this increases the likelihood that the two documents were authored by the same individual. While the method might therefore be sensitive to overfitting on low-frequency properties, this might be an attractive characteristic in certain (e.g., single-domain) authorship problems," Kestemont et al., "Authenticating the Writings of Julius Caesar," *Expert Systems with Applications* 63 (2016): 89.

⁶⁴ The similarity metric applied for the pairwise distances is the Minkowski metric, a Euclidean metric that is "a very general metric that can be used in a k -NN classifier for any data that is represented as a feature vector": see Pádraig Cunningham and Sarah Jane Delany, "k-Nearest Neighbour Classifiers," *Multiple Classifier Systems* 34 (2007): 1–17, at 4.

calculated, weighted in rank of smallest pairwise distance⁶⁵ and consequently mapped in space through force-directed graph drawing.⁶⁶ What a k -NN network ultimately captures is which texts are most akin—or have the closest connection—when it comes to writing style (as defined by the distribution of function words). It should be noted that k -NN nearly *always* finds relationships, as it is very much a closed game. It is designed to link candidates to one another in terms of distance (every text sample *needs* to find its five neighbors) and can presuppose ties that are rather coincidental or nonexistent (for example, in the case of outliers). The network visualization can therefore be biased by a misleading directionality.

Secondly, PCA is a technique that allows us to reduce a multivariate or multidimensional data set of many features, such as our function word frequencies, to merely two or three principal components, which disregard inconsequential information, or noise, in the data set and reveal its important dynamics (Figs. 2 and 4). The assumption is that the main principal components, our axes in the plot, point in the direction of the most significant change in our data, so that clustering and outliers become clearly visible. Each word in our feature vector is assigned a weighting, or loading, which reflects whether or not a word correlates highly with a PC and therefore gains importance as a discriminator in writing style. In a plot, the loadings or function words that overlap with the clustered texts of a particular author are the preferred function words of that author (see Figs. 5–6).⁶⁷ PCA is built to find the most meaningful variance of observations along the axes of its principal components. In this sense it is not always interested in finding links between candidates, as k -NN is, but rather in finding links between variables. Disadvantages are that

⁶⁵ The weights were derived directly from the calculated distances (see n. 64 for specifications on the metric). The intuition is then that the distances should be normalized to a (1,0) range. Note that this is not a (0,1) range, since smaller distances correspond to greater similarities and therefore require greater weighting: $distances = \frac{distances - \min(distances)}{\min(distances) - \max(distances)}$.

⁶⁶ See, for instance, Maciej Eder, “Visualization in Stylometry: Cluster Analysis Using Networks,” *Digital Scholarship in the Humanities* 32 (2015): 50–64. The algorithm used for the graphs in this article was Force Atlas 2, embedded in Gephi, an open-source tool for network manipulation and visualization: see Mathieu Bastian et al., “Gephi: An Open Source Software for Exploring and Manipulating Networks,” *Proceedings of the Third International Conference on Weblogs and Social Media (ICWSM)*, ed. Eytan Adar et al. (San Jose, CA, 2009): 361–62. It must be noted that Eder’s network iteratively runs over an increasing number of features (100–1000 MFW) to establish the consensus between different text samples (likewise by means of nearest neighbors). I have somewhat adjusted the method for the purpose of this article, since running up to 1000 MFW would surely overfit on the strong content-dependent connections that exist between Nicholas and Bernard’s texts. Matthew Lee Jockers likewise applied Gephi networks to detect stylistic differences in his seminal work *Macroanalysis: Digital Methods and Literary History* (Champaign, 2013). The difference with his approach and the one maintained in this article is that Jockers not only generated his networks through stylistic variables, but combined these with thematic linkages to discover trends of literary evolution on a far larger and diachronic scale than in this article (one could say that Jockers, as a quantitative formalist, demonstrated the popular formalist concept of the “defamiliarization” of literary language). As argued earlier, thematic variables should be disregarded in this case study. Nevertheless, I concur with Jockers’s view that networks are a powerful tool to “demonstrate literary imitation, intertextuality, and influence” (156).

⁶⁷ For an elaborate explanation of PCA and its applicability to stylometry, see José Nilo G. Binongo and M. Wilfrid A. Smith, “The Application of Principal Components Analysis to Stylometry,” *Literary and Linguistic Computing* 14 (1999): 446–66. The PCA plots were drawn with the Matplotlib package available for Python: see John D. Hunter, “Matplotlib: A 2D Graphics Environment,” *Computing in Science and Engineering* 9 (2007): 90–95.

PCA can never explain all the variance of the data, since it purposefully disregards many features and dimensions that it finds insignificant. It also has the tendency to produce somewhat nebulous scatter plots when texts are stylistically entangled (as is the case for Bernard and Nicholas).

THE LETTERS

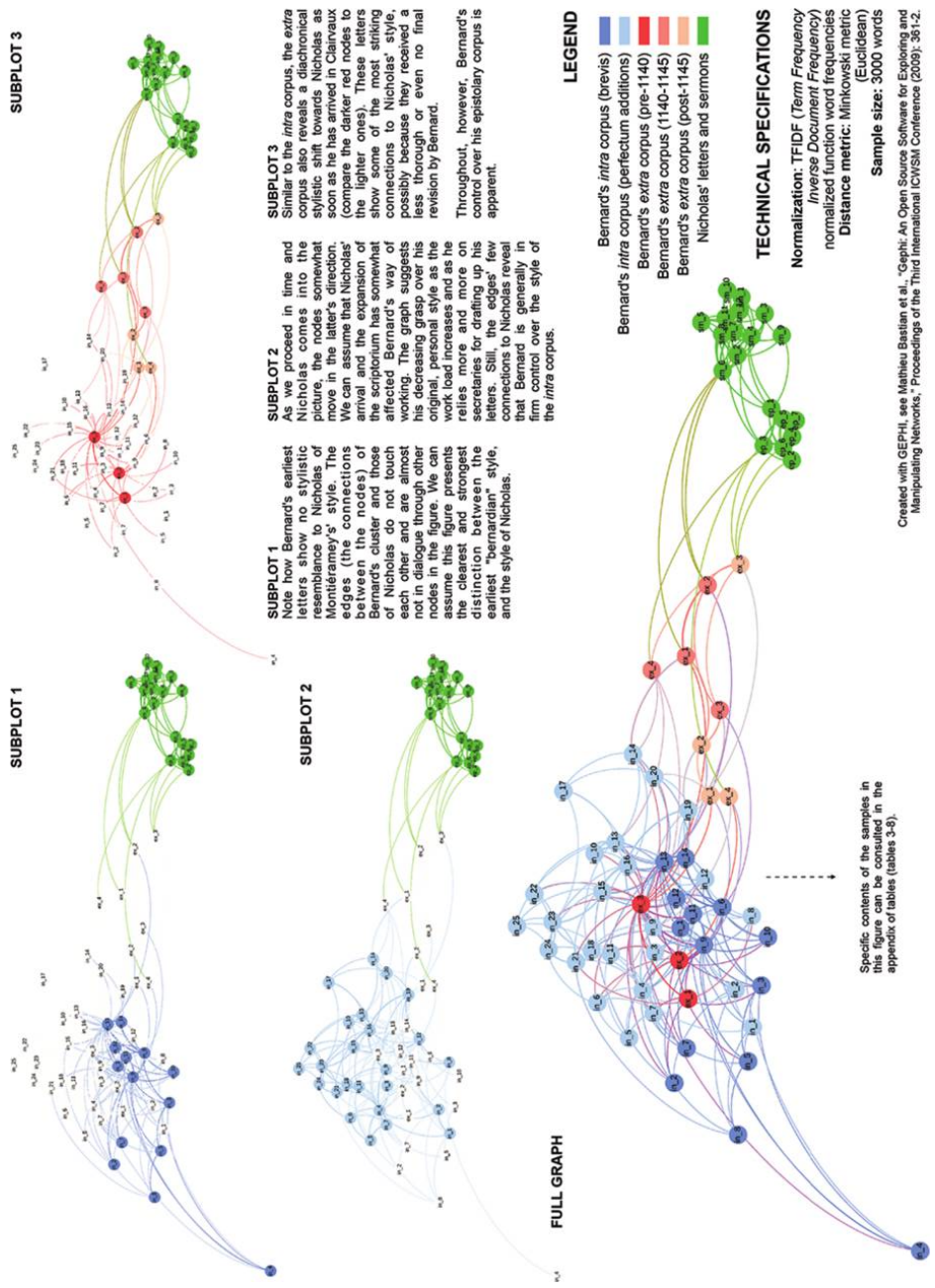
Bernard's epistolary corpus is very complex, but a coherent structure has been recognized thanks to Jean Leclercq's editorial achievements. Leclercq, whose terminology will be adopted here, has divided the corpus into a literary (*intra* corpus) and a nonliterary section (*extra* corpus).⁶⁸ Bernard intended the letters in the *intra* corpus to circulate as a literary collection and kept refining them intensely throughout his life, whereas the second group of letters is scattered across time and manuscript traditions.⁶⁹ Then, within the first section, the literary or *intra* corpus, we can make another division. Manuscript transmission allows us to distinguish between letters written before Nicholas's arrival in the scriptorium and letters inserted later.⁷⁰ The letters that date from before 1145 in an earlier first appearance are found in the *brevis* manuscripts, whereas those added to Bernard's literary corpus afterwards can be found in the *perfectum* manuscripts. The *perfectum* corpus was assembled after Bernard's death in 1153, possibly by Geoffrey of Auxerre, and contains, aside from the *brevis* letters, many new additions (Tables 3 and 4 give a detailed overview of which letters are included in either the *brevis* or *perfectum* samples).⁷¹ It is impor-

⁶⁸ Leclercq separated the *intra* corpus (Epp. 1–310) from the *extra* corpus (Epp. 311–547) in the SBO.

⁶⁹ Leclercq, SBO 8:233–38.

⁷⁰ Leclercq, "Lettres de S. Bernard: Histoire où littérature?," *Recueil d'études*, 4:158.

⁷¹ The main difference between the *brevis* and *perfectum* manuscripts is that the latter contain versions of these letters that were clearly amended and lengthened. In the introduction to the edition of the *intra* corpus, Leclercq gives a full account of the arrangement of the different transmissions, in which he distinguishes three more or less homogeneous collections, two of which are the *brevis* and *perfectum* cycles, whose names have served as inspiration to how we labeled our chronologically ordered data. It should be noted that Leclercq mentions a third intermediary publication that we have decided to exclude from our main argument, namely the *longior* corpus, which was compiled by Geoffrey of Auxerre and was presumably published in 1145. The *longior* corpus already contains quite a few of the *perfectum* additions. However, this corpus is hard to date or reconstruct, making it less interesting for us to include in this study: see Leclercq, SBO 7:xv. We decided to make a distinction between the early *brevis* publication, when Nicholas of Montiéramey was certainly absent from Clairvaux, and the later publication, when both of them, or an even more developed chancery, could have exerted influence on Bernard's style. More specifically, the letters falling under the heading of *brevis*—written before 1145 and not to be confused with Leclercq's *brevis* manuscript collection—can be consulted under the following indices: Epp. 1, 2, 4, 8, 11, 12, 24, 25, 42, 65, 67–70, 72, 73, 78, 79, 82, 83, 85, 87–89, 91, 95, 96, 98, 102, 104, 106, 107, 111, 113, 114, 117–19, 124–27, 129–33, 136, 139, 141, 143, 150, 152, 156, 158, 159, 168, 169, 178, 212, and 254. The letters that fall under the heading of *perfectum*—not to be confused with the *perfectum* cycle as it is described by Leclercq, but corresponding to all letters that were written after 1145 and remained unmentioned in the *brevis* data—are to be found under these indices: Epp. 3, 5–7, 9, 10, 13–23, 26–41, 43–64, 66, 71, 74–77, 80, 81, 84, 86, 90, 92–94, 97, 99–101, 103, 105, 108–10, 112, 115, 116, 120–23, 128, 134, 135, 137, 138, 140, 142, 144–49, 151, 153–55, 157, 160–67, 170–77, 179–211, 213–53, and 255–310. However, these letters were categorized according to their date of publication (*brevis* and *perfectum*) and split up into samples. A full inventory of which letters can be found under which sample in the figures is provided in the Appendix of Tables (see Tables 3–7).



SUBPLOT 1
 Note how Bernard's earliest letters show no stylistic resemblance to Nicholas of Montieramey's style. The edges (the connections) of Bernard's cluster and those of Nicholas do not touch each other and are almost not in dialogue through other letters. The figure assumes the clearest and strongest distinction between the earliest "bernardian" style, and the style of Nicholas.

SUBPLOT 2
 As we proceed in time and Nicholas comes into the picture, the nodes somewhat move in the latter's direction. We can assume that Nicholas' arrival and the expansion of the scriptorium has somewhat affected Bernard's way of working. The graph suggests his decreasing grasp over his letters as time passes and his work load increases and as he relies more and more on his secretaries for drafting up his letters. Still, the edges' few connections to Nicholas reveal that Bernard is generally in firm control over the style of the *intra* corpus.

SUBPLOT 3
 Similar to the *intra* corpus, the *extra* corpus also reveals a diachronical stylistic shift towards Nicholas as soon as he has arrived in Clairvaux (compare the darker red nodes to the lighter ones). These letters show some of the most striking connections to Nicholas' style, possibly because they received a less thorough or even no final revision by Bernard.

Throughout, however, Bernard's control over his epistolary corpus is apparent.

LEGEND

- Bernard's *intra* corpus (brevis)
- Bernard's *intra* corpus (perfectum additions)
- Bernard's *extra* corpus (pre-1140)
- Bernard's *extra* corpus (post-1145)
- Nicholas' letters and sermons

TECHNICAL SPECIFICATIONS

Normalization: TFIDF (Term Frequency Inverse Document Frequency) normalized function word frequencies
 Distance metric: Minkowski metric (Euclidean)
 Sample size: 3000 words

Specific contents of the samples in the figure are reported in the appendix of tables (tables 3-8).

Created with GEPHI, see Mathieu Bastian et al., "Gephi: An Open Source Software for Exploring and Manipulating Networks.", Proceedings of the Third International ICWSM Conference (2009): 361-2.

Fig. 1. Network visualization of Bernard of Clairvaux' epistolary corpus compared to Nicholas of Montieramey's letters and sermons.

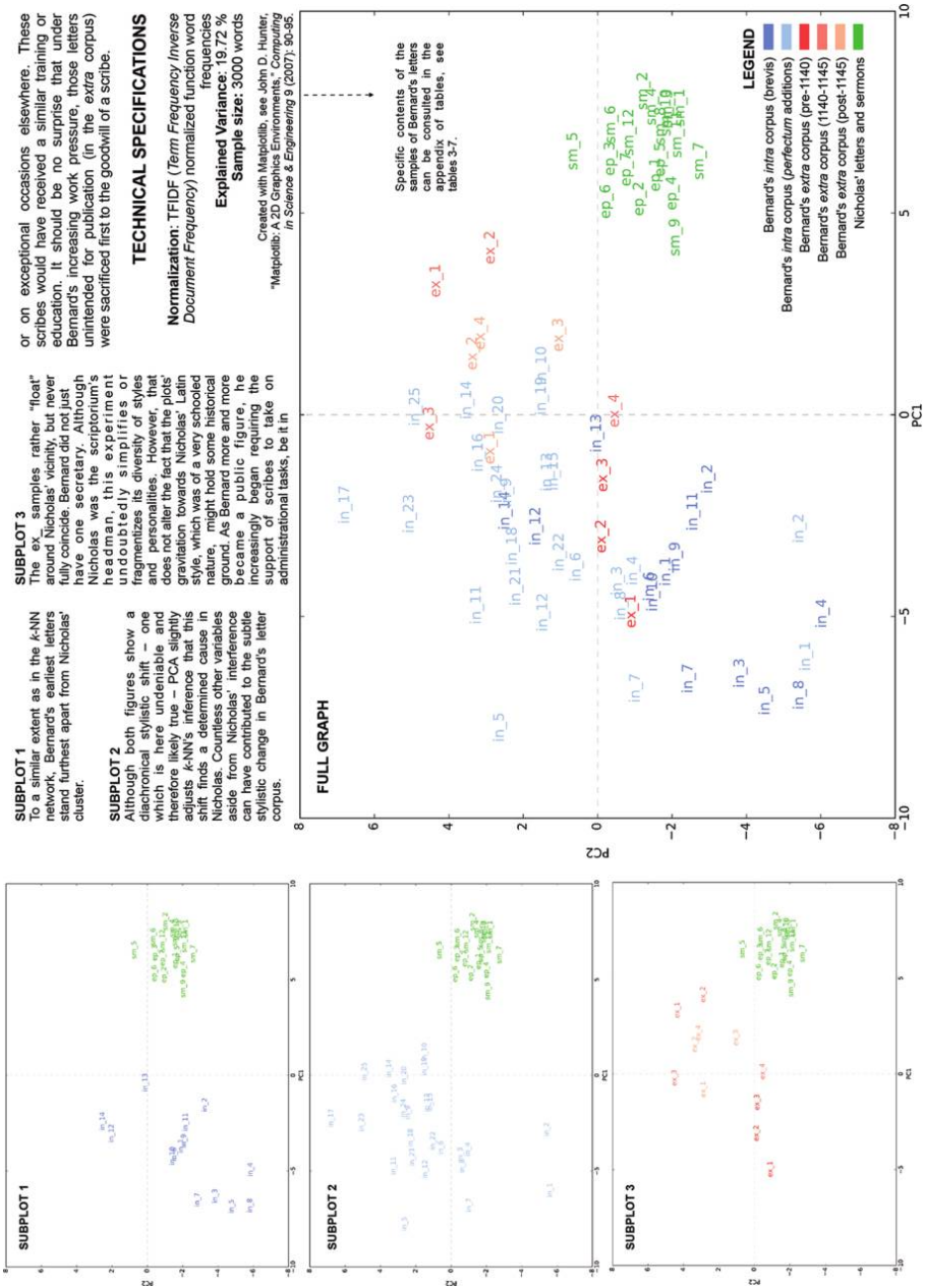


Fig. 2. Principal Components Analysis of Bernard of Clairvaux' epistolary corpus compared to Nicholas of Montieramey's letters and sermons.

tant to note that those letters which were already found in the *brevis* manuscripts and reoccur in the *perfectum* corpus have sometimes considerably changed in the eight years between the two appearances. After all, Bernard's aim was to compose a unified piece of literature. He corrected, rearranged, and selected throughout his life. Importantly, Leclercq's edition of Bernard's literary letter corpus, which we use in these experiments, is almost entirely based on the *perfectum* transmission, which enjoyed the most popular circulation.⁷² We therefore do not work strictly with the *brevis* corpus in its original pre-1145 form, but with a group of letters that was collectively reworked and jointly disseminated. This condition of the texts is a flaw in the experiment that should be kept in mind during the analysis. Moreover, as mentioned earlier, Bernard's *extra* letters have not known a homogeneous transmission but have been handed down to us under divergent circumstances. The corpus therefore required some reorganization. Since Leclercq's edition allows us to assign individual letters to discrete periods, we decided to divide the *extra* corpus into three time-bound parts to see if Nicholas's arrival came with a stylistic impact: the first part dates from before 1140, the second between 1140 and 1145, and the third from 1145 onwards. Those *extra* letters that are of questionable dating and addressee have been left out of our experiments, for they cannot contribute to a study of Bernard's stylistic evolution through the influence of his secretaries (Table 5 in the appendix gives a full overview of which *extra* letters were included or excluded).

In Fig. 1 (*k*-NN) and Fig. 2 (PCA) we have applied the statistical methods described above to calculate and visualize the stylistic differences between Bernard's letter corpus and Nicholas's authentic sermons and letters.⁷³ For each of these techniques, we have provided three additional subplots, which highlight how the different corpora are positioned within the clusters. There appear to be two general, observable dynamics, confirmed both by *k*-NN and PCA. Firstly, the writing style in Bernard of Clairvaux's letters is fairly coherent and forms a distinguishable cluster separated from Nicholas's works. Nevertheless—and this is the second, more hidden dynamic—our chronological and codicological rearrangements in the corpus have laid bare a gradual, subtle disturbance in Bernard's stylistic signal from 1140 onwards, corresponding to the approximate time of Nicholas's arrival in Clairvaux and seemingly moving towards the latter's cluster.⁷⁴ Yet, two major remarks are in order. Firstly, although the *perfectum* additions were indeed inserted into the literary corpus from 1140 onwards, some of them must have been first composed at a time before Nicholas's arrival. For example, sample *in_10* of the *perfectum* additions, which draws closest to Nicholas's cluster of all literary samples (only in the PCA, not in the *k*-NN network), contains letters that revolve around the schism between Antipope Anacletus and Pope Innocentius II, a series of events that occurred between 1130 and 1138.⁷⁵ Although Nicholas was not yet part of Bernard's entourage during these events and was therefore likely not involved in their first redac-

⁷² Leclercq, SBO 7:xvi.

⁷³ See n. 6 for a listing of the sermons under consideration. For Nicholas's letters, see PL 196:1593–1654.

⁷⁴ Turcan-Verkerk, "L'introduction de l'*ars dictaminis* en France," 70.

⁷⁵ The sample (*in_10*) contains Epp. 128, 134, 135, 137, 138, 140, 142, 144, 145, 146, 147 (SBO 7:321–51).

tion, he was nevertheless present when they were first sent out collectively with the other *perfectum* letters. However, if any refinement was imposed on these letters after 1140, it seems likely that Bernard, as the author, would have been the one to do so, not Nicholas, although technically the latter's interference is possible.

The second remark ties in to the problem we have just raised. Although both figures show a diachronic stylistic shift, PCA slightly adjusts *k*-NN's inference that this shift has a determined direction towards Nicholas. The *ex_* samples rather "float" around Nicholas's vicinity but never fully coincide. This suggests that the disturbances in Bernard's stylistic signal should not necessarily be as "monocausal" or "directional" as the *k*-NN network suggests, a nuance that reciprocates the historical skepticism raised in our first remark. Countless other variables aside from Nicholas's interference could have contributed to the subtle stylistic change in Bernard's letter corpus. One factor could be the lapse of time and Bernard's personal development.⁷⁶ Another is the respective corpus's divergent transmission history. But perhaps the most crucial reason for PCA's less outspoken directionality is that Bernard did not have just one secretary. Although Nicholas was the scriptorium's headman, this experiment undoubtedly simplifies or fragmentizes its diversity of styles and personalities. We might even be surprised that Bernard's letters—considering the circumstances under which they were conceived—still display this amount of stylistic coherence (although there might have been a more outspoken divergence if we had been able to oppose the very original *brevis* corpus to the published versions).

This does not alter the fact that the plots' gravitation towards Nicholas's Latin style, which was of a very schooled nature,⁷⁷ might hold some historical ground. As Bernard more and more became a public figure, he increasingly began requiring the

⁷⁶There is a considerable amount of literature that argues that such an evolution of personal style through time can be captured computationally by applying so-called stylochronometrical methods. For a concise overview of this subfield in computational stylistics, see Constantina Stamou, "Stylochronometry: Stylistic Development, Sequence of Composition, and Relative Dating," *Literary and Linguistic Computing* 23 (2008): 181–99.

⁷⁷Nicholas's style has often been deemed schooled and unoriginal: see Constable, *The Letters of Peter the Venerable*, 2:328; and Leclercq, "Les collections de sermons de Nicolas de Clairvaux," 1:55: "Ses exposés superficiels se développent selon un plan scolaire, en un style artificiel." Likewise, Dorette Sabersky argued that "the syntactical structure of his sentences is similar to Bernard's, but often clumsier, less clear, less elegant, and rhythmically less balanced. His frequent use of word plays is at times rather superficial and, in opposition to Bernard's use, of little importance to the development of the contents. Repetitions of certain phrases and topics occur every so often. He favors rather unusual words and likes to quote classical authors. His literary exertions are only too obvious. All these aspects evidence Nicholas' lack of Bernard's creative spontaneity and mastery of language," in "The Style of Nicholas of Clairvaux's Letters," in *Erudition at God's Service*, Studies in Medieval Cistercian History 11, Cistercian Studies Series 98, ed. John R. Sommerfeldt (Kalamazoo, 1987), 196. However, it is all the more peculiar and contradictory to these former statements that—even until very recently—attempts were made to attribute texts to Nicholas on the grounds of phrasing tics and certain lexical preferences (e.g., neologisms): see Patricia Stirnemann and Dominique Poirel, "Nicolas de Montiéramey, Jean de Salisbury et deux florilèges d'auteurs antiques," *Revue d'histoire des textes* 1 (2006): 173–88. Either such attribution methods should be challenged (perhaps rightly so; their word and phrase concordances form particularly dangerous grounds for attributing authorship in a twelfth-century context that boasts such a high degree of "plagiarism") or the statement that Nicholas has no style of his own should be withdrawn. I am convinced of the latter. Our computational experiments show that Nicholas, despite being an imitator, has a very controlled and rather clean authorial signal.

support of scribes to take on administrative tasks, be it in Clairvaux or on exceptional occasions elsewhere.⁷⁸ These scribes would have received a similar training or education. We can assume that most were under Nicholas's supervision, which meant that they departed from a common framework or set of rules from which they set out to imitate Bernard. This had become the nature of the epistolary writing art, or *ars dictaminis*.⁷⁹ Letters were constructed on the basis of similar formulas, abounded in clever wordplay, and the rhythms of their prose pulsed under comparable cadences.⁸⁰ Diplomats, ambassadors, and secretaries would inspire one another in a network of correspondence or share these rhetorical devices within their scriptoria.⁸¹ These practices might have considerably reshaped the stylistic homogeneity that is evident in the writings of Bernard from his earlier days, when he relied on a far smaller number of secretaries and had more time at his disposal so that he

⁷⁸ Bernard would not necessarily have found help only in Clairvaux. It is conceivable that when he was occupied with the turbulent matters of the schism and was traveling through Italy he called for the assistance of papal scribes to whom he could dictate his messages. The papal notaries, educated in the *ars dictaminis*, would in fact have been schooled in a similar tradition as Nicholas, who had visited Rome and moreover corresponded with at least three popes during his lifetime: see Constable, "Dictators and Diplomats," 43. This could also explain why *brevis* samples in_12 and in_13, which likewise have the schism as their subject, somewhat pair with letters that were added to the corpus later and not with the other *brevis* letters, which cling more closely together. The samples in_12 and 13 contain Epp. 126, 127, 129, 130, 131, 132, 133, 136, 139, 141, and 143 (SBO 7:309–43). For a concise overview of Bernard's interference in the papal schism and its importance for his public career, see the subchapter "A Leading Figure in the Papal Schism 1130–38," in Brian Patrick McGuire, "Bernard's Life and Works," in McGuire, *A Companion to Bernard of Clairvaux*, 40–47. For information on the papal chancery, see Christopher Robert Cheney, *The Study of the Medieval Papal Chancery: The Second Edwards Lecture Delivered within the University of Glasgow on 7th December* (Glasgow, 1966), 20–21. "Instructions about the framing of papal letters may be found in chancery ordinances and in guide-books for chancery clerks; these help to elucidate the legal principles which underly the phraseology." The writing style of the papacy's chancery must have served as an important model to all clerks and diplomats both in ecclesiastical and worldly contexts.

⁷⁹ See Ronald Witt, "Medieval 'ars dictaminis' and the Beginnings of Humanism: A New Construction of the Problem," *Renaissance Quarterly* 35 (1982): 1–35.

⁸⁰ With "comparable cadences" I am here referring to rhetorical devices such as the *cursus*: see Tore Janson, *Prose Rhythm in Medieval Latin from the 9th to the 13th Century* (Stockholm, 1975). The *cursus* has also been tested as a feature for authorship attribution: see Linda Spinazzè, "'Cursus in clausula,' an Online Analysis Tool of Latin Prose," *Proceedings of the Third AIUCD Annual Conference on Humanities and Their Methods in the Digital Ecosystem*, ed. Francesca Tomasi, Roberto Rosselli Del Turco, and Anna Maria Tammaro, Association for Computing Machinery (ACM) International Conference Proceedings Series (ICPS), (New York, 2014), 10:1–6.

⁸¹ On the subject of the *ars dictaminis*, see Giles Constable, *Letters and Letter-Collections* (Turnhout, 1976), 34–35: "This tendency towards a personalization of style and contents in eleventh- and twelfth-century epistolography was paralleled by a tendency, which was in some respects contradictory, towards formalization, which was represented by the emergence of the discipline known as the *dictamen* or *ars dictandi*, with teachers (*dictatores*), text-books (*artes* or *summae dictaminis*), and collections of model letters (formularies). Although *dictamen* now emerged for the first time as a discipline with clearly formulated rules, it had roots deep in the past and was connected in ways which are still not fully understood with the epistolographical rules and traditions which went back to Antiquity. . . . In the course of the twelfth century the number both of teachers and of text-books of *dictamen* spread rapidly, first in Italy and later, in the second half of the century, north of the Alps. Various schools developed with different styles, as at Bologna and Orleans; and although in the earlier twelfth century a certain number of writers, like St. Bernard and Peter the Venerable, who knew about *dictamen*, did not observe its rules, its influence was all but universal by the end of the century."

Speculum 92/S1 (October 2017)

could be present during the various phases of composition. We know of Bernard's increasing discomfort concerning the fact that he felt obliged to delegate the writing of his letters and sermons to assistants, and his dissatisfaction with some of them when it came to grasping the *sensus* of his message.⁸² Perhaps to his own frustration, Bernard was increasingly forced to have faith in the reliability of such scribes as Nicholas to reformulate his initial dictation in a letter that conformed to the style and content Bernard had intended. The *extra* letters would have received far less revision, resulting in the kind of hybrids that float towards middle ground in the figure.

THE SERMONS

In this second visualization we put Nicholas's word to the test. Firstly, assuming that the secretary speaks the "truth"⁸³ in his letter to Henry the Liberal, which Leclercq cited as the most striking example of his plagiarism, we expect that a small number of sermons that occur in Bernard's *De diversis*, namely 6, 7, 21, 62, 83, 100, and 104, could be attributed to him instead. On the side, we test if his claims to Hugh's commentaries on the Psalms, which were also mentioned in the letter, hold any ground.⁸⁴ In a second phase, we follow up on Henri Rochais' conclusions that Bernard—not Nicholas—wrote *De diversis* 40, 41, and 42.⁸⁵ In fact, the *De diversis* collection in its entirety is worth testing here, as it suffers from some considerable issues of authenticity, provenance, and dating and might contain other traces of Nicholas's presence. The corpus comprises an assembly of unpolished and rudimentary sermons found in various, heterogeneous manuscripts, conceivably written down by secretaries and granted little revision by Bernard (unless if they were reused elsewhere).⁸⁶ Bernard never disseminated the *De diversis* sermons himself. They were gathered after his death and passed on for several centuries until Jean Mabillon enumerated and published them in the seventeenth century. Leclercq and Rochais maintained Mabillon's structure in their edition.⁸⁷ Secondly, we have included the *Sermones super Cantica canticorum*, Bernard's literary masterpiece, as the cleanest possible specimen of Bernard's literary style to benchmark against these texts.⁸⁸

⁸² "Multitudo negotiorum in culpa est, quia dum scriptores nostri non bene retinent sensum nostrum, ultra modum acunt stilum suum, nec videre possum quae scribi praecepi," Bernard, Ep. 387 (SBO 8:355–56).

⁸³ "The medieval idea of truth . . . was subjective and personal rather than, as today, objective and impersonal": see Constable, "Forgery and Plagiarism," 23.

⁸⁴ Bernard, *De diversis* 6, 7, 21, 62, 83, 100, 104 (SBO 6/1:105–7, 107–11, 168–70, 295, 324, 367, 374–75). Nicholas used the phrase "aliosque sermones" in the prefatory letter in MS Harley 3073, *Recueil d'études*, 1:50, referring to the aforementioned sermons, a few other texts by Bernard, and, finally, Hugh of St. Victor's chapters on the Psalms gathered in the second book of his *Miscellanea* (PL 177:589).

⁸⁵ Rochais, "Saint Bernard est-il l'auteur des sermons 40, 41 et 42?," 324–45.

⁸⁶ Leclercq, SBO 6/1:59–71.

⁸⁷ Françoise Callerot, "Introduction," in Bernard of Clairvaux, *Sermons divers*, ed. Jean Leclercq, Henri Rochais, and Charles H. Talbot, 3 vols. (Paris, 2006), 1:21.

⁸⁸ Bernard must have started composing its beginnings around the end of 1135, but never commented the entire Song of Songs. They are, nevertheless, regarded as his life's work and greatest literary achievement: see Leclercq, SBO 1:xv–xvi. Leclercq argues Bernard must have passed away before he

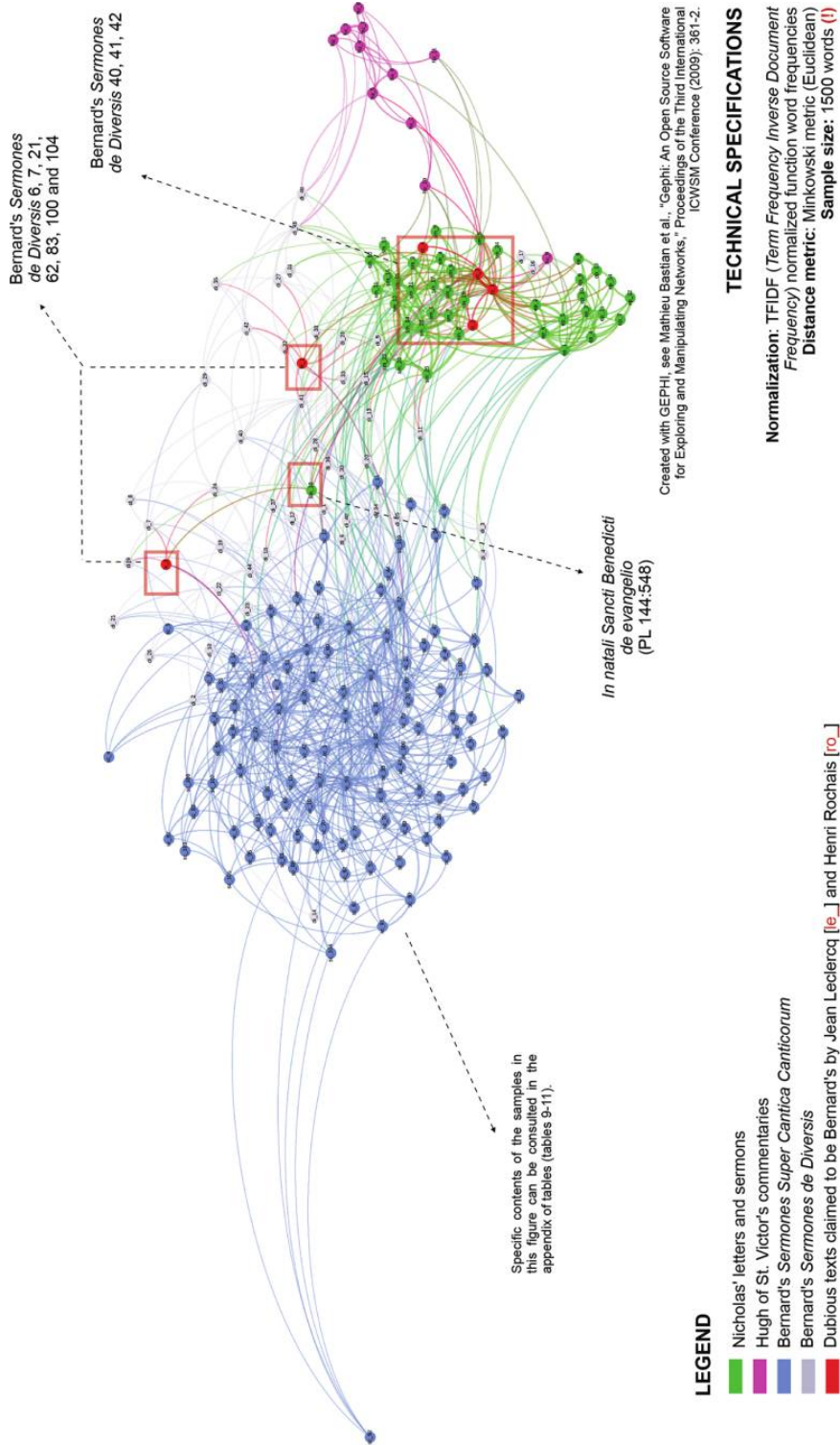


Fig. 3. Network visualization of Bernard of Clairvaux' *Sermones super Cantica Canticorum* and *Sermones de diversis* compared to Nicholas of Montieramey's letters and sermons and Hugh of St Victor's commentaries on the Psalms.

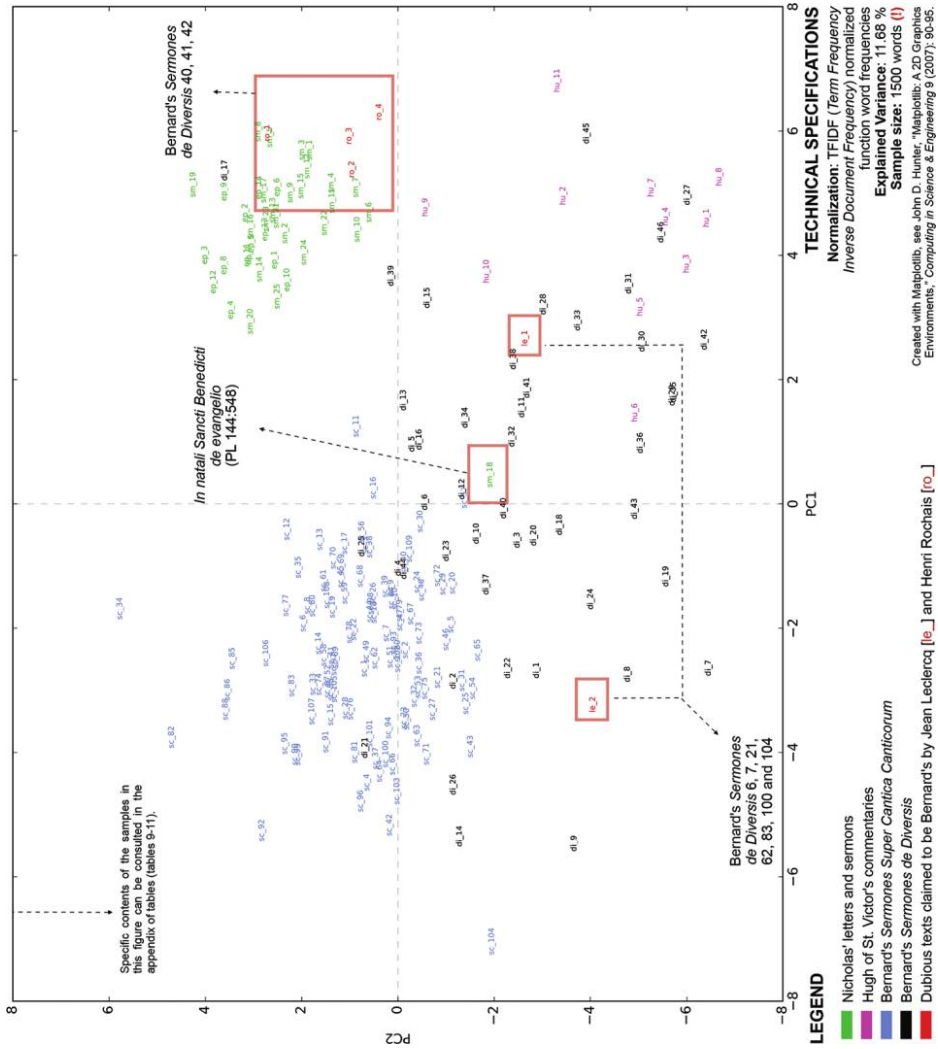


Fig. 4. Principal Components Analysis of Bernard of Clairvaux' *Sermones super Cantica Canticorum* and *Sermones de diversis* compared to Nicholas of Montieramey's letters and sermons and Hugh of St Victor's commentaries on the Psalms.

Fig. 3 (k -NN) and Fig. 4 (PCA) feature the results of matching up these texts. Firstly, when examining the visualizations, it is striking how the diversity of Bernard's *De diversis* is indeed captured. PCA, especially, demonstrates a discernible stylistic incoherence, as the samples burst open all over the plot (especially along the vertical axis of the second principal component), at times suggesting the interference of writers other than Nicholas or Bernard in their composition. Other samples gravitate in between Nicholas and Bernard, and in some cases Nicholas's influence on the style is undeniable. Before discussing some contingent subjects of interest, let us focus on the primary questions at hand. *De diversis* 6, 7, 21, 62, 83, 100, and 104, which Nicholas included in the letter to Count Henry the Liberal (they are split up in two red samples labeled with le_ of Leclercq), do not betray an obvious affinity to Nicholas's style (although le_1 is not far off). Neither are they unambiguously Bernard's. Both samples diverge strongly from Bernard's cluster and seem too hybrid in nature to be restrained to either of the authors' clusters. The case rather demonstrates how difficult it is to defend such concepts as "single authorship" and "textual theft" in a medieval context: the le_ samples are clearly not of a "singular" style (neither Nicholas's style nor Bernard's) but defy classification. In fact, if we compare both k -NN and PCA, Nicholas's influence in sample le_1 seems considerably larger than Bernard's. It has by now become an untenable simplification to argue that Nicholas has stolen these sermons, especially if we review the results of our second case, that of *De diversis* 40, 41, and 42 (four red samples labeled with ro_ of Rochais): although the sermons emanate from Bernardian thought, k -NN and PCA unambiguously cluster all three sermons together with those written by Nicholas, not Bernard.

There are some less straightforward developments on the side. Hugh of St. Victor's presence in both attribution problems remains somewhat unclear. Nicholas included Hugh's commentaries on the Psalms in his collection, yet Figs. 3 and 4 show that he was unlikely to have been the (only) author of this incohesive text (see the purple hu_ samples, of which hu_9 comes closest to Nicholas).⁸⁹ Vice versa, *De diversis* 40 (first part of the dubious ro_ samples) is collected in Hugh of St. Victor's *Miscellanea*. Would Nicholas have known Hugh well, and would they have collaborated before the latter's death in 1141? There is no proof of a direct acquaintance. Nicholas's musical sequences seem largely based on those of Adam of St. Victor, Hugh's choirmaster, but these texts enjoyed a popular circulation, so the sim-

had the chance to finish his work, but it is more likely that Bernard never had the intention of discussing all the *Canticles* and has delivered us a finished work of literature: see Wim Verbaal, "Les sermons sur le cantique de saint Bernard: Un chef d'oeuvre achevé?," *Collectanea Cisterciensa* 61 (1999): 167–85.

⁸⁹ Manuscript studies have argued that they can only be of Hugh's hand: see Joseph de Ghellinck, "Hugues de Saint-Victor," in *Dictionnaire de théologie catholique*, 18 vols. (Paris, 1922–72), 7:245. Although he admits that the *Miscellanea* is a confluence of the apocryphal and the authentic, de Ghellinck based his findings on the *Indiculum* of Hugh's writings. The commentaries on the Psalms often occur among Hugh's authentic works in the manuscript transmission. This has been confirmed in the exhaustive study of the dissemination of Hugh's oeuvre in Rudolf Goy, *Die Überlieferung der Werke Hugos von St. Viktor: Ein Beitrag zur Kommunikationsgeschichte des Mittelalters*, Monographien zur Geschichte des Mittelalters 14 (Stuttgart, 1976), 58–63.

ilarity does not necessarily presuppose a personal tie.⁹⁰ For Bernard and Hugh, however, the connections are less far-fetched. We know they corresponded.⁹¹ Hugh incorporated an entire letter he received from Bernard in his acclaimed masterpiece, the *De sacramentis*.⁹² Likewise, Figs. 3 and 4 show that samples di_46 and di_47 of Bernard's *De diversis* bear some affinity with Hugh's commentaries. These samples comprise the very last additions to Bernard's corpus, *De diversis* 112–25. They are shorter texts, which have not always been accompanied by the preceding sermons but must have circulated as a separate unit in manuscript transmission. Mabillon has argued that their provenance differs from that of the other *De diversis* sermons in a footnote,⁹³ thereby perhaps showing some wariness as to the authenticity of the works.⁹⁴ Although they might be Hugh's, we find that the textual style of both Bernard's *De diversis* and Hugh's commentaries is too unreliable to provide closure. The case for the triangular writing relationship between these authors is compelling, but there is insufficient historical proof to corroborate speculations of a collaboration between Nicholas and Hugh.

CONCLUSION

Jean Leclercq, in aspiring to discern the psychological personality of the author behind any given historical text, conceded the difficulty of infiltrating the “screen of rhetoric” so characteristic to twelfth-century literature, referring to its predilections of imitation and formal rigidity.⁹⁵ The surface of the medieval text can strike one as impenetrable. In a similar vein, Giles Constable has argued for medieval epis-

⁹⁰ “Since Nicolas is known for his plagiarism and incorporated the work of Hugh of St. Victor in the collection of his own opera dedicated to Count Henry, the suspicion arises that Nicolas modeled his work directly on that of Hugh's colleague, Adam of St. Victor”: John F. Benton, “Nicolas of Clairvaux and the Twelfth-Century Sequence,” *Traditio* 18 (1962): 149–179, at 154.

⁹¹ Bernard, Ep. 77, “Ad magistrum Hugonem de Sancto Victore,” SBO 7:184–200; also see Hugh Feiss, “*Bernardus Scholasticus*: The Correspondence of Bernard of Clairvaux and Hugh of Saint Victor on Baptism,” in *Bernardus Magister: Papers Presented at the Nonacentenary Celebration of the Birth of Saint Bernard of Clairvaux*, ed. John R. Sommerfeldt, *Commentarii Cistercienses* 135 (Spencer, MA, 1992), 349–78.

⁹² “Adding to the complications of *De sacramentis* as a text is Hugh's incorporation of passages not only from his own prior works but also from other theologians, patristic and contemporary, sometimes named but often without any attribution at all. In this respect, Hugh nicely represents the overall concern of twelfth-century authors to synthesize their sources”: Paul Rorem, *Hugh of St. Victor, Great Medieval Thinkers* (New York, 2009), 60.

⁹³ The earliest editions of Bernard's *Opera omnia* did not yet include these sermons. It was not until its publication by printer Johann Herwagen of Basel in 1566 that *De diversis* 112–25 found its place among the *Sermones de diversis*: see Gerhard B. Winkler, trans., *Bernhard von Clairvaux: Sämtliche Werke*, 10 vols. (Innsbruck, 1990–99), 9:882^{1h}, 884^{1o}. Jean Mabillon relied on Jacobus Pamelius's edition of these sermons; see note to PL 183:739.

⁹⁴ Jean Mabillon was aware of the fact that Herwagen's and Pamelius's editions were to be approached with great caution when it comes to attribution. Herwagen's and Pamelius's collections of Bede's works are examples of how these editors “ignored and altered rubrics, expurgated passages, disregarded section breaks, and lied outright about the Bedan origins of their material.” Nathan J. Ristuccia, “The Herwagen Preacher and His Homiliary,” *Sacris Erudiri* 52 (2013): 188.

⁹⁵ Jean Leclercq, “Modern Psychology and the Interpretation of Medieval Texts,” *Speculum* 48 (1973): 476–90, at 476.

tology—and he may well have found the statement applicable to all twelfth-century texts—that “style alone is not a reliable guide to authorship,” and that “even today some of the works of Nicholas of Montieramey, who was clearly an accomplished mimic, are not easy to distinguish from those of Bernard and other writers.”⁹⁶ Yet this trait of medieval texts, which is primarily qualitative and open to subjective interpretation, is elusive only in a close-reading approach⁹⁷ and seems not to present a problem when form is quantified in a distant-reading approach. Computational stylistics disables the distracting semantics in which Nicholas’s style is embedded, and patterns the turns of phrase that reveal his presence (or that of a chancery working under his lead). It can only follow, then, that Nicholas’s reputation of being Bernard’s pale shadow is a construction by readers who have undoubtedly experienced the difficulty of peering through the curtain of imitation, citation, and formalization when it comes to recognizing the author behind the text. It turns out that, if Nicholas’s style is not “distinguished,” in the sense that it can be judged as of a high literary value, it *is* nonetheless distinguishable. This does not simply mean that the application of computational stylistics results merely in giving an individualized coloration to the question of authorship. A glance at each of the figures in this article demonstrates the interconnectedness (or “infinite shadings,” in Constable’s words)⁹⁸ laid out as networks between these two authors. Computational stylistics therefore does not simply force us to choose a side in the medieval authorship dilemma, which is infinitely fought out along the axes of the “individual” and the “distributional.” It rather *becomes* these axes and reenacts the tension field *as is*. Neither is Nicholas’s and Bernard’s collaboration depicted as a hierarchical “author-scribe” relationship in one-sided text classifications, nor must we seek refuge in a stopgap conception of infinite authority and authorship. This approach embraces both an acknowledgment that the practice of cooperative medieval authorship is complex, and a refusal to believe that medieval authorship is interminably diffuse. Therefore, computational stylistics provides valuable tools with which to validate or contradict contrasting theories with objective material, taking the voices from the past at face value and opening up avenues to rethink our approach to medieval texts in literary theory, text editing, and historical studies.

⁹⁶ Constable, *Letters and Letter-Collections*, 50.

⁹⁷ Nicholas has often been accused of having a style filled with platitudes or even of having no style of his own at all; see n. 77.

⁹⁸ Constable, “Forgery and Plagiarism,” 3.

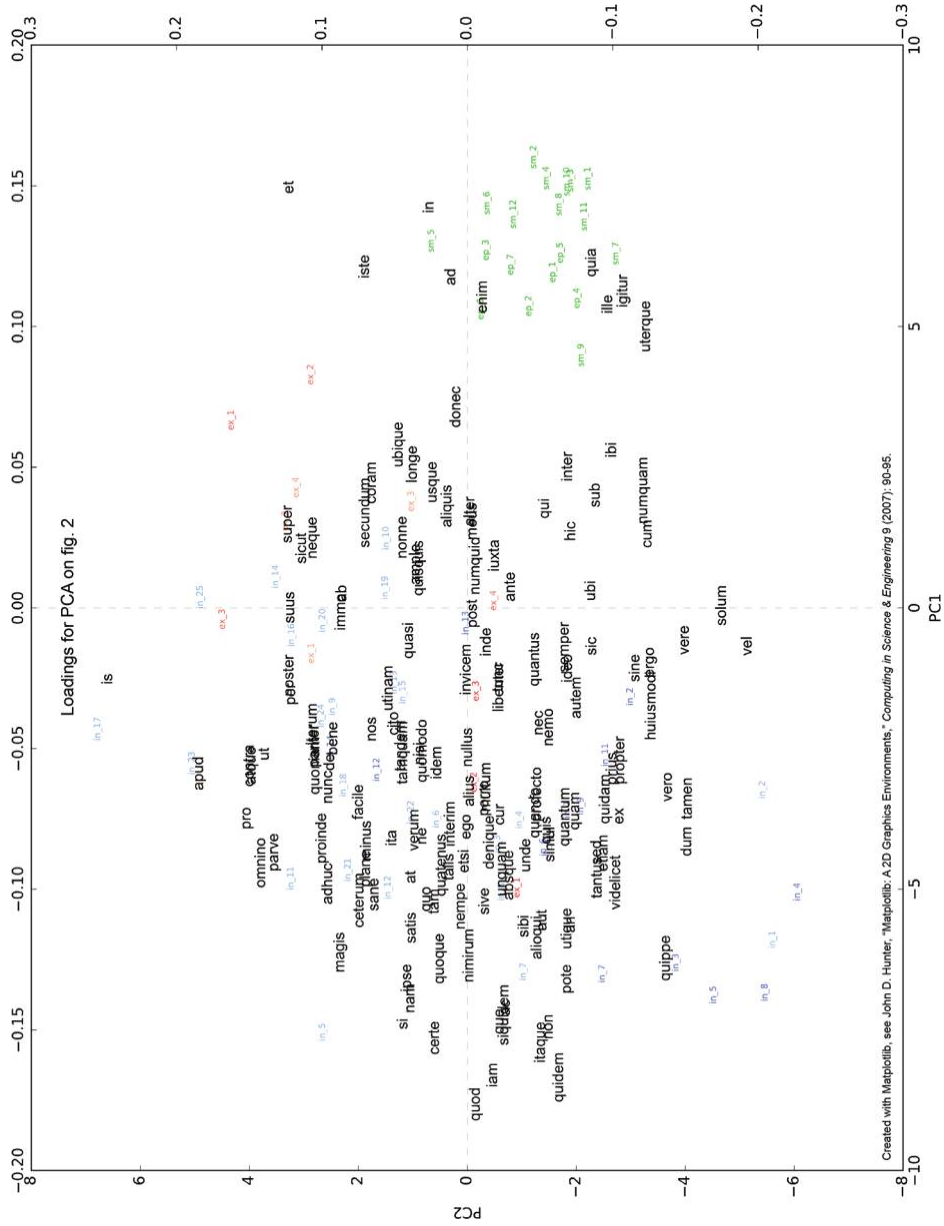


Fig. 5. Principal Components Analysis loadings for Fig. 2. Function words that swarm around author's cluster are most distinctive for that author.

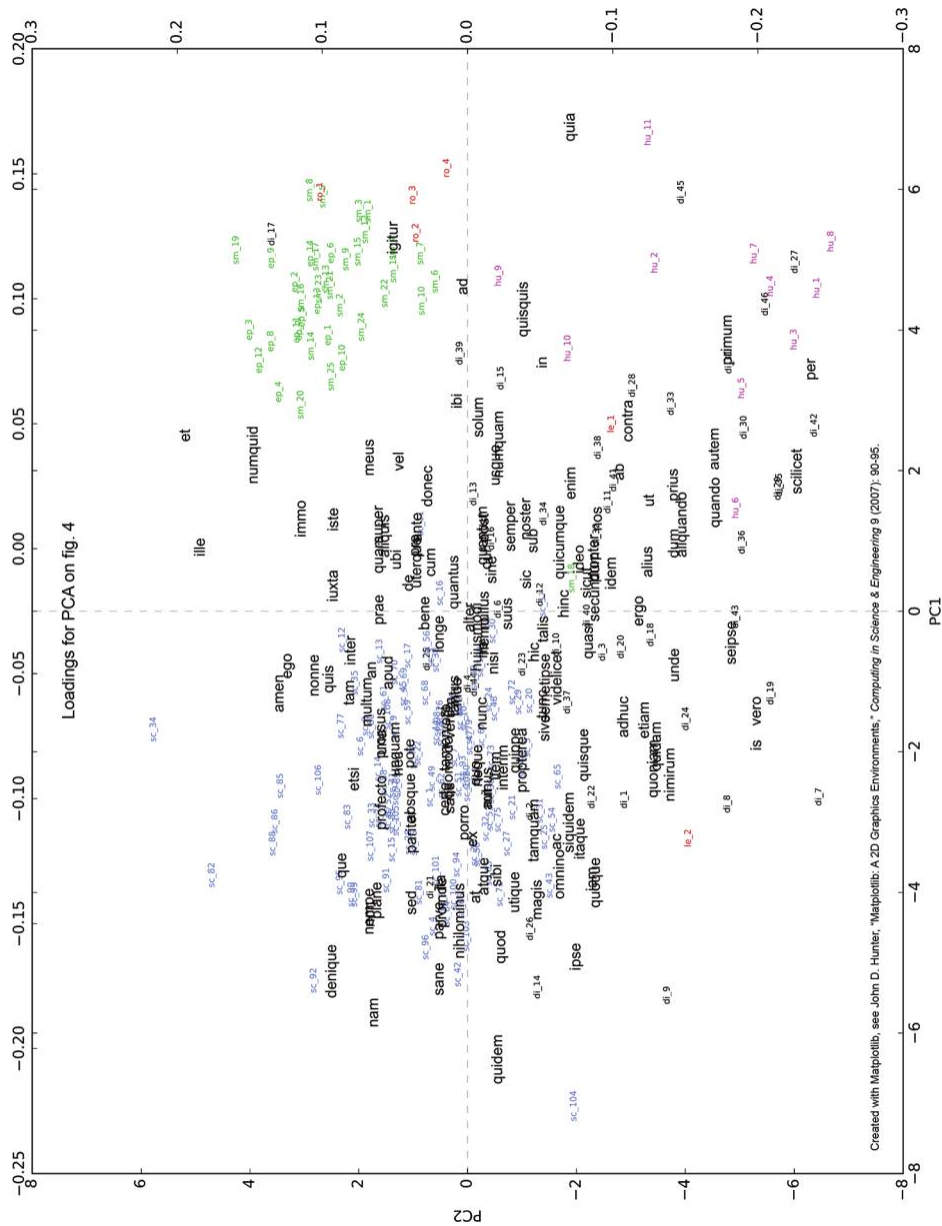


Fig. 6. Principal Components Analysis loadings for Fig. 2. Function words that swarm around author's cluster are most distinctive for that author.

APPENDIX OF TABLES

TABLE 1
Most Frequent Function Words for Figs. 1–2 (The Letters)

1–25	26–50	51–75	76–100	101–125	126–150
et	que	sine	tamquam	iuxta	donec
in	sibi	nam	ante	verum	cito
qui	pro	ita	utique	itaque	nimirum
non	enim	magis	contra	pote	numquam
is	vel	vero	nullus	secundum	cur
hic	ex	apud	igitur	multum	plane
quod	autem	bene	certe	quando	absque
ego	ne	tantus	aliquis	alter	quatenus
sed	per	inter	dum	ibi	proinde
de	aut	immo	semper	tunc	ceterum
ut	tamen	propter	videlicet	sane	longe
ad	iam	quippe	quidam	uterque	pariter
ille	quo	quoque	quisquis	nemo	facile
si	quidem	idem	siquidem	omnino	at
ab	sic	ac	sub	sive	inde
cum	iste	solum	satis	profecto	simul
quis	alius	denique	usque	nonne	ubique
suus	nisi	talis	quantum	prius	tandem
ipse	super	quoniam	numquid	porro	ideo
quam	tiam	adhuc	neque	ample	coram
meus	tam	atque	an	alioqui	huiusmodi
quia	ergo	quantus	post	vere	iterum
nec	ubi	quomodo	unquam	utinam	rursus
nos	sicut	etsi	quasi	libenter	quisque
noster	nunc	unde	minus	interim	parve

TABLE 2
Most Frequent Function Words for Figs. 3–4 (The Sermons)

1–25	26–50	51–75	76–100	101–125	126–150
et	nos	nam	uterque	iuxta	seipse
in	per	quoniam	aliquis	quisquis	item
qui	ex	inter	tunc	videlicet	quicumque
non	autem	denique	solum	apud	an
hic	noster	magis	sane	profecto	donec
is	que	nunc	quando	scilicet	certe
sed	vel	unde	igitur	prius	vere
ad	ergo	quidam	ante	nemo	quisque
ille	quidem	sine	talis	parve	absque
quod	tamen	propter	post	porro	interim
ut	iste	quasi	bene	plane	unquam
de	pro	tam	nullus	ibi	numquam
ego	iam	atque	sub	contra	quantum
cum	alius	quomodo	omnino	immo	pote
suus	ne	quoque	usque	nonne	prorsus

Speculum 92/S1 (October 2017)

TABLE 2 (*continued*)

1–25	26–50	51–75	76–100	101–125	126–150
ab	etiam	tamquam	semper	at	semetipse
si	aut	ac	quippe	nimirum	pariter
ipse	sic	tantus	sive	nihilominus	amen
quis	sicut	idem	alter	primum	proinde
quia	quo	neque	minus	propterea	satis
sibi	nisi	utique	etsi	verum	huiusmodi
meus	vero	adhuc	inde	nempe	numquid
enim	super	dum	siquidem	una	hinc
nec	ita	quantus	itaque	multum	aliquando
quam	ubi	secundum	ideo	longe	prae

TABLE 3
Description of Sample Contents (3,000 words) for Bernard's *intra* Corpus
(*brevis* Publication) in Figs. 1–2

Sample (3,000 words)	Contents	Sample (3,000 words)	contents
sample_#	SBO index and paragraph	sample_#	SBO index and paragraph
in_1	epp. 1.1–13	in_9	epp. 89.3ff., 91, 95, 96, 98, 102, 104, 106, 107.1–3
in_2	epp. 1.13ff., 2.1–12	in_10	epp. 107.3ff., 111, 113.1–2
in_3	epp. 2.12ff., 8, 11.1–5	in_11	epp. 113.2ff., 114, 117, 118, 119, 124, 125, 126.2
in_4	epp. 11.5ff., 12, 24, 25, 42, 65, 67.1–1	in_12	epp. 126.2ff., 127.1–2
in_5	epp. 67.1ff., 68, 69, 70, 72.1–3	in_13	epp. 127.2ff., 129, 130, 131, 132, 133, 136, 139, 141, 143.1–3
in_6	epp. 72.3ff., 73, 78.1–11	in_14	epp. 143.3ff., 150, 152, 156, 158, 159, 168, 169, 178.1–5
in_7	epp. 78.11ff., 79, 82, 83, 85, 87.1–2		
in_8	epp. 87.2ff., 88, 89.1–3		

TABLE 4
Description of Sample Contents (3,000 words) for Bernard's *intra* Corpus (*perfectum* Publication) in Figs. 1–2

Sample (3,000 words)		Contents	
sample_#	SBO index and paragraph	sample_#	SBO index and paragraph
in_1	epp. 3, 5, 6, 7.1–9	in_14	epp. 188.1ff., 189, 190, 191, 192, 193, 194, 195, 196.1
in_2	epp. 7.9ff., 9, 10	in_15	epp. 196.1ff., 197, 198, 99, 200, 201, 202, 203, 204, 205, 206, 207
in_3	epp. 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23.1–4	in_16	epp. 208, 209, 210, 211, 213, 214, 215, 216, 217, 218, 219, 220, 221.1
in_4	epp. 23.4ff., 26, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37.1	in_17	epp. 221.1ff., 222, 223, 224, 225, 226, 227
in_5	epp. 37.1ff., 38, 39, 40, 41, 43, 44, 45, 46, 47, 48, 49	in_18	epp. 228, 229, 230, 231, 232, 233, 234, 235, 236.1
in_6	epp. 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 66	in_19	epp. 236.1ff., 237, 238, 239, 240.2
in_7	epp. 71, 74, 75, 76, 77, 80, 81, 84, 86, 90, 92, 93, 94, 97.1	in_20	epp. 240.2ff., 241, 242, 243, 244.3
in_8	epp. 97.1ff., 99, 100, 101, 103, 105, 108.1	in_21	epp. 244.3ff., 245, 246, 247, 248, 249, 250, 251, 252, 253.1
in_9	epp. 108.1ff., 109, 110, 112, 115, 116, 120, 121, 122, 123, 128	in_22	epp. 253.1ff., 255, 256, 257, 258, 259
in_10	epp. 134, 135, 137, 138, 140, 142, 144, 145, 146, 147.1–2	in_23	epp. 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275
in_11	epp. 147.2ff., 148, 149, 151, 153, 154, 155, 157, 160, 161, 162, 163, 164, 165, 166.1	in_24	epp. 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286
in_12	epp. 166.1ff., 167, 170, 171, 172, 173, 174, 175	in_25	epp. 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301.1–2
in_13	epp. 176, 177, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188.1		

TABLE 5
Description of sample contents (3000 words) for Bernard's
extra corpus (pre-1140) in figs. 1–2

Sample (3,000 words)	Contents	Sample (3,000 words)	Contents
sample_ <i>n</i>	SBO index and paragraph	sample_ <i>n</i>	SBO index and paragraph
ex_1	epp. 311, 312, 313, 314, 315, 316, 317, 318, 319, 322.1	ex_3	epp. 397.4ff., 398, 399, 400, 406, 407, 410, 411, 413, 431, 432, 441, 449
ex_2	epp. 322.1ff., 324, 359, 391, 392, 394, 396, 397.1–4		

TABLE 6
Description of Sample Contents (3,000 words) for Bernard's *extra* Corpus
(1140–45) in Figs. 1–2

Sample (3,000 words)	Contents	Sample (3,000 words)	Contents
sample_ <i>n</i>	SBO index and paragraph	sample_ <i>n</i>	SBO index and paragraph
ex_1	epp. 320, 321, 323, 327, 330, 331, 332, 333, 334, 335, 336, 338.1	ex_3	epp. 356, 357, 358, 360, 362, 385, 393, 416
ex_2	epp. 338.1ff., 339, 340, 341, 342, 346, 347, 348, 349, 350, 351, 353, 354, 355	ex_4	epp. 433, 434, 435, 436, 437, 438, 439, 440, 447, 505, 520, 523, 525

TABLE 7
Description of Sample Contents (3,000 Words) for Bernard's *extra* Corpus
(Post-1145) in Figs. 1–2

Sample (3,000 words)	Contents	Sample (3,000 words)	Contents
sample_ <i>n</i>	SBO index and paragraph	sample_ <i>n</i>	SBO index and paragraph
ex_1	epp. 328, 329, 345, 361, 363, 364, 365.1–2	ex_3	epp. 377.2ff., 378, 379, 380, 381, 382, 383, 384, 387, 389, 390, 401, 402, 403.1–2
ex_2	epp. 365.2ff., 366, 367, 368, 369, 370, 371, 372, 374, 375, 376, 377.1–2	ex_4	epp. 403.2ff., 409, 417, 418, 419, 420, 421, 451, 455, 457, 458, 508, 509, 515, 521

TABLE 8
Description of Sample Contents (3,000 words) for Nicholas's Sermons
and Letters in Figs. 1–2

Sample (3,000 words)	Contents	Sample (1,500 words)	Contents
<i>sample_n</i>	<i>PL</i> (vol:col.)	<i>sample_n</i>	<i>PL</i> (vol:col.)
ep_1	ep. 1 (196:1593a–1594b) ep. 2 (196:1594b–1596a) ep. 3 (196:1596b–1597b) ep. 4 (196:1597b–1598c) ep. 5 (196:1598d–1600a) ep. 6 (196:1600b–1601b) ep. 7 (196:1601c–1601d)	ep_5	ep. 32 (196:1623a–1623c) ep. 33 (196:1623c–1625c) ep. 34 (196:1625d–1626c) ep. 35 (196:1626d–1631a) ep. 36 (196:1631b–1632c) ep. 38 (196:1632c–1635b)
ep_2	ep. 7 (196:1601d–1603a) ep. 8 (196:1603b–1605a) ep. 9 (196:1605b–1605d) ep. 10 (196:1606a–1607d) ep. 11 (196:1608a–1608c) ep. 12 (196:1608c–1609a) ep. 15 (196:1609b–1610a) ep. 15 (196:1610a–1610c) ep. 16 (196:1610d–1613c) ep. 17 (196:1613d–1616a) ep. 18 (196:1616b–1617c) ep. 19 (196:1617d–1618a)	ep_6	ep. 38 (196:1635b–1636c) ep. 40 (196:1636d–1639d) ep. 41 (196:1640a–1640b) ep. 42 (196:1640c–1641c) ep. 43 (196:1641c–1643b)
ep_3	ep. 15 (196:1610a–1610c) ep. 16 (196:1610d–1613c) ep. 17 (196:1613d–1616a) ep. 18 (196:1616b–1617c) ep. 19 (196:1617d–1618a)	ep_7	ep. 43 (196:1643b–1644a) ep. 44 (196:1644a–1645a) ep. 45 (196:1645b–1646d) ep. 46 (196:1647a–1648c) ep. 47 (196:1648d–1649a) ep. 50 (196:1649c–1650c) ep. 51 (196:1651a–1651d)
ep_4	ep. 23 (196:1618c–1619a) ep. 27 (196:1619c–1620a) ep. 29 (196:1620b–1621c) ep. 31 (196:1621d–1622d)	sm_8	sm. 23 (144:629c–637a) sm. 27 (144:649a–649b)
sm_1	sm. 69 (144:897c–902b) sm. 43 (144:732b–735b)	sm_9	sm. 27 (144:649c–652c) 9.42. hom. (144:548c–553a)
sm_2	sm. 43 (144:735c–736b) sm. 55 (144:811c–815c) sm. 56 (144:815d–818b)	sm_10	9.42. hom. (144:553b) sm. 29 (144:660b–666a) sm. 26 (144:646b–647d)
sm_3	sm. 56 (144:818c–822d) sm. 58 (144:828d–832a)	sm_11	sm. 26 (144:648a–649a) sm. 40 (144:717a–722c) sm. 44 (144:736b–737b)
sm_4	sm. 58 (144:832b–834c) sm. 59 (144:834d–838d) sm. 11 (144:557a–558a)	sm_12	sm. 44 (144:737c–740d) sm. 47 (144:761c–765c)
sm_5	sm. 11 (144:558b–563a) sm. 60 (144:839b–841d)		
sm_6	sm. 60 (144:842a–846a) sm. anonym. (144:848b–851d)		
sm_7	sm. anonym. (144:852a–853b) sm. 62 (144:853b–857c) sm. 23 (144:627b–629b)		

TABLE 9
Description of Sample Contents (1,500 words) for Bernard's
Sermones de diversis in Figs. 3–4

Sample (1,500 words)	Contents	Sample (1,500 words)	Contents
sample_n	SBO index and paragraph	sample_n	SBO index and paragraph
di_1	sm. 1.1–7	di_25	sm. 33.4ff., 34.1–3
di_2	sm. 1.7ff., 2.1–6	di_26	sm. 34.3ff., 45.1–5
di_3	sm. 2.6ff., 3.1–4	di_27	sm. 45.5ff., 47, 48, 49, 50.1–3
di_4	sm. 3.4ff., 4.1–2	di_28	sm. 50.3ff., 51, 52, 53, 54
di_5	sm. 4.2ff., 5.1–4	di_29	sm. 55, 56, 57.1
di_6	sm. 5.4ff., 8.1	di_30	sm. 57.1ff., 58, 59, 60, 61.1
di_7	sm. 8.1–8	di_31	sm. 61.1ff., 63, 64, 65, 66, 67
di_8	sm. 8.8ff., 10.1–2	di_32	sm. 69, 70, 71, 72.1–3
di_9	sm. 10.2ff., 11, 12.1–3	di_33	sm. 72.3ff., 73, 74, 75, 76, 77
di_10	sm. 12.3ff., 13, 14.1–4	di_34	sm. 78, 79, 80, 81, 82, 84, 85
di_11	sm. 14.4ff., 15.1–4	di_35	sm. 86, 87, 88.1
di_12	sm. 15.4ff., 16.1–6	di_36	sm. 88.1ff., 89, 90.1–5
di_13	sm. 16.6ff., 17.1–6	di_37	sm. 90.5ff., 91.1–7
di_14	sm. 17.6ff., 18, 19.1	di_38	sm. 91.7ff., 92, 93, 94.1
di_15	sm. 19.1ff., 20.1–3	di_39	sm. 94.1ff., 95, 96.1–3
di_16	sm. 20.3ff., 22.1–6	di_40	sm. 96.3ff., 97, 98
di_17	sm. 22.6ff., 23.1–4	di_41	sm. 99, 101, 102, 103.1–3
di_18	sm. 23.4ff., 24.1–4	di_42	sm. 103.3ff., 105, 106, 107.1–2
		di_43	sm. 107.2ff., 108, 109, 110, 111.1–4
di_19	sm. 24.4ff., 25.1–8	di_44	sm. 111.4ff., 112, 113, 115, 116, 117, 118
di_20	sm. 25.8ff., 26, 27.1–2	di_45	sm. 119, 120, 121, 122, 123, 124.1–2
di_21	sm. 27.2ff., 28.1–2	di_46	sm. 124.2ff., 125.1–3
di_22	sm. 28.2ff., 29.1–2		
di_23	sm. 29.2ff., 30, 31.1–3		
di_24	sm. 31.3ff., 32, 33.1–4		

TABLE 10
Description of Sample Contents (1,500 words) for Bernard's
Sermones super Cantica canticorum in Figs. 3–4

Sample (1,500 words)	Contents	Sample (1,500 words)	Contents
sample_n	SBO index and paragraph	sample_n	SBO index and paragraph
sc_1	sm. 1.1–1.11	sc_56	sm. 42.6–43.2
sc_2	sm. 1.11–2.8	sc_57	sm. 43.2–44.5
sc_3	sm. 2.8–4.1	sc_58	sm. 44.5–45.6
sc_4	sm. 4.1–5.6	sc_59	sm. 45.6–46.4
sc_5	sm. 5.6–6.6	sc_60	sm. 46.4–47.3
sc_6	sm. 6.6–7.7	sc_61	sm. 47.3–48.3
sc_7	sm. 7.7–8.6	sc_62	sm. 48.3–49.1
sc_8	sm. 8.6–9.4	sc_63	sm. 49.1–50.1
sc_9	sm. 9.4–10.4	sc_64	sm. 50.1–50.8

TABLE 10 (continued)

Sample (1,500 words)	Contents	Sample (1,500 words)	Contents
<i>sample_n</i>	<i>SBO</i> index and paragraph	<i>sample_n</i>	<i>SBO</i> index and paragraph
sc_10	sm. 10.4–11.2	sc_65	sm. 50.8–51.8
sc_11	sm. 11.2–12.1	sc_66	sm. 51.8–52.6
sc_12	sm. 12.1–12.9	sc_67	sm. 52.6–53.6
sc_13	sm. 12.9–13.4	sc_68	sm. 53.6–54.4
sc_14	sm. 13.4–14.1	sc_69	sm. 54.4–54.10
sc_15	sm. 14.1–14.8	sc_70	sm. 54.10–56.1
sc_16	sm. 14.8–15.6	sc_71	sm. 56.1–57.2
sc_17	sm. 15.6–16.4	sc_72	sm. 57.2–57.9
sc_18	sm. 16.4–16.14	sc_73	sm. 57.9–58.6
sc_19	sm. 16.14–17.7	sc_74	sm. 58.6–59.1
sc_20	sm. 17.7–18.6	sc_75	sm. 59.1–59.9
sc_21	sm. 18.6–19.7	sc_76	sm. 59.9–60.8
sc_22	sm. 19.7–20.5	sc_77	sm. 60.8–61.5
sc_23	sm. 20.5–21.2	sc_78	sm. 61.5–62.3
sc_24	sm. 21.2–21.10	sc_79	sm. 62.3–63.2
sc_25	sm. 21.10–22.6	sc_80	sm. 63.2–64.3
sc_26	sm. 22.6–23.1	sc_81	sm. 64.3–65.1
sc_27	sm. 23.1–23.8	sc_82	sm. 65.1–65.8
sc_28	sm. 23.8–23.15	sc_83	sm. 65.8–66.7
sc_29	sm. 23.15–24.5	sc_84	sm. 66.7–66.14
sc_30	sm. 24.5–24.8	sc_85	sm. 66.14–67.7
sc_31	sm. 24.8–25.5	sc_86	sm. 67.7–68.3
sc_32	sm. 25.5–25.9	sc_87	sm. 68.3–69.2
sc_33	sm. 25.9–26.5	sc_88	sm. 69.2–70.1
sc_34	sm. 26.5–26.10	sc_89	sm. 70.1–70.8
sc_35	sm. 26.10–27.2	sc_90	sm. 70.8–71.6
sc_36	sm. 26.10–27.9	sc_91	sm. 71.6–71.14
sc_37	sm. 27.9–28.3	sc_92	sm. 71.14–72.5
sc_38	sm. 28.3–28.10	sc_93	sm. 72.5–73.2
sc_39	sm. 28.10–29.3	sc_94	sm. 73.2–73.9
sc_40	sm. 29.3–30.1	sc_95	sm. 73.9–74.7
sc_41	sm. 30.1–30.8	sc_96	sm. 74.7–75.3
sc_42	sm. 30.8–31.2	sc_97	sm. 75.3–75.11
sc_43	sm. 31.2–31.9	sc_98	sm. 75.11–76.8
sc_44	sm. 31.9–32.6	sc_99	sm. 76.8–77.4
sc_45	sm. 32.6–33.3	sc_100	sm. 77.4–78.5
sc_46	sm. 33.3–33.10	sc_101	sm. 78.5–79.4
sc_47	sm. 33.10–34.1	sc_102	sm. 79.4–80.4
sc_48	sm. 34.1–35.3	sc_103	sm. 80.4–81.3
sc_49	sm. 35.3–36.2	sc_104	sm. 81.3–81.9
sc_50	sm. 36.2–37.2	sc_105	sm. 81.9–82.4
sc_51	sm. 37.2–38.2	sc_106	sm. 82.4–83.4
sc_52	sm. 38.2–39.4	sc_107	sm. 83.4–84.4
sc_53	sm. 39.4–40.3	sc_108	sm. 84.4–85.4
sc_54	sm. 40.3–41.4	sc_109	sm. 85.4–85.11
sc_55	sm. 41.4–42.6		

TABLE 11
Description of Sample Contents (1,500 words) for Nicholas's
Sermons and Letters in Figs. 3–4

Sample (1,500 words)	Contents	Sample (1,500 words)	Contents
sample_n	PL (vol:col.)	sample_n	PL (vol:col.)
ep_1	ep. 1 (196:1593a–1594b) ep. 2 (196:1594b–1596a)	ep_8	ep. 31 (196:1622c–1622d) ep. 32 (196:1623a–1623c)
ep_2	ep. 3 (196:1596b–1597b) ep. 4 (196:1597b–1598c) ep. 5 (196:1598d–1600a) ep. 6 (196:1600b–1601b) ep. 7 (196:1601c–1601d)	ep_9	ep. 33 (196:1623c–1625c) ep. 34 (196:1625d–1626c) ep. 35 (196:1626d–1631a) ep. 36 (196:1631b–1631c)
ep_3	ep. 7 (196:1601d–1603a) ep. 8 (196:1603b–1605a) ep. 9 (196:1605b–1605d)	ep_10	ep. 36 (196:1631c–1632c) ep. 38 (196:1632c–1635b)
ep_4	ep. 10 (196:1606a–1607d) ep. 11 (196:1608a–1608c) ep. 12 (196:1608c–1609a) ep. 15 (196:1609b–1610a)	ep_11	ep. 38 (196:1635b–1636c) ep. 40 (196:1636d–1639a)
ep_5	ep. 15 (196:1610a–1610c) ep. 16 (196:1610d–1613c) ep. 17 (196:1613d–1614a)	ep_12	ep. 40 (196:1639b–1639d) ep. 41 (196:1640a–1640b) ep. 42 (196:1640c–1641c) ep. 43 (196:1641c–1643b)
ep_6	ep. 17 (196:1614a–1616a) ep. 18 (196:1616b–1617c) ep. 19 (196:1617d–1618a)	ep_13	ep. 43 (196:1643b–1644a) ep. 44 (196:1644a–1645a) ep. 45 (196:1645b–1646d) ep. 46 (196:1647a–1647c)
ep_7	ep. 23 (196:1618c–1619a) ep. 27 (196:1619c–1620a) ep. 29 (196:1620b–1621c) ep. 31 (196:1621d–1622c)	ep_14	ep. 46 (196:1647d–1648c) ep. 47 (196:1648d–1649a) ep. 50 (196:1649c–1650c) ep. 51 (196:1651a–1651d)
sm_1	sm. 69 (144:897c–901c)	sm_14	sm. 62 (144:856a–857c)
sm_2	sm. 69 (144:901d–902b) sm. 43 (144:732b–735b)	sm_15	sm. 23 (144:627b–629b) sm. 23 (144:629c–633b)
sm_3	sm. 43 (144:735c–736b) sm. 55 (144:811c–814c)	sm_16	sm. 23 (144:633c–637a) sm. 27 (144:649a–649b)
sm_4	sm. 55 (144:814d–815c) sm. 56 (144:815d–818b)	sm_17	sm. 27 (144:649c–652c) 9.42. hom. (144: 548c–549b)
sm_5	sm. 56 (144:818c–822b)	sm_18	9.42. hom. (144:549c–553a)
sm_6	sm. 56 (144:822b–822d) sm. 58 (144:828d–832a)	sm_19	9.42. hom. (144:553b) sm. 29 (144:660b–663d)
sm_7	sm. 58 (144:832b–834c) sm. 59 (144:834d–836b)	sm_20	sm. 29 (144:664a–666a) sm. 26 (144:646b–647d)
sm_8	sm. 59 (144:836c–838d)	sm_21	sm. 26 (144:648a–649a)
sm_9	sm. 11 (144:557a–558a)	sm_22	sm. 40 (144:717a–719d) sm. 40 (144:720a–722c)
sm_10	sm. 11 (144:558b–561d) sm. 11 (144:562a–563a) sm. 60 (144:839b–841d)	sm_23	sm. 44 (144:736b–737b) sm. 44 (144:737c–740d)
sm_11	sm. 60 (144:842a–845d)	sm_24	sm. 47 (144:761c–761d)
sm_12	sm. 60 (144:846a) sm. anonym. (144:848b–851d)	sm_25	sm. 47 (144:762a–765c) sm. 47 (144:766a–766b)
sm_13	sm. anonym. (144:852a–853b) sm. 62 (144:853b–855d)		sm. 52 (144:794b–797b)

Jeroen De Gussem, Ghent University (jedgusse.degussem@ugent.be)

Speculum 92/S1 (October 2017)