

## Distant Reading of Gratian's *Decretum*

*Paul Evans*



GRATIANVS. MONACHVS. CAMALD  
BONONIA. IN. MONASTERIO. S. PELICIS.  
DECRETI. CONDITION. SUB. A. D. MCL.  
EX. DEI. M. G. M. ET. C. U. ET. B. R. O. M.

# **Chapter 0**

## **Introduction**

### **Origin of the Project**

This dissertation approaches a classic medieval text, Gratian's *Decretum*, in a distinctly untraditional way. I found my way to this topic through a specific and unique combination of academic interests and previous professional experience and against the backdrop of the rapid transformation between 2004 and 2009 of Humanities Computing into the new academic discipline of Digital Humanities.<sup>1</sup> Justifying the project and its findings therefore necessarily involves a more personal narrative than is typical for the introduction to a dissertation: the most straightforward way to discuss the development

---

<sup>1</sup> The term Digital Humanities came into general use in 2004 with the publication of *A Companion to Digital Humanities*, the National Endowment for the Humanities (NEH) created its Office of Digital Humanities (ODH) in 2008, and William Pannapacker's "The MLA and the Digital Humanities" in the December 28, 2009 issue of *The Chronicle of Higher Education* brought developments in DH to the attention of a widespread audience, notably including university administrators, for the first time. Matthew Kirschenbaum, "What Is Digital Humanities and What's It Doing in English Departments?" in *Debates in the Digital Humanities*, ed. Matthew K. Gold (Minneapolis: University of Minnesota Press, 2012), 3–11.

of the digital methods used in the project and the scholarly context from which they emerged is through my first-hand experiences with them.

The most significant finding of my dissertation is that the author who wrote the thirty-six case statements introducing the hypothetical cases that make up the second part of Gratian's *Decretum* is very unlikely to have been the same as the author who wrote the *dicta* in either the first or second recension of the work. The statistical method used to make this determination assigns probable authorship on the basis of frequencies of common function words like prepositions and conjunctions in a sample of text; the method will be explained in full detail in Chapter 4.

I did not start work on this project thinking that the authorship of the case statements was in any way a research problem. I assumed that by definition the author of the case statements was one and the same person as the author of the first-recension *dicta*. It is therefore worth explaining in some detail how I came to make this completely unexpected finding.

I worked in information technology as a system administrator and manager for most of the twenty-three years after I graduated with an undergraduate degree in History from



UC San Diego in 1984. Stanley Chodorow had been the advisor for my undergraduate senior thesis on the role of the cardinals in the thirteenth and fourteenth centuries, and I knew that he had written a book about Gratian's *Decretum*.<sup>2</sup> I was therefore aware of Gratian in a general sort of way, although the only use I made of the *Decretum* in connection with my thesis was to consult Emil Friedberg's 1879 edition for the Latin text of Nicholas II's 1059 decree on papal elections (D.23 c.1).

Chodorow urged me to use computer-aided typesetting for the project, and in this way I acquired a then-unusual skill that led directly to my IT career. In the mid- to late-1980s I went on to take most of the required courses for the undergraduate Computer Science major at UC San Diego (e.g., Data Structures, Compiler Construction, Operating Systems), although I did not enroll in a degree program. During my professional career, I was never primarily a programmer, but from time to time my job responsibilities did include programming projects in C and Perl and ultimately servlet-based web applications in Java.

---

<sup>2</sup> Stanley Chodorow, *Christian Political Theory and Church Politics in the Mid-Twelfth Century; the Ecclesiology of Gratian's Decretum*, Publications of the Center for Medieval and Renaissance Studies, U.C.L.A., 5 (Berkeley: University of California Press, 1972).



In October 2003, quite by accident, I became aware of Anders Winroth's *The Making of Gratian's Decretum*.<sup>3</sup> I had done a Google search for Chodorow's contact information, and in the process came across his review of Winroth's book in *The English Historical Review*.<sup>4</sup> From the review I learned that Winroth had identified five twelfth-century manuscripts as a first recension of the *Decretum*, shorter and more coherent than later more widely-circulated versions of the text. In addition, I became aware of Winroth's claim that two different authors, Gratian 1 and Gratian 2, were responsible for the first and second recensions. It was clear to me that there had been a revolution in Gratian studies.

From September 2007 to May 2009, I was a student in the History of Christianity master's program at Yale Divinity School. Among the courses I took was a one on Latin Paleography that Richard and Mary Rouse of UCLA taught in the Beinecke Rare Book and Manuscript Library. In October 2009, I attended a talk by David Ganz (then of

---

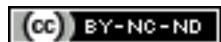
<sup>3</sup> Anders Winroth, *The Making of Gratian's Decretum* (Cambridge: Cambridge University Press, 2000).

<sup>4</sup> Stanley Chodorow, "Review of the Making of Gratian's Decretum by Anders Winroth," *The English Historical Review* 118, no. 475 (February 2003): 174–76.



King's College, London) who pointed out that there were two different versions of the text of the *Capitulare Carisiacense* (873) in Beinecke MS 413 with numerous variant readings. He suggested that transcribing and comparing the two versions would be a worthwhile project for a graduate student. Because of my paleography training with the Rouses, I felt qualified to undertake the project, and set to work on the manuscript right away. Although I had a general interest in applying my computing background to my academic work before I graduated from YDS, the Beinecke 413 project was my first opportunity to do so. Within a month, I had created a custom text-encoding format for my transcriptions and written a prototype textual difference visualizer in Perl to compare them. My notes from the project indicate that by January 2010 I was using the term Digital Humanities to describe my work.

In August 2010, I started on my PhD in the Medieval and Byzantine Studies (MBS) program at The Catholic University of America (CUA) in Washington, DC. I went to CUA specifically to work with Kenneth Pennington on Gratian's *Decretum*. Even before moving from New Haven to Washington, I had participated in Winroth's class on law in medieval Europe at Yale, and, once at CUA, I took Pennington's classes on canon and Roman law, and his sources seminar (twice). From 2010 through 2012, then, I



thoroughly immersed myself in the scholarly debates surrounding the identity of Gratian and the recensions and dating of the *Decretum*. These studies produced a certain level of personal discomfort at being unable to reconcile the contradictory positions staked out by Pennington and Winroth.

Pennington and his students Melodie Harris Eichbauer and Atria A. Larson argued that the *Decretum* was the result of a long process of continuous revision. They therefore saw the first recension as one stage in a series of stages in the composition of Gratian's text and argued that the work entered circulation at an early date, in the 1130s. Pennington in particular argued strongly that a single author, Gratian, compiled and wrote both the first and second recensions of the *Decretum*. Winroth and his student John Wei argued that the first and second recensions represented discrete and discontinuous stages in the composition process of the *Decretum* and that the two recensions were compiled and written by two different authors, Gratian 1 and Gratian 2. Winroth has insisted on a late date, around 1140, for the first recension. Much of the debate over whether the *Decretum* was the result of a continuous or discontinuous process of composition focused on the



Sankt Gallen 673 (Sg) manuscript. The text in Sg is shorter than the first recension (somewhat less than 1,050 canons as opposed to 1,860),<sup>5</sup> and Pennington and some of his students have argued that it represented, at some unknown number of removes, an earlier version of the *Decretum* than Winroth's first recension. Winroth and Wei have argued that Sg was a relatively uninteresting abbreviation of a first recension manuscript with some second recension interpolations.<sup>6</sup>

In a January 2011 advising conversation, Jennifer Davis, director of graduate studies for MBS at the time, suggested that, given my professional background, it would be strategically advantageous for the purpose of whatever academic career I might hope to have to position myself as a Digital Humanities specialist. In the summer of 2010, I had taught myself to write Python web applications on the Google App Engine (GAE) platform, so in the first half of 2011, I developed Ingobert, a Python/GAE web

---

<sup>5</sup> Carlos Larraínzar in "El Borrador de La 'Concordia' de Graciano: Sankt Gallen, Stiftsbibliothek MS 673 (=Sg)," *Ius Ecclesiae: Rivista Internazionale Di Diritto Canonico* 11, no. 3 (1999): 601, describes Sg as having "poco menos de 1,050 *auctoritates* y en torno a los 650 *dicta*." "The second recension contains 3,945 canons (including the paleae) in the editions. The first recension contains only 1,860 canons (47 percent)." Winroth, *The Making of Gratian's Decretum*, 122.

<sup>6</sup> See Melodie H. Eichbauer, "Gratian's Decretum and the Changing Historiographical Landscape," *History Compass* 11, no. 12 (December 2013): 1111–25 for a good recent overview of these debates.



application to visualize textual differences in Beinecke 413 in connection with an independent study project supervised by Pennington and Davis.<sup>7</sup> Largely on the strength of the Ingobert project, Neil Fraistat of the University of Maryland hired me as a graduate assistant at the Maryland Institute for Technology in the Humanities (MITH) to work as a Scala/Lift programmer on the Active OCR project.<sup>8</sup>

I finished my PhD comprehensive examinations in October 2012 and advanced to candidacy in January 2013. I had not yet made a definite decision to pursue a dissertation project with a Digital Humanities component but audited Matt Kirschenbaum's graduate introduction to Digital Humanities course at the University of Maryland in Spring 2013, with the idea that an overview of the field might suggest a potential project.

---

<sup>7</sup> Ingobert was named after the Carolingian scribe of the Bible of San Paolo fuori le Mura. Some scholars have suggested that he was responsible for Beinecke 413; the hand is certainly similar to his. The Ingobert project is still under active development: see my GitHub [Ingobert2](#) repository for the source code of the current version of the Python web application ported to the Django platform.

<sup>8</sup> NEH ODH Grant number: [HD-51568-12](#)



One step in the direction of a digital project was to obtain an electronic version of the *Decretum* text. In the mid- to late-1980s, Timothy Reuter and Gabriel Silagi edited the *Wortkonkordanz zum Decretum Gratiani* for the *Monumenta Germaniae Historica* (MGH) in Munich, a computer-generated concordance in the tradition of Father Roberto Busa's *Index Thomisticus*.<sup>9</sup> As part of the project, the MGH undertook to scan, correct, and encode the 1879 Friedberg edition of the *Decretum*, in the now-obsolete and non-tree-structured Oxford Concordance Program (OCP) format. In spring 2013, Winroth and Lou Burnard of the Oxford Text Archive (OTA) each provided me with a copy of the Reuter and Silagi e-text. The two copies, however, differed in many places, and I had to go through a process similar to preparing a critical edition to restore the e-text to a state as close as possible to what I thought the editors intended. I then began to experiment with writing Python programs that used regular expressions to extract textual features of interest. The fact that the OCP e-text format is not tree-structured the way XML is—

---

<sup>9</sup> Timothy Reuter and Gabriel Silagi, eds., *Wortkonkordanz Zum Decretum Gratiani*, *Monumenta Germaniae Historica. Hilfsmittel* 10 (München: Monumenta Germaniae Historica, 1990). See Susan Hockey, "The History of Humanities Computing," in *A Companion to Digital Humanities*, ed. Susan Schreibman, Raymond George Siemens, and John Unsworth, *Blackwell Companions to Literature and Culture* 26 (Malden, MA: Blackwell Pub, 2004), 3–19, for Busa (4) and OCP (8).



textual features have start tags but do not have end tags—makes it extremely difficult to parse, so this was a slow process.<sup>10</sup>

My initial focus was on the use of David Mimno's Mallet (MAchine Learning for LanguagE Toolkit) to topic model *dicta* and canon texts from the first and second recensions of Gratian's *Decretum* as a way to identify new topics added in the second recension.<sup>11</sup> The inspiration was Pennington's observation that most passages in the *Decretum* dealing with the legal status of Jews, particularly those dealing with forced conversion, were introduced only in the second recension.<sup>12</sup> My goal was to see whether Mallet could bring more such topics to the surface, by topic modeling the first and second parts of the vulgate *Decretum*, topic modeling the first recension, and seeing

---

<sup>10</sup> See Appendix 1 and Appendix 2 for Python source code of OCP parsers written for this project.

<sup>11</sup> Placeholder for Mallet footnote.

<sup>12</sup> Kenneth Pennington, "The Law's Violence Against Medieval and Early Modern Jews," *Rivista Internazionale Di Diritto Comune* 23 (2013): 23–44; and Kenneth Pennington, "Gratian and the Jews," *Bulletin of Medieval Canon Law* 31, no. 1 (2014): 111–24.



what topics were left when the first recension topics were subtracted from the vulgate topics.<sup>13</sup> While simple in concept, this proved prohibitively difficult in practice.<sup>14</sup>

In July 2013, I was working at MITH, and following the DH 2013 conference at University of Nebraska-Lincoln out of general interest. One presentation in particular

---

<sup>13</sup> “Vulgate” in this context refers to the version of the text of Gratian’s *Decretum* found in Emil Friedberg’s 1879 edition. The vulgate includes approximately 150 canons (the so-called “palea”) added after the completion of the second recension.

<sup>14</sup> This project was attractive to Pennington because although the results would be obtained computationally, they could be verified by someone doing a close reading of the text of the *Decretum*. There were three insurmountable barriers to carrying out the project as originally conceived: the time required to prepare the necessary text samples; the difficulty in determining the number of topics to look for (a necessary precondition for unsupervised topic modeling); and the fact that there was no obvious way to subtract topics.

While a stylometric analysis for authorship attribution requires only the *dicta* (*ante*, *post* and *initiale*) thought to have been written by Gratian himself, a topic can be present in any text in the *Decretum*, inscriptions and canons as well as rubrics and *dicta*. It took six weeks—twice—just to prepare a proxy text for the first-recension *dicta*. (In late Summer 2015 I discovered quality anomalies in the *dicta* samples I had hand-edited in Fall 2013, so in Fall 2015, I regenerated the *dicta* samples from scratch by rigorously cross-checking all of the hand-edited *dicta* against a data set automatically generated using Python regular expressions until no differences remained between the two sets of samples.) There is about four times as much text by word count in the canons as there is in the *dicta*, so I estimated that it would take just under six person-months to prepare a proxy text for the first-recension canons.

The Latent Dirichlet Allocation (LDA) algorithm that MALLET uses to generate topic models has to be provided with an exact number of topics to look for. In February 2014, I carried out a preliminary experiment to obtain a rough estimate of the number of topics in the *Decretum*, inspired by the metaphor of focusing a telescope. I took the second-recension *dicta* and repeatedly ran MALLET on them, looking for values of the number of topics at which Pennington’s topic on the legal status of Jews came into focus. Pennington’s topic started to appear at somewhere over 200 topics.



caught my attention: "Stylometry and the Complex Authorship in Hildegard of Bingen's Oeuvre" by Mike Kestemont, Sara Moens, and Jeroen Deploige. Their work was later published as a paper, but the conference website had an unusually detailed abstract, and a video was made available as part of the presentation.<sup>15</sup>

The applicability of Kestemont's methodology to the intractable problem of the authorship of the *Decretum* was immediately obvious to me; it seemed to finally offer a way past endless debates based on indirect evidence about whether there had been one Gratian or two. I would extract the first- and second-recension *dicta*, those parts of the text of the *Decretum* thought to have actually been written (depending on whether one accepted Pennington's or Winroth's argument) by Gratian or by Gratian 1 and Gratian

---

<sup>15</sup> Abstract: Mike Kestemont, Sara Moens, and Jeroen Deploige, "Stylometry and the Complex Authorship in Hildegard of Bingen's Oeuvre," in *Digital Humanities 2013: Conference Abstracts* (Lincoln, NE: University of Nebraska-Lincoln, 2013), 255–58, <http://dh2013.unl.edu/abstracts/ab-126.html>. Video: Mike Kestemont, "Documentary: 'Hildegard of Bingen: Authorship and Stylometry' [HD]," July 18, 2013, <https://vimeo.com/70881172>. Paper: Mike Kestemont, Sara Moens, and Jeroen Deploige, "Collaborative Authorship in the Twelfth Century: A Stylometric Study of Hildegard of Bingen and Guibert of Gembloux," *Literary and Linguistic Computing* 30, no. 2 (June 2015): 199–224.



<sup>2<sup>16</sup></sup> and run the same kind of analysis that Kestemont had run for Hildegard of Bingen and Guibert of Gembloux. I expected the results to provide an unambiguous answer, sufficiently compelling to both Pennington and Winroth to settle the debate one way or the other as to whether there had been one or two authors.

In August and September of 2013, I replicated the working software environment with which Kestemont had obtained his Hildegard results, installing R, R Studio, and the stylometry for R package that Kestemont had written with Maciej Eder and Jan Rybicki.<sup>17</sup> I started extracting text samples from Reuter and Silagi's e-text of the Friedberg edition of the *Decretum*. The fact that the e-text was encoded in the obsolete (and not tree-structured) Oxford Concordance Program format made this an extremely difficult and time-consuming process. In fact, the only parts of the e-text that could both be easily extracted using Python regular expressions and, once extracted, quickly

---

<sup>16</sup> To the extent that there is some one person we can point to as corresponding to our idea of "Gratian," it's the author of the first-recension *dicta*. "The *dicta* in Gratian's *Decretum* bring the reader closer to its author than any other part of the text." Winroth, *The Making of Gratian's Decretum*, 187.

<sup>17</sup> Maciej Eder, Mike Kestemont, and Jan Rybicki, "Stylometry with R: A Suite of Tools," in *Digital Humanities 2013: Conference Abstracts* (Lincoln, NE: University of Nebraska–Lincoln, 2013), 487–89, <http://dh2013.unl.edu/abstracts/ab-136.html>.



verified to be correct were the case statements. This made the case statements an obvious first choice for a test sample, although my ultimate goal was to compare only the first- and second-recension *dicta*.

Next, I needed a distraction text presumably not written by Gratian. For that purpose, I chose extracts from the pseudo-Augustinian *De vera et falsa penitentia* quoted by Gratian in his *de Penitentia*, a treatise on penance inserted at C.33 q.3 in the second part of the *Decretum*. In the interest of getting results quickly, I hand-edited the excerpts directly out of the Reuter and Silagi e-text. With the case statements and the *De vera* extracts in hand, I now had enough in the way of text samples to verify that I had installed and configured R, R Studio, and stylo correctly. I have to admit that I was somewhat disappointed that the results of the first test were exactly what I should have expected: the case statements and the excerpts from *De vera* displayed a marked left-right separation along the horizontal x-axis representing the first principal component, indicating that they were written by two different authors. Because *De vera* is an anonymous work that predated the *Decretum* by no more than a decade or so, and because Gratian was one of the earliest authors to quote extensively from it (although



not the earliest, as I mistakenly believed at the time), I thought it would make an excellent dissertation topic if it could be shown that Gratian had forged *De vera*.

Having confirmed that my test environment could correctly distinguish the authorship of the case statements from that of the pseudo-Augustinian excerpts from *De vera*, I moved on to the much slower process of hand-editing text samples of the first- and second-recension *dicta* from the Reuter and Silagi e-text.<sup>18</sup>

By mid-September 2013, I had edited the first- and second-recension *dicta* for the first part of the *Decretum* (D.1-101). When I ran stylo on the samples, however, I got neither of the two results I had expected: either a tight clustering of all *dicta* (first- and second-recension as well as case statements) indicating a single author and confirming all of Pennington's arguments for the unity of Gratian, or alternatively, a bimodal distribution confirming Winroth's arguments for a Gratian 1 and a Gratian 2. Instead,

---

<sup>18</sup> For the purpose of comparing the first- and second-recension *dicta*, I define the first-recension *dicta* as the *dicta* (*ante* and *post*, but not *initiale*) in the first and second parts of the Friedberg edition of the *Decretum* to which I apply the transformations defined by Winroth's appendix. I define the second-recension *dicta* as the *dicta* (*ante* and *post*, but not *initiale*) in the first and second parts of Friedberg remaining after the proxy first-recension text generated by applying the Winroth transformations has been subtracted.



these preliminary results seemed to suggest that the first recension *dicta* had many authors, perhaps one or two of whom went on to write the second recension *dicta*. What was completely unexpected, however, was that the case statements clustered far away from the *dicta*, extremely strong evidence that they had not been written by the same author. I immediately realized that if this accidental result held up under further testing it would be both significant and controversial. (See Figure 1 below.)<sup>19</sup>

---

<sup>19</sup> The statistical technique of principal components analysis (PCA) projects or flattens an n-dimensional vector space representing the total variation between a set of samples into a more easily-visualized 2-dimensional plot. In this case, 65 vectors representing the variation in the frequency of occurrence of the 65 most frequent words in the text samples were collapsed into a smaller number of synthetic principal components. The horizontal x-axis represents the first principal component (PC1), which represents 16.9% of the total variation between the samples. The vertical y-axis represents the second principal component (PC2), which represents 12.5% percent of the total variation between the samples. The units along the x- and y-axes are standard deviations away from the means (indicated by the dashed lines) for each of the two principal components. Principal components analysis and its application to the problem of authorship attribution will be covered in depth in Chapter 4, Stylometry.



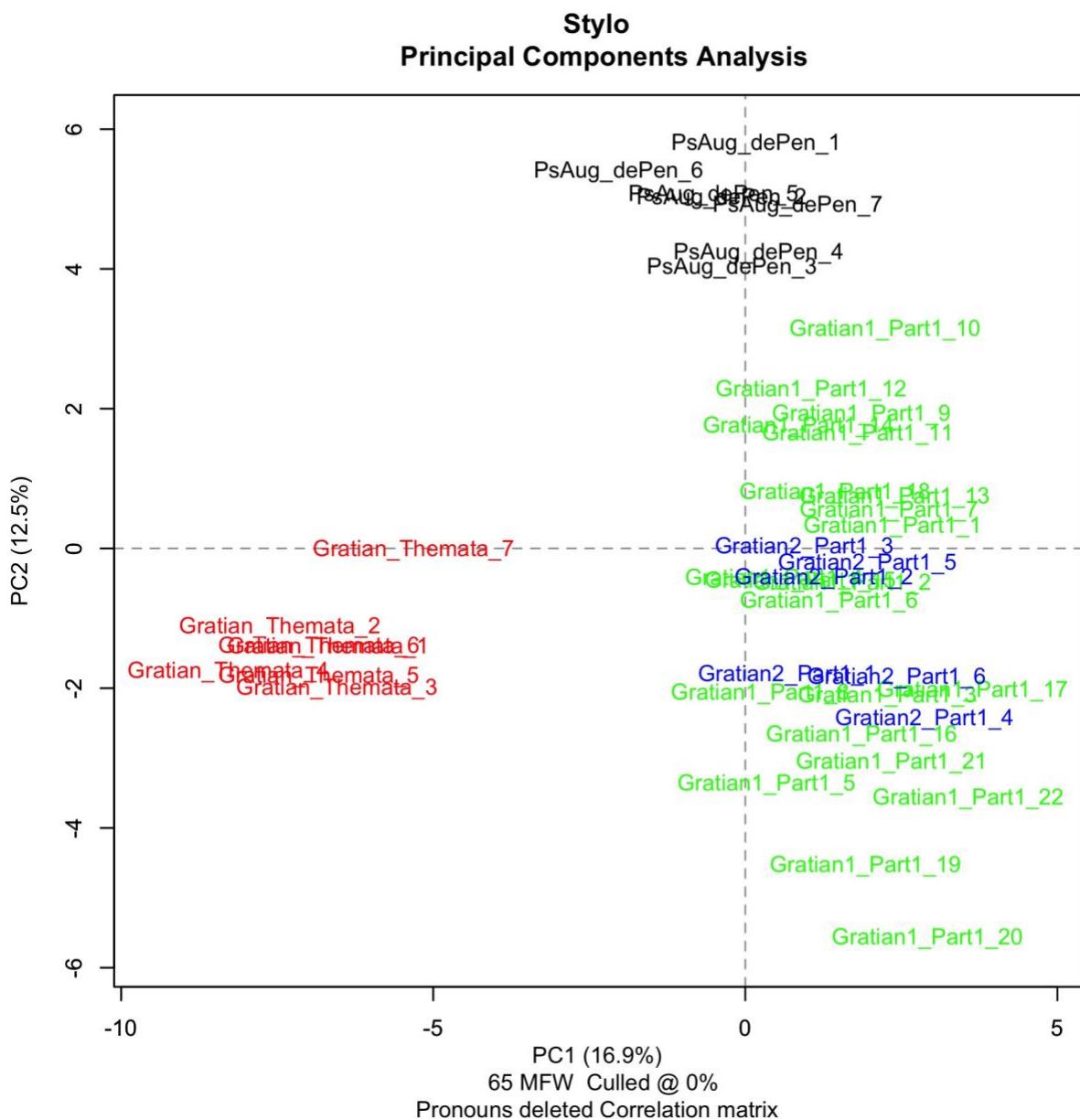


Figure 1 10 Sep 2013

Scholars working in the field of medieval canon law have long been accustomed to thinking of the author of the *dicta* (or after Winroth's discovery, at least the author of the first-recension *dicta*) as Gratian. My initial interpretation of these surprising results was therefore that Gratian had not written the case statements. Soon, however, I came to see the image produced by stylo as telling a different and very specific "likely story"—a phrase borrowed from Plato's *Timaeus*—or what Pennington calls a "conjectural novella" about the earliest beginnings of Gratian's project and, by extension, about the dawn of the formal, academic study of canon law and of the European university, the moment when the medieval school run by a lone master began to evolve into a faculty whose members taught a standardized program.

Many scholars, notably Noonan and Pennington, have seen the thirty-six cases that make up the second part of the *Decretum*, each organized around a case statement, as Gratian's unique, original, contribution to the teaching of canon law.<sup>20</sup> There is also a scholarly consensus foundational to most recent work on the composition of the

---

<sup>20</sup> John T. Noonan, "Catholic Law School - A.D. 1150," *Catholic University Law Review* 47 (1997): 1201; and Kenneth Pennington, "The Biography of Gratian, the Father of Canon Law," *Villanova Law Review* 59 (2014): 689.



*Decretum* that Gratian drew on just five formal sources for the bulk of the authorities he cited.<sup>21</sup> These observations prompted me to reframe my initial interpretation and consider the possibility that the eponymous Gratian who gave his name to the entire project had written *only* the case statements.

Noonan ended his article “Gratian Slept Here” with a contemporary report of an 1143 case argued at San Marco in Venice in which a Gratian participated as a consultant to the judge. Many subsequent books and articles have referred to Noonan’s discussion of the courtroom sighting of “the silent figure in the shadows of S. Marco.”<sup>22</sup> I saw the plot generated by the stylometry software as an indirect but compelling classroom sighting of Gratian: seated at a table with his case statements in hand and their lists of questions as his syllabus, he harmonized the canons for his students directly out of the formal sources in the form of a pile of books on the table in front of him.

---

<sup>21</sup> Winroth, *The Making of Gratian's Decretum*, 15. Roughly one-fifth of the text of the *Decretum* has traditionally been attributed to Gratian himself; the other fourth-fifths of the text is made up of excerpts from the authorities Gratian cited.

<sup>22</sup> John T. Noonan, “Gratian Slept Here: The Changing Identity of the Father of the Systematic Study of Canon Law,” *Traditio* 35 (January 1979): 171–72.



This conjectural novella provides a way to make sense of the fact that the author of the case statements does not appear to have written either the first- or-second recension *dicta*. In the beginning, the *Decretum* existed only in the form of the master expounding the canons to his students in a classroom presentation guided by the case statements and the questions they posed. The overall organization, the wording of the case statements and questions, and the methodology of the *Decretum* are all Gratian's, and his students clearly thought it worthwhile to preserve the substance of his arguments, but the words are not his. The first recension of the *Decretum* "may be a record of the first 'university course' in canon law ever taught,"<sup>23</sup> but the results of this experiment in authorship attribution suggest that we owe the written form of that record to the students rather than to their master. The strong evidence is that Gratian's direct involvement in the project came to an end, whether through death, declining health, or ecclesiastical promotion, before the first-recension *dicta* were preserved in their permanent written form.

---

<sup>23</sup> Winroth, *The Making of Gratian's Decretum*, 194.



## Outline of Chapters

Background, the *Decretum*, Gratian, Stylometry, Next steps.

## Note on the Title of the Dissertation

University policy required me to decide on the final title of my dissertation, “Distant Reading of Gratian’s *Decretum*,” years before I could possibly have known what the outcome of my research was going to be. In fact, another policy actually prohibited “proceed[ing] beyond the preliminary stage in the investigation of the topic” until my dissertation proposal had been approved, but the final title still had to be submitted as part of the proposal. The “distant reading” of the title is a nod to Franco Moretti’s book of the same name<sup>24</sup> and refers to my early plans to use MALLET to perform unsupervised topic modeling on the first and second recensions of the *Decretum* and to identify new topics added to the second recension by comparing the results. As the project evolved and the methodological emphasis shifted from unsupervised topic modeling to stylometry using principal components analysis, the original title became

---

<sup>24</sup> Franco Moretti, *Distant Reading* (London: Verso, 2013).



obsolete. If I were to choose a title today, “Computer-aided Close Reading of Gratian’s *Decretum*” would more accurately reflect the results of the project as delivered.

## Note on Translations

I have, wherever possible, supplied for each Latin passage quoted the corresponding passage from a published English translation.<sup>25</sup> In cases where no such translation was available, or I considered the available translation seriously misleading, I have supplied my own translation, indicated with the notation (trans. PLE). Special thanks to Atria A. Larson for her suggestions regarding the translation of the *Marturi placitum*.

---

<sup>25</sup> Katherine Ludwig Jansen, Joanna H. Drell, and Frances Andrews, eds., *Medieval Italy: Texts in Translation*, The Middle Ages Series (Philadelphia: University of Pennsylvania Press, 2009); Robert Somerville and Bruce Clark Brasington, eds., *Prefaces to Canon Law Books in Latin Christianity: Selected Translations, 500-1245* (New Haven, Conn: Yale University Press, 1998); and Augustine Thompson and James Gordley, trans., *The Treatise on Laws: (Decretum DD. 1-20)*, Studies in Medieval and Early Modern Canon Law, v. 2 (Washington, D.C: Catholic University of America Press, 1993) have been particularly helpful resources in this regard.



# Chapter 1

## Background

Around 1140, a revolutionary textbook transformed the study of canon law into a systematic academic discipline. It did not have an attribution of authorship. It did not even have a title. However, as it entered widespread circulation in the middle decades of the twelfth century, the new textbook gave rise to a considerable body of commentary, and some early glosses referred to it as the *Concordia discordantium canonum* (“Concordance of discordant canons”). The same glosses identified Gratian, an otherwise unknown teacher thought to have worked in Bologna, as its author.<sup>1</sup> As a

---

<sup>1</sup> Eight early manuscripts contain the gloss: “*Concordia discordantium canonum iuxta determinationem Gratiani episcopi quae in duas partes principaliter est divisa* (The agreement of disagreeing canons according to the determination of Bishop Gratian, which is principally divided into two parts.)” See below concerning whether, where, and when Gratian served as a bishop. Bl = Baltimore, Walters Art Gallery 777; Gt = Ghent, Bibl. der Rijksuniversiteit 55; Mt = Montecassino, Bibl. Abbaziale 66; Pf = Paris, Bibl. Nationale lat. 3884 I and II; Po = Pommersfelden, Bibl. des Grafen Schönborn 142 (2744); Ro = Rouen, Bibl. municipale E 21 (707); Sl = St. Paul im Lavant, Stiftsbibl. 25/1 (XXV.2.6); Tr = Trier, Stadtbibl. 906 (1141). Gt, Pf, and Tr indicate that the *Decretum* has two parts; Bl, Mt, Po, Ro, and Sl indicate that it has three. The earliest version of the *Decretum* to survive in more than one manuscript (the first recension) has two parts, while the most widely circulated version (the second recension) has three. Part III of the *Decretum (de Consecratione)* is outside the scope of this project, the intent of which is to compare the earliest (first-recension) version with the most widely circulated (second-recension) version for the purpose of determining authorship of those parts of the text traditionally attributed to Gratian himself. Part III is not

result, the book came to be commonly known as the *Decretum Gratiani* ("Gratian's *Decretum*").

The *Decretum* is not the kind of document, like the Fournier Register—the record of an episcopal inquisition into Albigensian heretics surviving in a single manuscript—that many non-specialists, with a journalistic eye for the exotic and the extreme, take to be the central concern of medieval studies.<sup>2</sup> The *Decretum* was, instead, both a normal and

---

found in the first recension, therefore direct comparison is not possible. In any event the class of text traditionally attributed to Gratian personally (the sayings or *dicta*) are absent from Part III. Therefore, it is not subject to comparative analysis using the techniques available for examining Parts I and II. John T. Noonan, "Gratian Slept Here: The Changing Identity of the Father of the Systematic Study of Canon Law," *Traditio* 35 (January 1979): 154. Noonan lists Mazarine 1289 (possibly a typo for Pm = Paris, Bibl. Mazarine 1287?) in place of Sl. Rudolf Weigand, "Frühe Kanonisten Und Ihre Karriere in Der Kirche," *Zeitschrift Der Savigny-Stiftung Für Rechtsgeschichte. Kanonistische Abteilung* 76 (1990): 135–55. Kenneth Pennington, "The Biography of Gratian, the Father of Canon Law," *Villanova Law Review* 59 (2014): 698–700, supplements Noonan's discussion of these eight glosses.

The author of the preface to the *Summa Parisiensis* incorrectly thought that Gratian himself had given the collection the title *Concordia discordantium canonum*: "Magister Gratianus, in hoc opere antonomasice dictus Magister, loco proemii talem suo præmisit libro titulum: Concordia discordantium canonum, in quo materiam et intentionem breviter exponit (Master Gratian, antonomastically called in this work 'the Master', in place of an introduction prefaced his book with the following title, *Concord of discordant canons*, by which he concisely sets forth subject matter and intention.)" Terence P. McLaughlin, ed., *The Summa Parisiensis on the Decretum Gratiani* (Toronto: Pontifical Institute of Mediaeval Studies, 1952), 1. English translation from Robert Somerville and Bruce Clark Brasington, eds., *Prefaces to Canon Law Books in Latin Christianity: Selected Translations, 500-1245* (New Haven, Conn: Yale University Press, 1998), 201.

<sup>2</sup> The manuscript is Vat. Lat. 4030. Jean Duvernoy, ed., *Le Registre d'inquisition de Jacques Fournier, évêque de Pamiers (1318-1325)*, Bibliothèque Méridionale. 2. Ser, t. 41 (Toulouse: É. Privat, 1965). See also [Inquisition](#)



a normative text in its own time, and for centuries thereafter. The *Decretum* survives in an unusually large number of manuscripts—around 150-200 from the twelfth century and around 600 for the medieval period as a whole.<sup>3</sup> The *Decretum* was the fundamental textbook for first-year university instruction in canon law through the seventeenth century. As the first volume of the *Corpus iuris canonici*, the *Decretum* remained valid law in the Roman Catholic Church until 1917 and exercised enormous influence over the 1917 and 1983 codifications of canon law that replaced the medieval *Corpus*.

## School, Faculty, University

Both the content and the form of the *Decretum* had an enormous impact on the civilization of medieval Europe. Gratian's methodological breakthrough transformed the study of canon law into a rigorous academic discipline. His imposition of a

---

Records of Jacques Fournier ("An on-going English translation of the Inquisition Records of Jacques Fournier, Bishop of Pamiers, France, 1318-1325") by Nancy P. Stork (SJSU).

<sup>3</sup> "My listing of more than 600 manuscripts containing the *Decretum* will appear in Kenneth Pennington and Wilfried Hartmann, eds., *History of Medieval Canon Law*, 11." Anders Winroth, *The Making of Gratian's Decretum* (Cambridge: Cambridge University Press, 2000), 122n2. This list has never made it into print to my knowledge. For a reasonably complete and current list, see the entry for [Gratian](#) in Kenneth Pennington's Bio-Bibliographical Guide to Medieval and Early Modern Jurists hosted by the Ames Foundation at Harvard University.



consistent interpretive framework (hermeneutic) on the inherited and internally contradictory mass of canonical texts transformed them into a coherent system of substantive law. A generation before Gratian, Ivo of Chartres (+1115) had proposed the possibility of such a program in the preface to his own canonical collections, but he left its application to his readers. Gratian demonstrated that Ivo's program could in fact be carried out in a truly systematic way and did so at an historical moment when the need for such a sweeping synthesis was particularly compelling. Gratian was working in the immediate aftermath of the Concordat of Worms (1122), which recognized the formal juridical independence of the Church from secular authority. Gratian's *Decretum* provided a comprehensive blueprint for the legal machinery by which the Church ultimately came to govern the many aspects of life in Christian society over which it claimed exclusive jurisdiction.

The form in which Gratian's intellectual achievement was transmitted had as great an impact as its content. The *Decretum* was arguably the first book in the European tradition written from the ground up as a textbook. Gratian's own teaching was probably typical of the relatively informal and unstructured environment of the medieval schools, organized around the activity of a gifted master, learned and



charismatic enough to gather a critical mass of students. The existence of a textbook, however, made it possible for someone other than the original master to teach the same, or a similar, course. A textbook that circulates widely enough becomes a *de facto* standard, and classes taught from it, by whomever, become likewise standardized. The *Decretum*, and classes taught from it or its abbreviations, attained the status of *de facto* standards across much of Europe very quickly—within the two decades between 1140 and 1160. I believe that the new evidence previewed above suggests that while we owe the intellectual substance of at least the first recension of the *Decretum* to Gratian, we owe the words in which that substance was expressed to his students.<sup>4</sup>

The compilation and circulation of the *Decretum* led directly to the formation of faculties of canon law across Europe. Faculties came into being as an institutional response to a new problem raised by standardized classes taught from standardized textbooks: what are the requisite qualifications for the teaching of canon law? In some respects, early faculties were similar to other medieval guilds that regulated competition between

---

<sup>4</sup> Gratian provided at minimum the outline for second part of the *Decretum*. How closely the logic and arguments of the *dicta* follow Gratian's classroom presentation, as opposed to being the student's own, cannot be determined using statistical authorship attribution methods.



masters in places where material and political conditions favored the concentration of many masters practicing the same craft in one place.<sup>5</sup> Faculties and craft guilds, however, differed in one important respect: while the most important product of the masters of the guild of shoemakers was shoes, the most important product of the masters of the faculty of canon law was, at least in the first generation, the next generation of masters of the faculty of canon law. The distinctive innovation of the faculty was the conferring of credentials or qualifications (to call them degrees is perhaps anachronistic) on students who completed a prescribed sequence of standardized courses, defined in such a way that the requirements for completion for students coincided with those for admission to teaching. The faculty was a crucial intermediate stage in the institutional evolution from school to university, one that has been obscured by the retrospective myth-making of famous universities, which, once established, sought to enhance the lustre of their names by projecting their origins as far

---

<sup>5</sup> For example, see R. W. Southern, *Scholastic Humanism and the Unification of Europe*, vol. 1 (Oxford, UK ; Cambridge, Mass., USA: Blackwell, 1995), 310–18, for a discussion of the role that material and political conditions played in the rise of Bologna and Paris as major academic centers.



as possible into the past.<sup>6</sup> It is perhaps going too far to claim that the *Decretum* was the textbook that created the university. It is not, however, going too far to recognize

---

<sup>6</sup> "By the first decade of the thirteenth century the private schools of canon law had become part of the corporate structure of the universities in all three places [Bologna, Paris, and Oxford]." James A. Brundage, "The Teaching and Study of Canon Law in the Law Schools," in *The History of Medieval Canon Law in the Classical Period, 1140-1234: From Gratian to the Decretals of Pope Gregory IX*, ed. Wilfried Hartmann and Kenneth Pennington, *History of Medieval Canon Law* (Washington, D.C: Catholic University of America Press, 2008), 98–99. Brundage here somewhat overstates the degree of institutional evolution the proto-universities had attained at this time. The corporate structure of the university had fully emerged by the first decade of the thirteenth century only at Bologna. It did not do so at Paris and Oxford until the second decade of the century. The standard handbooks on the history of the university in medieval Europe, Hastings Rashdall, *The Universities of Europe in the Middle Ages*, A new edition, ed. by F. M. Powicke and A. B. Emden (London: Oxford Univ. Press, 1936). and Hilde de Ridder-Symoens, ed., *Universities in the Middle Ages*, History of the University in Europe, v. 1 (Cambridge [England] ; New York: Cambridge University Press, 1992), are less helpful on this point than might be hoped.

Bologna: "Bologna vies with Paris for the title of oldest and most distinguished university of medieval Europe, but whereas Paris was essentially a guild of masters, Bologna was in origin an association of students. The *universitas scholarium* emerged around 1190 as an organization of the non-Bolognese law students resident in the city, and rapidly developed a complex constitution rather like that of contemporary communes." J. K. Hyde, "Bologna, University of," in *Dictionary of the Middle Ages*, ed. Joseph R. Strayer, vol. 2 (New York: Scribner, 1982), 311.

Oxford: "The origins of the university are shrouded in obscurity." Damian Riehl Leader, "Oxford University," in *Dictionary of the Middle Ages*, ed. Joseph R. Strayer, vol. 9 (New York: Scribner, 1982), 320. "Oxford probably developed in the twelfth century out of an informal group of masters teaching near St. Mary's Church. Among the first known masters was Theobald of Étampes, who was lecturing before 1100 (ca. 1095) and continued until about 1125. This group of scholars seems to have increased markedly following Henry II's prohibition of foreign study in 1170, and by 1185 Gerald of Wales (Giraldus Cambrensis) recorded that they were organized into several faculties (*doctores diversarum facultatum*), which, although likely an exaggeration, indicates there was a varied group of scholars. This development is implied by several other late twelfth-century authorities as well." Leader, 320. "The university was first recognized as a legal corporation by a legatine ordinance of 1214, ending a dispersal of the university following a riot with the townsmen in 1209." Leader, 320.



that its existence was a necessary precondition for the emergence of the faculty of canon law.

In creating his new textbook, Gratian blazed a trail for canon law that theology followed about two decades later. The appearance around 1160 of Peter Lombard's *Sentences* served as the catalyst for a similar transformation of theology into a rigorous academic discipline taught from a standardized textbook in the context of a formally organized curriculum by a degree-granting faculty, and for the displacement of more loosely organized regional schools, such as the one at Laon, by the preeminent theology faculty of Paris.

---

Paris: "The school of Notre Dame Cathedral on the Cité, and the Abbey of St. Victor and the collegiate church of Ste. Geneviève on the Left Bank of the Seine had already by the end of the twelfth century attracted such famous masters as Hugh of St. Victor, Peter Abelard, Stephen of Tournai, and Peter Lombard. In 1180 Alexander Neckham was able to write of an 'honorable society of masters' who were teaching arts, theology, canon law, and medicine, disciplines that would later form the four faculties of the university. At this time, the right to teach (*licentia docendi*) and therefore to become a master was given out by the chancellor of the cathedral chapter of Notre Dame." Astrid L. Gabriel, "Paris, University of," in *Dictionary of the Middle Ages*, ed. Joseph R. Strayer (New York: Scribner, 1982), 408. "The masters organized themselves into an association between 1180 and 1210, later called *universitas magistrorum et discipulorum*, and both kings and popes granted it protective privileges." Gabriel, 408. "In the beginning, Paris church officials opposed the formation of this independent society, which claimed legal status and its own seal. The papacy, however, sided with the university. Rules regulating its operation, called 'statutes', were granted in 1215 by Robert Courson, papal legate, and these dealt specifically with curriculum and textbooks." Gabriel, 408.



In the first decades of the thirteenth century, both the faculties of canon law and the *Decretum* settled into their permanent places in the organization of medieval academic life, as the faculties of canon law became part of the university (a corporation encompassing several faculties such as those of arts, law, and theology), while the focus of legal scholarship in canon law gradually shifted from the “old law” of Gratian to the “new law” of papal decretals.<sup>7</sup>

Recent complaints about the corporatization of the university notwithstanding, *universitas* means corporation.<sup>8</sup> We think of a corporation as a commercial entity, but medieval legal thinkers looked upon the corporation at a more basic level, as a means for legally creating an artificial, collective person, and then investing that “person” with rights and privileges that could otherwise only be held by a natural person or persons, such as the right to own property and the standing of a litigant in courts of law. The

---

<sup>7</sup> This distinction between old and new law comes from the preface of *Bernardus Papiensis* (†1213) to his Breviary of *Extravagantia* or First Compilation, c.1191. Somerville and Brasington, *Prefaces to Canon Law Books in Latin Christianity*, 219, 230–31

<sup>8</sup> Jacques Verger, “Patterns,” in *Universities in the Middle Ages*, ed. Hilde de Ridder-Symoens, History of the University in Europe, v. 1 (Cambridge [England] ; New York: Cambridge University Press, 1992), 37–41.



most fundamental right of a corporation, however, was the right of self-governance, including the right of the corporation to select its own leadership. The classic medieval example of corporate self-governance was the right of a cathedral chapter to elect a bishop. For an academic faculty, the exercise of the right of corporate self-governance took the particular form of determining the qualifications for teaching, something often contested by the local bishop.<sup>9</sup> The incorporation of several faculties as a university provided a form of organization well-adapted for achieving both permanence and

---

<sup>9</sup> "Four steps would seem to have been pre-eminently necessary to give to mere customary meetings of masters for the initiation of new members or similar purposes the character of a definite and legally recognized corporation: (1) the reduction of their unwritten customs to the form of written statutes or by-laws, (2) the recognition or (if authoritative recognition was unnecessary) the exercise of the right to sue and be sued as a corporation, (3) the appointment of permanent common officers, (4) the use of a common seal." Rashdall, *The Universities of Europe in the Middle Ages*, 299.

Medieval law knew no regular procedure for incorporation. Every effort to form a corporation therefore involved an intensive and frequently protracted effort to persuade the relevant authority, whether pope, emperor, king or commune, to grant the privilege. In addition, medieval legal thought tended to conceive of rights rather more concretely than modern legal systems do. Every right or privilege was thought of as coming at the expense of some other right-holder, and therefore attempts by the early universities to form corporations were contested by those, usually local bishops, at whose expense the exclusive right to confer a license to teach was being secured.

Although the right of a faculty or university to own property might be considered necessary to realizing any aspirations toward institutional permanence, in practice, academic corporations avoided property ownership in the twelfth and thirteenth centuries. Indeed, the fact that they tended to rent rather than own the lands and buildings they used could become a powerful weapon in the event of "town vs. gown" conflict, allowing the university to credibly threaten to relocate elsewhere, as the masters and students of Oxford did in 1209. Rashdall, 406.



independence from local ecclesiastical authority through the accumulation of legal rights and privileges. The faculties ensured their own long-term survival by incorporating within the larger university.

The *Decretum* remained the standard textbook for the prerequisite first course in canon law throughout the Middle Ages. Around 1190, however, while the evolution from faculty to university was still in progress, the focus of academic canonists shifted away from Gratian and toward the new jurisprudence of papal decretals. Scholarly attention devoted to the *Decretum* after 1190 tended to focus on improving its utility as a teaching text. Between 1214 and 1217, [Johannes Teutonicus](#) (c.1170-1245) compiled more than a half-century's worth of accumulated commentary on Gratian into the *Glossa Ordinaria*,<sup>10</sup> and [Bartholomeus Brixensis](#) (fl.1234-1258) revised the ordinary gloss between 1234 and 1241.<sup>11</sup> By then, the interest of thirteenth-century teachers and students centered on a

---

<sup>10</sup> Rudolf Weigand, "The Development of the *Glossa Ordinaria* to Gratian's *Decretum*," in *The History of Medieval Canon Law in the Classical Period, 1140-1234: From Gratian to the Decretals of Pope Gregory IX*, ed. Wilfried Hartmann and Kenneth Pennington, *History of Medieval Canon Law* (Washington, D.C.: Catholic University of America Press, 2008), 82-86.

<sup>11</sup> Weigand, 88-91.



new advanced second course in canon law taught from Raymond de Peñafort's *Decretales Gregorii IX* (1234), edited from five earlier compilations of papal decretals that had become a teaching library of decretal law. The *Decretum* retained its status as a classic textbook but never again held the intellectual attention of the discipline in the way it had during the second half of the twelfth century.

The emergence of faculties of canon law in the second half of the twelfth century was a feature of a wider intellectual and social revolution. There had been a revival in Bologna of the formal study of Roman law during the second and third decades of the twelfth century. The immediate impetus for the revival was the recovery of the great *libri legales* from which law had been taught in the late Roman world, most important, the *Digest* of Justinian, because it provided the pattern for jurisprudence as a systematic intellectual activity.<sup>12</sup> The early Romanists therefore did not need to create textbooks for their students as Gratian had to for his.

---

<sup>12</sup> The Digest was not recovered all at once, but in three parts: the *Digestum vetus* (Dig. 1.1-24.2), the *Infortiatum* (Dig. 24.3-38.17), and the *Digestum novum* (Dig. 39.1-50.17). Manlio Bellomo, *The Common Legal Past of Europe: 1000-1800, Studies in Medieval and Early Modern Canon Law*, v. 4 (Washington, D.C.: Catholic University of America Press, 1995), 62–63. Wolfgang P. Müller, “The Recovery of Justinian’s



The twelfth-century revolution in legal science did not take place in a vacuum. It took place because of the emergence, especially in Italy, of an increasingly urban and commercial society with many moving parts. Such a society needed a correspondingly complex and sophisticated legal system, and, because it had much in common, both culturally and materially, with the society of the late Roman world, it was able to borrow extensively from Roman law.

The emergence of distinct faculties of arts, law, and theology in the middle decades of the twelfth century gave rise to a rapidly proliferating number of major universities throughout Europe, recognizably continuous in form with those of the present day. The

---

Digest in the Middle Ages," *Bulletin of Medieval Canon Law* 20 (1990): 1–29. Michael H. Hoeflich and Jasonne M. Grabher, "The Establishment of Normative Legal Texts: The Beginnings of the *Ius Commune*," in *The History of Medieval Canon Law in the Classical Period, 1140-1234: From Gratian to the Decretals of Pope Gregory IX*, ed. Wilfried Hartmann and Kenneth Pennington, *History of Medieval Canon Law* (Washington, D.C: Catholic University of America Press, 2008), 5. The standard modern edition of the *Digest* can be found in Paul Krüger and Theodor Mommsen, eds., *Corpus Iuris Civilis* (Berolini: apud Weidmannos, 1928). There is a recent English translation of the *Digest* in Alan Watson, ed., *The Digest of Justinian* (Philadelphia, Pa: University of Pennsylvania Press, 1985). [See review in Charles Donahue, "On Translating the "Digest"," *Stanford Law Review* 39, no. 4 (1987): 1057–77.] There is an excellent facsimile edition of the *Codex Florentinus* (Firenze, Biblioteca Laurenziana, *sine numero*), the sixth- or seventh-century manuscript that served at least indirectly as the basis for the eleventh- and twelfth-century recovery of the *Digest* in Alessandro Corbino and Bernardo Santalucia, eds., *Justiniani Augusti Pandectarum Codex Florentinus* (Firenze: Olschki, 1988).



existence of standardized textbooks such as Gratian's *Decretum* and later Peter Lombard's *Sentences* was an absolutely essential precondition for the formation of academic faculties organized along formal disciplinary lines, offering standardized courses of instruction for their students. Those faculties in turn supplied the foundation for the medieval, and, by extension, the modern European university. Gratian's *Decretum* can be seen from the point of view of this series of historical developments as the textbook that set the stage for the emergence of the university.

## Roman Law

There is good evidence that canon law emerged as a distinct, formally organized, academic discipline about two decades after Roman law and about two decades before theology. Therefore, any information that allows us to refine our knowledge of the dates by which those disciplines, but especially Roman law, emerged has considerable evidentiary value for the dating of the emergence of canon law as well. The current consensus among historians is that the teaching of Roman law got under way as an organized activity in Bologna between 1100 and 1110 and that the teaching of canon law got under way, also at Bologna, a little later, perhaps between 1120 and 1130.



It is first necessary to distinguish the body of Roman law formally codified in the sixth century under the emperor Justinian (+565) from the forms of Roman law that had survived through continuous use and adaptation in southern Europe from late antiquity into the eleventh century. Early medieval Europe had not forgotten Roman law, and, in particular, the *Lex Romana Visigothorum* or *Breviary of Alaric* (506), derived from the Theodosian Code (438), was a major source of Roman law concepts and terminology throughout the early middle ages. The codification produced by Justinian's commissioners under the guidance of Tribonian (+545?) received limited reception in the sixth century in those parts of Italy then under Byzantine control. However, it disappeared from view around the beginning of the seventh century. The last known reference to the *Digest* in the Latin-speaking and -writing West is found in a letter of pope Gregory I in 603.<sup>13</sup>

---

<sup>13</sup> The reference is to Dig. 48.4.7.3. "XIII, 49 Iohanni defensori exempla legum tradit, secundum quas iudicet" in Dag Ludvig Norberg, ed., *S. Gregorii Magni Registrum Epistularum, Libri VIII-XIV, Appendix*, Corpus Christianorum. Series Latina, 140 A (Turnholti: Brepols, 1982), 1058–64. Translated as "13.49 Gregory to John, his defender, going to Spain. An example of a law," August 603, in John R. C. Martyn, trans., *The Letters of Gregory the Great*, Mediaeval Sources in Translation 40 (Toronto: Pontifical Institute of Mediaeval Studies, 2004), 863–66.



What was new in Italy from the late eleventh century on was the revival of the study and application of Roman law as codified in the *Corpus iuris civilis* of Justinian: the *Institutes*, the *Codex*, the *Digest* or *Pandects*, and the *Novels*. (The *Novels*, unlike the other volumes of Justinian's *corpus*, were originally written in Greek, and were known in this period in the form of a Latin translation, the *Authenticum*.) The first documentary evidence for this revival is a reference to the *Digest* found in the *Marturi placitum*, the record of a court case held in 1076 in present-day Poggibonsi in Tuscany.<sup>14</sup>

---

<sup>14</sup> Cesare Manaresi, ed., *I Placiti Del "Regnum Italiae."*, Fonti Per La Storia d'Italia, Pubblicate Dall' Istituto Storico Italiano Per Il Medio Evo 97 (Roma: Tip. del Senato, 1955), 333–35. (no. 437) The *Marturi placitum* does not explicitly identify the passage in the *Digest* to which the citation refers, but the fact that the legal principle being invoked is *restitutio in integrum* is by itself sufficient to narrow it down to *Dig.* 4. Radding and Ciaralli identify the passage as *Dig.* 4.6.26.4 (with some wording borrowed from *Dig.* 4.1.5). Charles Radding and Antonio Ciaralli, *The Corpus Iuris Civilis in the Middle Ages: Manuscripts and Transmission from the Sixth Century to the Juristic Revival*, Brill's Studies in Intellectual History, v. 147 (Leiden ; Boston: Brill, 2007), 183–84. Nordillus, the judge, decided the case “*lege Digestorum libris inserta considerata* (after he considered the law inserted into the books of the *Digest*)” [translation suggested by Atria Larson]. The translation turns on the interpretation of the participial adjective *inserta*. “I think the straight-up ‘inserted’ makes sense here. The writer would seem to be noticing that the *Digest* is not itself *lex* but includes commentary on the *lex*, the written law, and refers to and inserts many elements of *lex* throughout.” (Atria Larson, email to Paul Evans, July 21, 2017.) The alternative is to translate *inserta* as “introduced,” describing a passage from the *Digest* that an advocate for one or the other of the litigants brought to the attention of the court in support of their case as having been “introduced.” For an English translation of the *Marturi placitum*, see Bruce Clark Brasington, ed., *Order in the Court: Medieval Procedural Treatises in Translation*, Medieval Law and Its Practice 21 (Leiden: Brill, 2016). (See [West Texas A&M University: Order in the Court: Medieval Procedural Treatises in Translation](#).)



Because of the title *Corpus iuris civilis*, “the body of civil law,” Roman law was also referred to as civil law, and students, teachers, and practitioners of Roman law were referred to as civilians. Gratian, following Isidore of Seville (†636), used a much older definition: “*Ius civile est, quod quisque populus vel civitas sibi proprium divina humanaque causa constituit* (Civil law is what each people and each commonwealth establishes as its own law for divine or human reasons).”<sup>15</sup> Gratian, again following Isidore, referred to Roman law as the law of the Quirites (*Ius Quiritum*).<sup>16</sup>

The scholarly consensus that the teaching of Roman and canon law were going concerns in Bologna by around 1110 and 1130 respectively has had two notable recent dissenters, Charles Radding and Anders Winroth. Radding has proposed that the epicenter of the revival of the study of Roman law was not Bologna, but Pavia, and places the date somewhat earlier, in the late eleventh century. Radding’s claim is that research into

<sup>15</sup> D.1 c.8. *Decretum Gratiani*, First Recension, edition in progress. Anders Winroth, 3/21/2017, 2. English translation from Augustine Thompson and James Gordley, trans., *The Treatise on Laws: (Decretum DD. 1-20)*, Studies in Medieval and Early Modern Canon Law, v. 2 (Washington, D.C: Catholic University of America Press, 1993), 7.

<sup>16</sup> D.1 c.12. *Decretum Gratiani*, First Recension, edition in progress. Anders Winroth, 3/21/2017, 3.



Roman law developed from its having been used as an analytical tool for the explication of the central text of the Lombard legal tradition, the *Liber Legis Langobardorum* or *Liber Papiensis*. Radding's argument has not met with widespread acceptance.<sup>17</sup>

Anders Winroth is by far the more consequential dissenter from the consensus. He draws attention to the underdeveloped way in which Gratian uses Justinianic Roman law in what Winroth calls the first recension, the earliest version of the *Decretum* to survive in more than one manuscript.

There are only three places in the first recension of the *Decretum* that incorporate material taken directly from Justinianic Roman law sources: C.2 q.6 c.28 (an excerpt from the *Authenticum*), and C.15 q.3 cc.1-3 and C.15 q.3 c.4 (three excerpts from the *Codex* and four from the *Digest*).<sup>18</sup>

<sup>17</sup> Charles Radding, *The Origins of Medieval Jurisprudence: Pavia and Bologna, 850- 1150* (New Haven: Yale University Press, 1988).

<sup>18</sup> Winroth, *The Making of Gratian's Decretum*, 146–48.



Gratian does not deploy the proper Justinianic shades of meaning to the concepts of ownership and possession when discussing the issue of *restitutio in integrum* in the first-recension version of C.3 q.1 d.p.c.2. Instead, he uses concepts and vocabulary that, Winroth argues, derive from pre-Justinianic Roman law sources such as the *Lex Romana Visigothorum*.<sup>19</sup>

Furthermore, there are places in the first recension where Gratian, in Winroth's opinion, would have reached different conclusions had he used Justinianic Roman law. He points to Gratian's treatment of false (or at least unsuccessful) accusers in C.2 q.3, and the penalties, including infamy, incurred by them. Gratian's analysis —he distinguishes three categories of false accusers based on material indirectly derived from the *Lex Romana Visigothorum*— is incompatible with the well-developed Justinianic jurisprudence on the subject.<sup>20</sup>

---

<sup>19</sup> Winroth, 148–51.

<sup>20</sup> Winroth, 153–56.



But there is one question—C.15 q.3—in the first recension of the *Decretum* in which Gratian makes substantive use of Justinianic Roman law concepts and vocabulary as a part of his argument in the *dicta* and does so more or less correctly in Winroth's judgment. This is significant from Winroth's point of view, because it demonstrates that Gratian did not have an ideological objection to the use of Roman law, as Vetulani and Chodorow argued.<sup>21</sup> Rather, Winroth argues that where Gratian either failed to use Roman law, or used it incorrectly, it was because he lacked the knowledge of that tradition.<sup>22</sup>

For reasons that I will discuss more fully below, Winroth is convinced that the first recension of the *Decretum* cannot have been completed or entered circulation earlier than around 1140. The underdeveloped way in which Justinianic Roman law is used in

---

<sup>21</sup> Chodorow explains Gratian's "deletion" of Roman law as a consequence of what he sees as Gratian's adherence to the party of the papal chancellor Haimeric and proponents of the Concordat of Worms (1122). Stanley Chodorow, *Christian Political Theory and Church Politics in the Mid-Twelfth Century; the Ecclesiology of Gratian's Decretum*, Publications of the Center for Medieval and Renaissance Studies, U.C.L.A., 5 (Berkeley: University of California Press, 1972), 60–63.

<sup>22</sup> Winroth, *The Making of Gratian's Decretum*, 151–53.



the first recension therefore leads Winroth to argue that the study and teaching of Roman law in Bologna had not progressed nearly as far by 1140 as the consensus holds.

Some corollaries to Winroth's argument that Roman law teaching got off to a late start are fairly far out of the mainstream. A notable example is his dating of the famous letter of an anonymous monk to abbot B (probably Bernard III Garin) of Saint-Victor, Marseille —the content of which Winroth has so memorably summarized as “[m]y donkey is dead, so I am going to law school”—to the 1180s rather than the 1120s. But it is extremely difficult to imagine that a letter that shows no apparent awareness that either civil (Roman) or canon law were being taught in Bologna can have been written in the 1180s.<sup>23</sup>

---

<sup>23</sup> Anders Winroth, “Law Schools in the Twelfth Century,” in *Mélanges En L’honneur d’Anne Lefebvre-Teillard*, ed. Bernard d’ Alteroche et al. (Paris: Éd. Panthéon-Assas, 2009), 1060. For the text of the letter, see Jean Dufour, Gérard Giordanegno, and André Gouron, “L’attrait Des ‘Leges’: Note Sur La Lettre d’un Moine Victorin (Vers 1124/1127),” *Studia et Documenta Historiae et Iuris* 45 (1979): 504–29. For an English translation of the letter, see Anonymous, Letter to Abbot B[ernard III] of Saint-Victor, Marseille (1124/27) in Katherine Ludwig Jansen, Joanna H Drell, and Frances Andrews, eds., “Roman Law and Legal Studies: Three Texts (ca. 1124-66),” in *Medieval Italy: Texts in Translation*, trans. Sean Gilsdorf, The Middle Ages Series (Philadelphia: University of Pennsylvania Press, 2009), 167–72.



Kenneth Pennington has pointed to the *Authentica Habita* (1155/1158) issued by the emperor Frederick I (†1190), which has traditionally been taken as extending imperial protection to the law school of Bologna, its faculty, and its students.<sup>24</sup> Pennington makes the common-sense point that a privilege like this is unlikely to have been granted to an

---

<sup>24</sup> “[O]mnibus, qui causa studiorum peregrinantur, scolaribus et maxime divinarum atque sacrarum legum professoribus hoc nostre pietatis beneficium indulgemus, ut ad loca, in quibus litterarum excentur studium, tam ipsi quam eorum nuntii veniant et habitent in eis securi.” Heinrich Appelt, ed., *Die Urkunden Der Deutschen Könige Und Kaiser*, vols. X, 2, *Monumenta Germaniae Historica. Diplomata Regum et Imperatorum Germaniae* (Berlin: Weidmannsche Verlagsbuchhandlung, 1979), 39. “[W]e bestow this pious gift upon all who travel for the purpose of study, students and especially teachers of divine and sacred laws: namely, that they as well as their messengers may travel safely to the place where they are engaged in the study of letters and safely dwell there.” Jansen, Drell, and Andrews, “Roman Law and Legal Studies,” 167–72.

The date of *Authentica Habita* is uncertain. Frederick ordered the constitution inserted into the *Codex* as an act of imperial propaganda. No medieval manuscripts of the *Codex* give a date for the proclamation. Early modern print editions of the *Codex*, however, associated it with the Diet of Roncaglia in November 1158, and modern editions have retained the ascribed date. Kenneth Pennington, “The Beginnings of Law Schools in the Twelfth Century,” in *Les écoles Du XIIe Siècle*, ed. Cédric Giraud (Leiden: Brill, 2018). (p.20, fn.53 in MS Word version.) The uncertainty arises because a contemporary poem, the *Carmen de gestis Frederici I. Imperatoris in Lombardia* gives a first-hand account of an incident that took place near Bologna in 1155 in which Frederick, in response to a petition brought by a deputation of doctors and scholars from the city, issued a proclamation matching the description of *Authentica Habita*. Some modern scholars attempt to split the difference by accepting the November 1158 date for the version of the text that we have, but characterizing it as a re-issue of the 1155 proclamation described the *Carmen de gestis*. There appears to be no good reason to believe that *Authentica Habita* was promulgated twice, in 1155 and 1158, as opposed to just once, in 1155. In any event, the uncertainty (within a narrow and circumscribed range) about the date *Authentica Habita* was promulgated does not take anything away from the validity of the point Pennington is making.



institution just getting off the ground.<sup>25</sup> Winroth counter-argues that the privilege mentions neither Bologna nor those involved in legal studies specifically.<sup>26</sup>

I agree with the consensus that the teaching of Roman and canon law were going concerns in Bologna by around 1110 and 1130 respectively, but the reader should keep in mind that Winroth, a scholar whose point of view is not to be lightly dismissed, thinks otherwise.

In recent years, it has become obligatory to mention at this point in any discussion of the revolution in legal science the case of one Mengho, a resident of Bologna who was hanged in December 1299 after having been tortured into confessing to the theft of some

---

<sup>25</sup> Kenneth Pennington, “A Short History of Canon Law from Apostolic Times to 1917,” 2002, 17, <http://legalhistorysources.com/Canon%20Law/PenningtonShortHistoryCanonLaw.pdf>.

<sup>26</sup> It is true that *Authentica Habita* does not mention Bologna explicitly, but instead refers to “the place where they are engaged in the study of letters.” But *Authentica Habita* does specifically mention “students and especially teachers of divine and sacred laws” so it is not clear to me why both Pennington and Winroth contest the latter point. “It [*Authentica Habita*] did not mention Bologna nor law students.” Pennington, “The Beginnings of Law Schools in the Twelfth Century.” (p.18 in MS Word version.) “The law [*Authentica Habita*] does not mention Bologna at all, nor does it mention studies of law.” Anders Winroth, “The Teaching of Law in the Twelfth Century,” in *Law and Learning in the Middle Ages*, ed. Mia Münster-Swendsen and Helle Vogt (Copenhagen: DJØF, 2006), 41–62. (pp.7-8 in MS Word version.) “The Teaching of Law in the Twelfth Century” has since been superseded by “Law Schools in the Twelfth Century,” which does not mention *Authentica Habita*.



bolts of fabric. There was not in 1299, nor is there now, much doubt as to Mengho's guilt—a search of his residence was carried out during the course of the investigation into the crime and he was found to be in possession of the stolen goods, a state of affairs that he was unable to credibly explain. Neither can it be seriously doubted that Mengho's case was conducted according to the due process standards prevailing in that time and place. Full proof necessary to convict a criminal defendant (*reus*) under Roman law required either a confession or the testimony of two witnesses, testimony unlikely to be forthcoming in the case of a crime of stealth, such as Mengho's, committed at night. In the absence of a voluntary confession or the testimony of witnesses, torture of the defendant in order to obtain an involuntary confession was simply the next routine step of due process. Indeed, Hermann Kantorowicz (†1940) brought the story of Mengho to the attention of medieval legal historians, and it remains well known today, precisely because Kantorowicz considered it a textbook example of inquisitorial (as opposed to adversarial) procedure under the *ordo iudicarius* derived from Roman law.<sup>27</sup>

---

<sup>27</sup> Hermann Kantorowicz, *Albertus Gandinus Und Das Strafrecht Der Scholastik*, vol. 1: Die Praxis (Berlin: J. Guttentag, 1907), 203–18. See also Kenneth Pennington, “Law, Criminal Procedure,” in *Dictionary of the Middle Ages. Supplement 1*, ed. William C. Jordan (New York: Scribner, 2004), 309–20.



The fact that Mengho was tortured and executed was not the point of the story for Kantorowicz, but it is for us. In the twenty-first century, the story of Mengho is brought up to make a particular point. There is a powerful scholarly tradition going back at least to Charles Homer Haskins of presenting the revolution in legal science as an intellectual triumph that took place in the classroom.<sup>28</sup> All revolutions, including the revolution in legal science, have winners and losers, and some of the losers we do not feel very sorry for.<sup>29</sup> But what happened in the classroom ultimately had consequences in the courtroom, consequences that were enacted on the rights and property, on the bodies and lives, of women and men in the real world.

---

<sup>28</sup> Charles Homer Haskins, *The Renaissance of the Twelfth Century* (Cambridge: Harvard University Press, 1927).

<sup>29</sup> R.I. Moore observes that the establishment of the new order brought about by the legal revolution “required another change, no less profound: the replacement of warriors by literate clerks as the agents of government and the confidants of princes.” R. I. Moore, *The Formation of a Persecuting Society: Authority and Deviance in Western Europe, 950-1250*, 2nd ed (Malden, MA: Blackwell Publishing, 2007), 128.



# Chapter 2

## The *Decretum*

The *Decretum* represents a towering intellectual achievement of the renaissance of the twelfth century that ushered in the high middle ages in Europe. But whose achievement was it? In particular, is the *Decretum* the achievement of a single author or was it the product of serial or even collaborative authorship? Because so little is reliably known about the historical Gratian, and because almost everything that previous generations thought was known about Gratian has proved, on closer examination, to be myth,<sup>1</sup> it is best to make our first approach (*accessus*) to the author through his text.<sup>2</sup> The approach is made more difficult by the fact that, like many modern university textbooks, the *Decretum* underwent at least one major revision and was probably the product of a process of continuous revision.

---

<sup>1</sup> John T. Noonan, "Gratian Slept Here: The Changing Identity of the Father of the Systematic Study of Canon Law," *Traditio* 35 (January 1979): 145–72.

<sup>2</sup> "Given this lack of reliable extrinsic evidence, our best source of evidence about Gratian is his book(s)." John C. Wei, *Gratian the Theologian*, Studies in Medieval and Early Modern Canon Law, Volume 13 (Washington, D.C: Catholic University of America Press, 2016), 33.

In 1996, Anders Winroth discovered that four surviving twelfth-century manuscripts—Florence, Biblioteca Nazionale Centrale, Conv. Soppr. A. 1.402 (Fd); Barcelona, Arxiu de la Corona d’Aragó, Santa Maria de Ripoll 78 (Bc); Paris, Bibliothèque Nationale de France, nouvelles acquisitions latines 1761 (P); and Admont, Stiftsbibliothek 23 and 43 (Aa)—preserve the text of what Winroth called the first recension of the *Decretum*.<sup>3</sup> In 1998, Carlos Larrainzar identified Paris, Bibliothèque Nationale de France, latin 3884 I, fo. 1 (Pfr) as a one-page fragment of a first-recension

---

<sup>3</sup> Manuscripts of Gratian’s *Decretum* are frequently referred to in recent scholarly literature by two-letter abbreviations or *sigla* derived from their shelfmarks. A reference to Sg (to use a particularly controversial example) is much more compact and memorable than one to Sankt Gallen, Stiftsbibliothek 673. Rudolf Weigand (†1998) generated the original list of *sigla* in the course of his groundbreaking study of early glossed manuscripts of the *Decretum*, *Die Glossen Zum "Dekret" Gratians: Studien Zu Den Frühen Glossen Und Glosskompositionen*, Studia Gratiana 25–26 (Rome, 1991). In the context of Weigand’s study early means before 1216, when Johannes Teutonicus (†1245) finalized the ordinary gloss, or standardized commentary, on the *Decretum*. Manuscripts falling outside the original scope of Weigand’s investigation are now assigned *sigla* following the pattern he set. For example, Paris, Bibliothèque Nationale de France, nouvelles acquisitions latines 1761, although an extremely important and early (possibly the earliest) manuscript of the *Decretum*, was not glossed, and was therefore not assigned a *siglum* by Weigand. On the advice of Weigand, Winroth supplied the *siglum* P by which the manuscript is now commonly identified: “it was Rudolf Weigand who came up with it. I had asked him what *siglum* this ms should have, since I thought all of the two-letter ones that started with a P had been taken. He said to use only P, ‘because it is such an important manuscript.’” (Anders Winroth email to Paul Evans, October 1, 2018) (Although the *sigla* in Weigand’s *Handschriftenliste* adhere consistently to the two-letter convention, some of the *sigla* supplied by other scholars deviate from it, e.g., P, Pfr.)



manuscript of the *Decretum*.<sup>4</sup> In 2011, Atria Larson discovered that München, Bayerische Staatsbibliothek, lat. 22272 (Mw) contains an abbreviation of the first recension of the *Decretum*.<sup>5</sup> The four first-recension manuscripts Aa, Bc, Fd, and P were previously thought to have been abbreviations and are by one imperfect but frequently-invoked metric about half the length of the vulgate *Decretum*, 1,860 as opposed to 3,945 canons.<sup>6</sup>

Winroth's discovery of the first recension suggests one obvious way in which it might be problematic to refer simplistically to Gratian as the author of the *Decretum*, as I have done thus far. Winroth rejects any presumption of monolithic authorship on

<sup>4</sup> Anders Winroth, *The Making of Gratian's Decretum* (Cambridge: Cambridge University Press, 2000), 32.

<sup>5</sup> Atria A. Larson, "An Abbreviatio of the First Recension of Gratian's Decretum in Munich?" *Bulletin of Medieval Canon Law* 29 (2011): 51–118.

<sup>6</sup> "The second recension contains 3,945 canons (including the paleae) in the editions. The first recension contains only 1,860 canons (47 percent)." Winroth, *The Making of Gratian's Decretum*, 122. The division of distinctions and questions into canons is a matter of editorial opinion and therefore unavoidably conventional. The generally-cited number of 3,945 comes from Friedberg's introduction (edF 1.xii). Furthermore, many canons include more than one canon text, and the conventional number obfuscates the distinction between the canon defined as a container or package and the canon text or texts included within it. The Reuter and Silagi e-text of the Friedberg edition has 3,848 canons (identified by the OCP <4 N> tag, where N is the canon number) containing 4,394 distinct canon texts (identified by the OCP <T T> tag). (Note: egrep "<T T>.\*<T T>" edF.txt to find the three lines where the canon text tag occurs twice on the same line.)



stratigraphic grounds. He argues that the first recension is “coherent and complete” and that its author intentionally released it into circulation as a “finished product.”<sup>7</sup> (In other words, it did not just prematurely escape into the wild like parts of Augustine’s *de Trinitate*.) Winroth thinks of the first and second recensions as distinct textual layers and argues that two different authors, Gratian 1 and Gratian 2, compiled the two recensions.<sup>8</sup>

Several decades before Winroth discovered the first recension of the *Decretum* and argued that the Gratian of the first recension was not the same as the Gratian of the second recension, John Noonan pointed out another obvious way in which it might be

---

<sup>7</sup> “The first recension of the *Decretum* was not a living text. It was a finished product which its author considered ready to be circulated. This is evident from its text, which is as much a finished and polished product as could be expected of any twelfth-century text. Further, it is also evident from the fact that the first recension survives in one version only; what differences there are among the manuscripts are all minor (the apparent exception of Aa will be discussed below). They are differences one would expect to find in any manuscript tradition, arising from scribal mistake or ingenuity. In other words, the manuscripts do not represent different stages in the development of the text, in the manner of ‘classically’ living texts, such as the *Song of Roland*, where each different manuscript version has an equally valid claim to authenticity.” Winroth, 130.

<sup>8</sup> “In the interests of simplicity and clarity, I have therefore chosen to call the author of the first recension Gratian 1 and the author of the second recension Gratian 2. These labels are not intended to suggest that Gratian 1 and Gratian 2 could not have been the same person.” Winroth, 122. “It is impossible to draw any certain conclusions, but the evidence presented in the last two chapters supports the view that that two recensions had difference authors.” Winroth, 194–95.



problematic to conceptualize the *Decretum* as the product of monolithic authorship.

As was noted briefly above and will be discussed in further detail below, the second recension of the *Decretum* had three distinct parts. Noonan warned that the Gratian of Part I was not necessarily the same as the Gratian of Part II or Part III. (And in point of fact, the Gratian of Part III, the compiler of *de Consecratione*, is very unlikely to have been the same as the Gratian of Parts I and II since the treatise was such a late addition to the *Decretum*.) Furthermore, nothing logically excludes the possibility that Winroth and Noonan are both right and that the Gratian of the first-recension of Part I is not necessarily the same as the Gratian of the second-recension of Part I, and so on through all the possible permutations and combinations of recensions (first and second) and parts (I, II, and III).<sup>9</sup>

Long before the discovery of the first recension in the late twentieth century, however, the first twelfth-century readers of the *Decretum* were aware of the limited extent to which it could be thought of as the work of a single author. Even early readers, such as

---

<sup>9</sup> Excepting only that there is no first-recension version of Part III.



Stephen of Tournai and the author of the *Summa Parisiensis*, who implicitly endorsed the single author theory by subsuming the entire vulgate text of the *Decretum* under the eponym “Gratian,” recognized that whoever Gratian was, he was not directly responsible either for the bulk of the text or for certain notable formal features of the work, like the division of Part I into *distinctiones*.<sup>10</sup>

---

<sup>10</sup> “Compositorem huius operis recte dixerim Gratianum, non auctorem. Capitula namque a sanctis patribus edita in hoc volumine composuit, i.e. ordinavit. Non eorum auctor vel conditor fuit, nisi forte quis eum auctorem idcirco dicere velit, quoniam multa ex parte sua sanctorum sententias distinguendo et exponendo in paragraphis suis ponit.” Johann Frierich Schulte, ed., *Die Summa über Das Decretum Gratiani* (Aalen: Scientia Verlag, 1965), 5. “I should properly say that Gratian is the compiler of this work, not the author, for he brought together in this volume—that is, he arranged—rulings enacted by the holy Fathers. He was not their author or framer, unless perchance someone wishes to say that he is thus the author since he offered on his own in his dicta many things for clarifying and explicating the opinions of the holy [Fathers].” Robert Somerville and Bruce Clark Brasington, eds., *Prefaces to Canon Law Books in Latin Christianity: Selected Translations, 500-1245* (New Haven, Conn: Yale University Press, 1998), 200–201. Somerville is not translating from the 1891 Schulte edition but from Herbert Kalb, *Studien Zur Summa Stephani von Tournai: Ein Beitrag Zur Kanonistischen Wissenschaftsgeschichte Des Späten 12. Jahrhunderts*, Forschungen Zur Rechts- Und Kulturgeschichte, Bd. 12 (Innsbruck: Universitätsverlag Wagner, 1983), 113–20.

“Distinctiones apposuit in prima parte et ultima Paucapalea ...” Terence P. McLaughlin, ed., *The Summa Parisiensis on the Decretum Gratiani* (Toronto: Pontifical Institute of Mediaeval Studies, 1952), 1. (Paucapalea arranged the distinctions in the first part and the last, trans. PLE)



Gratian was working within a well-established genre of academic writing in medieval Latin literature, the canonical collection.<sup>11</sup> The *Decretum* followed in the footsteps of any number of systematic canonical collections that had been in circulation since the beginning of the eleventh century in which the material was organized by topic.<sup>12</sup>

Twelfth-century readers of a canonical collection did not necessarily expect its compiler to have written all (or even any) of the words they read. They expected instead to find the distinctive contribution of the compiler in the selection, editing, and arrangement of texts not his own: canons of councils, decrees of popes, and extracts from patristic authorities and secular law, often accompanied by traditional inscriptions ascribing the

---

<sup>11</sup> Cf. Atria A. Larson, *Master of Penance: Gratian and the Development of Penitential Thought and Law in the Twelfth Century*, Studies in Medieval and Early Modern Canon Law, volume 11 (Washington, D.C: The Catholic University of America Press, 2014), 12–13, n30. “My refusal to apply the standard label of ‘canonical collection’ to the *Decretum* is meant to make clear that I do not classify the rest of the *Decretum Gratianii* as a canonical collection in the exact way that the work of Regino, Burchard, Anselm of Lucca, or even Ivo of Chartres (and many anonymous compilers) was, and yet it was a canonical collection in many respects, especially considering much of its source material and how it was used. Many abbreviations of the *Decretum*, for instance, are clear testimony that some religious houses and episcopal courts wanted the canons, not the *dicta*; they wanted a pure canonical collection that could serve as a reference manual to the church’s law.”

<sup>12</sup> In earlier collections such as the *Collectio Dionysiana* (ca.500) the material was organized chronologically.



texts (accurately or inaccurately) to recognized authoritative sources. Gratian met this expectation. Around seventy percent of the text of the vulgate *Decretum* comes from the tradition of canonical texts accumulated over the first millennium of the church's history.<sup>13</sup> An extremely important part of the authorial activity of Gratian, then, did consist in the collection and selection of material from the canonical tradition and in his presentation and organization of the texts that he had collected and selected.

If the author of a canonical collection had anything to say on his own authority, he was expected to do so in a prologue. The Prologue of Ivo of Chartres, for example, was a work of considerable theoretical sophistication that continued to circulate independently and find an audience even after Gratian's *Decretum* superseded the collections to which it had originally been prepended.<sup>14</sup> In this respect, Gratian did not conform to the expectations of the genre in which he was working. For starters, he did

---

<sup>13</sup> 302,384 words or 71.2% of the vulgate *Decretum*.

<sup>14</sup> The prologue was prepended (in slightly different forms) to both the *Decretum* and *Panormia* of Ivo of Chartres. For the edited Latin text, see Bruce Clark Brasington, ed., *Ways of Mercy: The Prologue of Ivo of Chartres; Edition and Analysis*, Vita Regularis, Bd. 2 (Münster : Piscataway, N.J: LIT ; Distributed in North America by Transaction Publishers, 2004). For an English translation, see Somerville and Brasington, *Prefaces to Canon Law Books in Latin Christianity*, 132–58.



not include a preface, something that did not escape the notice of contemporaries like the author of the *Summa Parisiensis*.

In presenting and organizing the canonical texts that he had selected, however, Gratian added a significant amount of first-person commentary (*dicta*), amounting to around twenty percent of the vulgate *Decretum*.<sup>15</sup> Gratian's commentary is what binds the *Decretum* together into a coherent whole, and his *dicta* form the backbone of the arguments that he makes. The *dicta* (sayings) are a feature that Gratian borrowed from another genre, the theological treatise, specifically Alger of Liège's *De misericordia et iustitia*. In addition to the *dicta*, Gratian also wrote the rubrics, so-called because they were written in red ink in manuscripts of the *Decretum*. The rubrics are one-line summaries of canons that they introduce.

As a first-order approximation, then, Gratian could be considered the author, in the modern sense of the word, of the *dicta* and rubrics, and the compiler and arranger of the canons and inscriptions. (In practice, *dicta*, inscriptions, and rubrics cannot be

---

<sup>15</sup> 81,008 words or 19.1%.



distinguished quite so cleanly as we would like: dicta sometimes trail off into inscriptions, and rubrics and inscriptions are sometimes poorly separated, especially in *de Penitentia*.) But, as Noonan further pointed out, the Gratian who compiled, arranged, and edited the canons and inscriptions was not necessarily the same as the person who wrote the *dicta*, who in turn was not necessarily the same as the person who created the rubrics.<sup>16</sup> All of these considerations suggest that the presumption that the *Decretum* was the product of a single author ought to be entertained with extreme caution. As I turn my attention to the collection and selection, presentation, and organization of the canonical texts found in the *Decretum*, I will continue to use the name Gratian as a conventional label for its author or authors, while withholding judgment as to whether one person or many stand behind the name

---

<sup>16</sup> "There are two difficulties when this authorship comes to be determined. One is that the *Concordia* is a book in three distinctive parts, with several distinctive sections within the parts, and these parts have undergone editing as to chapter heading, chapter insertions, and chapter order. Whatever the later formal ascription, is Gratian responsible for all these parts? The other difficulty is the range of roles the one 'responsible' person might have had—compiler, commentator, rubricator, reviser. Was Gratian one or all? The difficulties relate to one another, because evidence that Gratian was, say, the commentator of Part II does not establish what role he had in Part I." Noonan, "Gratian Slept Here," 162–63.



## Collection and Selection

A canon is an authoritative text that provides a rule or norm for Christian life, either for individual Christian believers in their clerical or lay station or for the Church corporately at whatever level—local church or religious foundation, diocese, province, or the universal Church. The Greek work *kanon* literally means a ruler or straightedge used to mark a line, and Christian believers individually and the Church corporately follow a straight path when they follow the rules marked out by the *kanones*. “*Canon grece, latine regula nuncupatur. Regula dicta est eo quod recte dicit, nec aliquando aliquo sum trahit. Alii dixerunt regulam dictam, vel quod regat, vel normam recte vivendi prebeat, vel quod distortum pravumque est corrigat.*”<sup>17</sup> (“Canon” is Greek for what is called a ‘rule’ in Latin. It is called a rule because it leads one aright and never takes one astray. But others say that it is called a rule because it rules, presents a norm for right living, or sets aright what is twisted and bent.) The late first- or early second-century *Didache* is an example of a very early Christian text that contains such rules or norms. After the official recognition

---

<sup>17</sup> D.3 c.1-2. *Decretum Gratiani*, First Recension, edition in progress. Anders Winroth, 3/21/2017, p.5, l.13-18.



of Christianity by Constantine in 313, the Church had a greater need for an agreed-upon set of rules. Such rules were provided by the canons of provincial and ecumenical councils, and somewhat later by papal decretals.

Gratian defines canons as either *decreta Pontificum* ("decrees of pontiffs") or *statuta conciliorum* ("statutes of councils").<sup>18</sup> From a relatively early date, the bishops of Rome had cultivated the practice of issuing decretal letters, consciously modeled after Roman imperial rescripts, in which they laid down disciplinary and doctrinal norms in response to questions put to them, usually by other bishops.<sup>19</sup> The earliest surviving example is a letter from Pope Siricius (†399) to Bishop Himerius of Tarragona, written in 385 in response to a letter that Himerius had sent to Siricius's predecessor, Pope Damasus (†384).<sup>20</sup> The letter, sometimes referred to by the first word of its text as the

<sup>18</sup> "Porro canonum alii sunt *decreta Pontificum*, alii *statuta conciliorum*." D.3 d.p.c.2. *Decretum Gratiani*, First Recension, edition in progress. Anders Winroth, 3/21/2017, p.5, l.19-20.

<sup>19</sup> In addition to genuine papal material, for example a large number of excerpts from the register containing the letters of Gregory I, Gratian used a number of pseudo-papal sources, which will be discussed below in the context of Gratian's formal sources.

<sup>20</sup> JK 255. Philipp Jaffé, ed., *Regesta Pontificum Romanorum*, 2nd ed, vol. 1 (Graz: Akademische Druck, 1956), 40. ["Register of papal letters to 1198. Reflecting the editors who contributed to various time



*Directa decretal*, remains an important source for norms concerning clerical celibacy in late antiquity. For modern readers, the most familiar example of an imperial rescript, the genre after which the papal decretal letter was modeled, is the early second-century (ca. 112) response of the emperor Trajan (†117) to a letter from Pliny the Younger inquiring as to how Pliny, then Roman provincial governor of Bithynia and Pontus, ought to proceed against those accused of being Christians.

Another extremely important source for the *Decretum* is canons from the ecumenical councils and from historically important provincial synods: *Conciliorum vero alia sunt universalia, alia provincialia*. (“Some councils are universal, others provincial.”)<sup>21</sup> Although the Council of Nicaea (325) is primarily remembered for its doctrinal settlement, memorialized in the Nicene Creed, of the Arian controversy over the divine nature of Christ, it also promulgated twenty canons of a disciplinary nature. Provincial synods tended to assume an outsized importance in the canonical tradition when

periods, the work is conventionally given an abbreviated citation as JK to the year 590, JE for 590-882 and JL for 883-1198.” [Papal Documents: A Finding Aid | Columbia University Libraries](#)

<sup>21</sup> D.3 d.p.c.2. *Decretum Gratiani*, First Recension, edition in progress. Anders Winroth, 3/21/2017.



figures such as Augustine of Hippo (†430) or Caesarius of Arles (†542) either participated in or presided over them: “*Etiam S. Augustinus Yponensis episcopus in eadem sinodo legitur fuisse.*” (It is read that St. Augustine, bishop of Hippo, also attended this synod.)<sup>22</sup> By the time the *Decretum* was compiled, the cumulative output of the ecumenical councils and provincial synods amounted to a considerable mass of material, and Gratian devoted most of *Distinctiones* 15 and 16 to enumerating the canonically authoritative councils and synods.

In addition to textual material strictly defined as canons (“decrees of pontiffs” and “statutes of councils”), Gratian drew on other sources which, if not canons according to the narrower technical definition, could be treated as authoritative in a more general sense. The most important of these *auctoritates* were drawn from the writings of major patristic figures such as Ambrose (†397), Jerome (†419 or 420), Augustine (†430), and Gregory (†604), whom later generations considered doctors of the church.<sup>23</sup> It is easy to

<sup>22</sup> D.16 c.11. *Decretum Gratiani*, First Recension, edition in progress. Anders Winroth, 3/21/2017. p.32, 1.7-9.

<sup>23</sup> Ambrose, Augustine, Gregory the Great, and Jerome were formally recognized as doctors by Boniface VIII in 1298 in the *Liber Sextus*, VI 3.22, *Gloriosus Deus in sanctis suis in De reliquiis et veneratione sanctorum* (edF 2.1059-1060). Leo the Great was only added to the list in 1754.



see how works such as Ambrose of Milan's *De officiis* ("On duties")—a Christian answer to the work of Cicero with the same title—could be a rich source of norms. Gratian provides an explicit list of "the works of the holy fathers that are received in the Catholic Church."<sup>24</sup> (It is necessary to distinguish the papal from the patristic when dealing with figures like Gregory. While textual excerpts from the letters found in the registers of Gregory clearly derive their authority from the fact that Gregory was pope, a text like the *Moralia in Job* has an authority to a large extent independent of the formal ecclesiastical office Gregory held.)<sup>25</sup>

---

<sup>24</sup> First recension D.15 c.3, §1-§16 in Friedberg, §2-§17 in Thompson and Gordley.

<sup>25</sup> "The *Moralia*: based on talks Gregory gave on the Book of Job to his 'brethren' who accompanied him to Constantinople while he held the office of papal *apocrisiarius* (see above). The work as we have it is the result of Gregory's revision and completion of it soon after his accession to the papal office." R. A. Markus, *Gregory the Great and His World* (Cambridge ; New York: Cambridge University Press, 1997), 15. "In the case of Gratian's chapters from Gregory I's *Moralia in Job*, the unusually precise citations in Gratian's inscriptions lead to the conclusion that he used this work directly (as Munier already argued)." Peter Landau, "Gratian and the *Decretum Gratiani*," in *The History of Medieval Canon Law in the Classical Period, 1140-1234: From Gratian to the Decretals of Pope Gregory IX*, ed. Wilfried Hartmann and Kenneth Pennington, *History of Medieval Canon Law* (Washington, D.C: Catholic University of America Press, 2008), 22–54, 35. "There are six chapters taken from the *Moralia*, and Gratian noted the book and chapter of the text in his inscription: D.13 c.2; D.45 c.9 and c.14; D.46 c.1 and c.2; D.47 c.3. Each text is also in the first recension." Landau, 35n22.



Finally, in addition to excerpts from works of patristic authors, another source for extra-canonical authorities was secular law, primarily (pre-Justinianic) Roman law but also including the capitulary legislation of the more important Carolingian emperors — Charlemagne (†814), Louis the Pious (†840), and Charles the Bald (†877)—and their successors.

Up to this point, I have been implicitly referring to Gratian's material sources, texts such as the *Letters* of Gregory I or the *de Officiis* of St Ambrose from which the canons in the *Decretum* were ultimately derived. And in a small number of cases, Gratian did work directly with material sources. Isidore of Seville's *Etymologies* is an example of a text from which Gratian probably collected excerpts directly from the material source. It is also possible that he collected excerpts directly from Gregory the Great's *Moralia in Job*. But Gratian drew the overwhelming majority of the canons he compiled in the *Decretum* from formal sources, predecessor collections containing patristic, conciliar, and papal (as well as pseudo-papal) authorities, predigested into canon-sized units of text.

Peter Landau has suggested that Gratian relied primarily on just five formal sources in writing the Decretum: Anselm of Lucca's *Collectio canonum*, the pseudo-Ivonian *Collectio*



*Tripartita*, Ivo of Chartes's *Panormia*, Gregory of San Grisogono's *Polycarpus*, and an anonymous *Collection in Three Books* (3L).<sup>26</sup> This hypothesis is not universally accepted: Pennington, for example, believes that Gratian may have also relied on one or more now-lost Central Italian collections similar to the *Collection in Nine Books* (9L).<sup>27</sup>

By the end of the Patristic period, the canons of the ecumenical and of historically important provincial councils constituted a substantial body of canon law, which were gathered into collections such as the *Collectio Dionysiana* (ca. 500). A later revision of the *Dionysiana*, the *Collectio Dionysio-Hadriana* (774), exercised enormous influence on the transmission of canon law to the Carolingian world and through it to medieval Western Europe.

---

<sup>26</sup> Peter Landau, "Gratians Arbeitsplan," in *Iuri Canonico Promovendo: Festschrift Für Heribert Schmitz Zum 65. Geburtstag*, ed. Winfried Aymans and Karl-Theodor Geringer (Regensburg: F. Pustet, 1994), 691–707. See also Winroth, *The Making of Gratian's Decretum*, 15–17.

<sup>27</sup> For the *Collectio IX librorum* of the Ms Vat. lat. 1349 (9L), see Linda Fowler-Magerl, *Clavis Canonum: Selected Canon Law Collections Before 1140: Access with Data Processing*, Monumenta Germaniae Historica. Hilfsmittel 21 (Hannover: Hahnsche, 2005), 79–82



Collectors like Dionysius Exiguus (†ca. 540), the original compiler of the *Collectio Dionysiana*, located all authority in the distant past. This attitude prevailed throughout the period during which the canonical sources later used by Gratian were being compiled. Law, however, ultimately has to correspond with and respond to the needs of contemporary society, and by the ninth century the Church, especially north of the Alps, was operating in a very different political and social environment from that of the Mediterranean world of late antiquity, which had produced the sources for Dionysius's collection. When faced with the need for new law to cope with new circumstances, some enterprising Carolingian churchmen took more recent material—mostly canons from provincial synods—and repackaged them, attributing them to popes from the first century through Gregory I (†604). The compilers of these collections are usually referred to as “forgers,” a convention I will observe in the following discussion, but they were not simply inventing their sources. Most of the material in the collections was genuine but of relatively recent origin—the goal of the forgers was simply to retroject it far enough into the past to meet the standard they set for authoritativeness.

In order to understand the intent of the forgers, one has to understand the immediate political situation to which they were reacting. Charlemagne's son and successor,



Emperor Louis the Pious (†840), was dethroned by three of his sons in a ceremony of public penance in 833. When Louis was unexpectedly restored less than six months later, he moved quickly to depose many of the bishops who had participated in imposing the penance on him, starting with Ebbo, archbishop of Rheims.<sup>28</sup> The forgers wanted to protect bishops from being deprived of office by emperor, king, or metropolitan. They did this by building up the pope as the only superior who could judge a bishop. This was convenient, because while popes in the ninth century had considerable moral authority north of the Alps, they had little real power. The intent of the forgers' program therefore was to render bishops effectively impervious to judgment. The collections that resulted from this effort, the so-called Pseudo-Isidorian *Decretals* and *Benedictus Levita*, were accepted everywhere as genuine in an age lacking any serious historical-critical awareness, and they survived to become important

---

<sup>28</sup> Mayke De Jong, *The Penitential State: Authority and Atonement in the Age of Louis the Pious, 814-840* (Cambridge, UK ; New York: Cambridge University Press, 2009).



sources for canonical collectors in the eleventh century.<sup>29</sup> Many of these forged canons eventually found their way into Gratian's *Decretum*.

The reform papacy that reached its apogee with the pontificate of Hildebrand, who reigned as Gregory VII from 1073 to 1085, provided a new impetus for the collection and study of canons. The special concerns of the Gregorian reformers also shaped their approach to the collection and presentation of canons. The consuming interest of the reform generation was the relationship between *regnum* and *sacerdotium* (church and state). This interest in the problem of church-state relations was instrumental: the

<sup>29</sup> For the much-criticized 1863 edition, see Paul Hinschius, ed., *Decretales Pseudo-Isidorianaes, et, Capitula Angilramni: Ad Fidem Librorum Manuscriptorum Recensuit, Fontes Indicavit, Commentationem de Collectione Pseudo-Isidori Praemisit* (Aalen: Scientia Verlag, 1963). Recent efforts to replace the Hinschius edition include [Zum Inhalt von "Projekt Pseudoisidor"](#) by Karl-Georg Schon; and [Pseudo-Isidore | An edition-in-progress of the False Decretals](#) by Eric Knibbs. The work of Klaus Zechiel-Eckes (+2010) forms the basis for much of our current understanding of the place and date at which the pseudo-Isidorian collection was compiled (Corbie in the mid-830s), and therefore of the political motivations of the compilers; see "Ein Blick in Pseudoisidors Werkstatt. Studien Zum Entstehungsprozeß Der Falschen Dekretalen. Mit Einem Exemplarischen Editorischen Anhang (Pseudo-Julius an Die Orientalischen Bischöfe, JK +196)," *Francia* 28, no. 1 (2001): 37–90. Horst Fuhrmann, "The Pseudo-Isidorian Forgeries," in *Papal Letters in the Early Middle Ages*, ed. Detlev Jasper and Horst Fuhrmann, History of Medieval Canon Law (Washington, D.C.: Catholic University of America Press, 2001), 137–95 provides a good general introduction to the forged decretals; see also [Introduction to Pseudo-Isidore - Decretum Gratiani](#) by Eric Knibbs. See Horst Fuhrmann, *Einfluß Und Verbreitung Der Pseudoisidorischen Fälschungen : Von Ihrem Auftauchen Bis in d. Neuere Zeit*, Schriften Der Monumenta Germaniae Historica, Deutsches Institut Für Erforschung Des Mittelalters ; Bd. 24 (Stuttgart: Hiersemann, 1972) on the influence of the forged decretals.



ultimate concern of the eleventh-century reformers, as it had been for their eighth- and ninth-century Carolingian predecessors, was for the reform and renewal of monastic and clerical life. Unlike the Carolingians, however, who had thought that cooperation between ecclesiastical and secular authorities was essential for the reform of religious life, the eleventh-century reformers believed that secular rulers had to acknowledge the jurisdictional supremacy of the papacy and the independence of the church from lay interference as a necessary precondition for effective clerical and monastic reform. It is significant that the most important canon law collections of the reform period were compiled by prelates sympathetic with the reform movement, such as Anselm, bishop of Lucca (†1086), compiler of the *Collectio canonum*; Gregory, cardinal of San Grisogono (†1113), compiler of the *Polycarpus*; and Ivo, bishop of Chartres (†1115), compiler of the *Panormia*. Burchard of Worms (†1025) was a forerunner of this group. Although he died before the reform papacy or the investiture controversy, his *Decretum* has more in common with collections compiled during that period, insofar as it is systematic and reflects concerns similar to those of his successors.

Unsurprisingly, given their concern with asserting the jurisdictional supremacy of the papacy and the independence of the church from lay interference, the eleventh-century



collectors showed considerable interest in the forged decretals of the ninth century.

For example, 252 of the 315 chapters of the anonymous *Collection in 74 Titles* (ca. 1050) are drawn from the Pseudo-Isidorian *Decretals*.<sup>30</sup>

The program of the Gregorian reformers collided with two practical realities of eleventh-century life. Abbotts and bishops were large landowners and therefore important vassals of secular rulers. Furthermore, secular rulers were almost completely dependent on the Church for administrative personnel until the rise of university faculties of Roman law in the twelfth century. It was therefore essential from the point of view of secular rulers that they should be able to control appointments to key benefices in order to ensure the loyalty of their holders. The controversy over this issue dominated papal-imperial relations for half a century. It was finally resolved by the Concordat of Worms (1122), which provided that key benefices were to be filled through regular canonical processes free from lay interference (such as election by a

---

<sup>30</sup> "Altogether I have estimated that 252 out of the 315 *capitula* were taken from Pseudo-Isidore." J. T. Gilchrist, ed., *The Collection in Seventy-Four Titles: A Canon Law Manual of the Gregorian Reform*, Mediaeval Sources in Translation 22 (Toronto: Pontifical Institute of Mediaeval Studies, 1980), 15



cathedral chapter) but that the benefice holder should swear fealty to the secular ruler for his fiefs.

Gratian depended on a relatively small number of eleventh- and early twelfth-century systematic canonical collections, his formal sources, and those predecessor collections, in turn, depended on an older stratum of mostly chronologically arranged canonical collections. Gratian was thus selecting from a body of canonical material that had already passed through a fairly rigorous filtering process, and two of the most important filters—the ninth-century pseudo-Isidorian forgers and the eleventh-century Gregorian reformers—had had extremely strong ideological orientations.



# Chapter 2

## The *Decretum*

### Presentation

Having considered how Gratian sourced and collected his materials—the question of what canons are and where Gratian got them—we now turn to the question of how Gratian presented and organized the canons he had collected. Here, it is important to distinguish between presentation, the small-grain or micro structure of the *Decretum*, and organization, the large-grain or macro structure of the *Decretum*. The organization of the *Decretum* is very regular, so structurally regular, in fact, that Reuter and Silagi's OCP e-text of the Friedberg edition can be taken apart using a recursive descent parser<sup>1</sup>

---

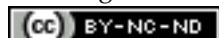
<sup>1</sup> See Appendix 2 for the Python code listing for my implementation of the recursive-descent parser. Thanks to Patricio Simari of the Electrical Engineering and Computer Science Department at The Catholic University of America, who provided helpful suggestions on parser implementation. In Computer Science-theoretic terms, the parsed *Decretum* can be thought of as an ordered tree, with parts, distinctions, cases, and questions as interior nodes. A canon or chapter, properly conceptualized as a node encapsulating rubrics, inscriptions, canon texts, and *dicta post*, is also an internal node. A *dictum (ante, post, or initiale)* is a terminal node. So are rubrics, inscriptions, and canon texts. A case statement (*dictum initiale*) is always the first terminal node of a case. A *dictum ante* is always the first terminal node of a distinction or question. The anomalous C.16 q.2 d.a.c.8 is actually the *dictum ante* introducing a vestigial fifth question positioned between C.16 q.2 and C.16 q.3 (see Friedberg, 1.787). The traditional notation

and should be thought of as a sequence of hierarchically nested containers. Taking a bottom-up approach, we will first consider the presentation or small-grain structure, and start with the canons, the atomic containers out of which the higher-order containers in the *Decretum*—parts, distinctions, cases, and questions—are built up.

The way in which canons are presented in the *Decretum* differs somewhat from the way in which they were presented in the formal sources from which Gratian selected his material. A canon in the formal sources typically included an inscription, a rubric, and the canon text. Gratian added a *dictum post canonem* (plural *dicta post canones*) containing his own commentary on the preceding canon or canons, a feature borrowed from Alger of Liège's theological treatise *De misericordia et iustitia* but without an immediate precedent in the canonical literature.<sup>2</sup>

misleadingly implies that d.a.c.1 is a leaf node of c.1. In fact, d.a.c.1 and c.1 are both child nodes at the same nesting level under one parent distinction or question node. A *dictum post* is an optional, and usually the last, terminal node of a canon.

<sup>2</sup> Gratian is known to have used Alger as both a source, especially in C.1, and, to some extent, as a methodological model. Anders Winroth, *The Making of Gratian's Decretum* (Cambridge: Cambridge University Press, 2000), 17, 39, 144. See Robert Kretzschmar's edition, *Alger von Lüttichs Traktat "de Misericordia et Iustitia": Ein Kanonistischer Konkordanzversuch Aus Der Zeit Des Investiturstreits: Untersuchungen Und Edition, Quellen Und Forschungen Zum Recht Im Mittalder*, Bd. 2 (Sigmaringen: J.



## Inscription

Most canons have an inscription, which identifies the ultimate source of authority for that canon, usually a papal decretal, a canon from an ecumenical council or an important provincial synod, or a patristic text.<sup>3</sup> This raises the need to distinguish Gratian's material sources—an original letter of Pope Gregory I, for example—from his formal sources—the collection from which Gratian actually copied his text.<sup>4</sup> Because

---

Thorbecke, 1985). Somewhat confusingly for Gratian scholars, Kretzschmar uses regular typeface for Alger's *dicta* and italics for the canons, the opposite of the convention adopted by Friedberg and maintained by Winroth.

<sup>3</sup> The distinction between rubrics and *dicta* is blurry: "Gratian rarely took his rubrics from earlier collections. Rather he created his own and often melded the rubrics with the dicta." Melodie Harris Eichbauer, "St. Gall Stiftsbibliothek 673 and the Early Redactions of Gratian's *Decretum*," *Bulletin of Medieval Canon Law* 27 (2007): 107. "Both J. Rambaud-Buhot and John Noonan, Jr. have highlighted the similarity between dicta and rubrics, that is, a rubric very often echoes the dictum that immediately preceded it" and "These isolated instances in the first cluster show that Gratian felt that the dictum was sufficient for summarizing the following *auctoritas*." Eichbauer, 115. **Expand to include inscriptions.**

<sup>4</sup> "Since Gratian frequently took fragments of letters from the Register of Gregory I—266 in all—using the inscription 'in registro', older research assumed that he must have used this important source in the form of the *Registrum Hadrianum*. According to more recent research (Landau), even these texts from the Register of Gregory I found in Gratian derive almost without exception from canonical collections predating Gratian; the direct use of the Register is probable in only a single case." [C.27 q.1 c.19 (JE 1496)] Peter Landau, "Gratian and the *Decretum Gratiani*," in *The History of Medieval Canon Law in the Classical Period, 1140-1234: From Gratian to the Decretals of Pope Gregory IX*, ed. Wilfried Hartmann and Kenneth Pennington, *History of Medieval Canon Law* (Washington, D.C: Catholic University of America Press, 2008), 34.



Gratian collected his texts almost exclusively from formal sources, the inscription was sometimes historically inaccurate, especially if the text in question was one derived from the Pseudo-Isidorian collections.

### Rubric

A canon is often, though not always, introduced by a rubric. The name refers to the fact that rubrics were conventionally written in red ink in manuscripts. A rubric is a very short summary of what the following canon is about. In many cases the rubric simply reads “*de eodem*”, which means “about the same thing” as the preceding canon.<sup>5</sup>

<sup>5</sup> The role of the *de eodem* rubrics in Winroth’s argument that the Aa, Bc, Fd, and P mss. of the *Decretum* are a first recension rather than an abbreviation of the vulgate. Winroth, *The Making of Gratian’s Decretum*. Kuttner’s statement of the “untidy seams” problem in “Acta and Agenda”, and how Winroth solves it. Stephan Kuttner, “Research on Gratian: Acta and Agenda,” in *Studies in the History of Medieval Canon Law, Collected Studies CS325* (Aldershot, Hampshire, Great Britain : Brookfield, Vt., USA: Variorum ; Gower, 1990). Winroth observed that there are 398 *de eodem* rubrics in the vulgate *Decretum*, see Anders Winroth, “Uncovering Gratian’s Original *Decretum* with the Help of Electronic Resources,” *Columbia Library Columns* 46, no. 1 (1997): 28; and Winroth, *The Making of Gratian’s Decretum*, 127. The precision of this frequently-cited number can be refined in ways that demonstrate the true power of the careful use of electronic resources for the study of this, or any, text. (In the following examples, *edF.txt* is the filename of the OCP-format Reuter and Silagi e-text of the Friedberg edition. The commands used in the examples should work on any macOS or Linux-based system.) Winroth’s figure of 398 simply represents the number of occurrences of the substring “*de eodem*” in the file: `grep -i "de eodem" edF.txt | wc -l` returns 398. Most, but not all, of the occurrences of the substring “*de eodem*” do, in fact, appear in the context of a rubric. In one case, *de Cons. D.2 c.3*, the inscription reads *Idem de eodem* (“the same person about the same thing”), and refining the search to count those



(Pennington and Eichbauer believe that Gratian wrote the rubrics in the *Decretum* rather than reusing those he found in formal source collections.<sup>6</sup> The fact that the rubrics are short, discontinuous, texts means that they are of effectively no value for the purpose of authorship attribution. On the other hand, the fact that they are believed to have been written by Gratian, and therefore reflect his choice of wording, means that they are potentially useful evidence in the search for new topics introduced between the first and second recensions.)

---

occurrences that do **not** appear alongside the OCP rubric <T R> or inscription <T I> tags reveal that in 8 instances, the words *de eodem* are simply part of the text of a canon or *dictum*: `grep -i "de eodem" edF.txt | egrep -v "<T R>|<T I>" | wc -l` returns 8. Setting aside, then, those instances in which the words *de eodem* occur in canons, *dicta*, or inscriptions (and acknowledging that the criteria for distinguishing between inscriptions and rubrics can be blurry) leaves 389 rubrics containing the substring "de eodem": `grep -i "de eodem" edF.txt | grep "<T R>" | wc -l` returns 389. Of those 389, there are 373 *De eodem* rubrics and 13 *Item de eodem* rubrics (12 of those 13 introduce second-recension canons in D.30 in the first part of the *Decretum*): `fgrep "De eodem." edF.txt | grep "<T R>" | wc -l` returns 373, and `fgrep "Item de eodem." edF.txt | grep "<T R>" | wc -l` returns 13. In the three remaining cases, words *de eodem* are part of a longer rubric: `grep -i "de eodem" edF.txt | grep "<T R>" | fgrep -v "De eodem." | fgrep -v "Item de eodem."` returns the rubrics for D.23 c.32 (*De eodem, et ut clericci comam non nutrient*), C.22 q.5 c. 16 (*De eodem, et ut a ieunis iuramenta prestentur*), and *de Cons.* D.1 c.17 (*De eodem, et quod octo diebus dedicationum*). Winroth's larger point, however, had to do with the place of the *de eodem* rubrics in the "untidy seams" problem. For that purpose, only the 137 *de eodem* rubrics in the first recension of the *Decretum* are potentially relevant evidence. See GitHub [Sand](#) repository. This result (137) is potentially off by one, i.e., the total number of *de eodem* rubrics in the first and second recension rubric data sets is 388 instead of 389, with no way to tell whether the missing rubric is from the first or second recension.

<sup>6</sup> Placeholder for footnote referencing Eichbauer's observation that Gratian wrote his own rubrics.



### Canon text

(The nature of the canon texts proper has been described above in the section on the collection and selection of canons.)

Finally, there are the *dicta post canones*, usually abbreviated d.p.c. and literally meaning “something said after the canons.” The *dicta post* are statements that Gratian made on his own authority as a jurist, carry the thread of his argument, and do his interpretive work. The *dicta* are texts that Gratian (whether we think of him as one person or many) actually wrote, and are therefore subject to analysis for authorship attribution. “The *dicta* in Gratian’s *Decretum* bring the reader closer to its author than any other part of the text.”<sup>7</sup> (The *dicta post* are not found in Part III, which is in this respect more structurally similar to the formal source collections Gratian relied on than to the rest of the *Decretum*).

---

<sup>7</sup> Winroth, 187.



# Chapter 2

## The *Decretum*

### Organization

[Move all discussion of d.a.c. and d.init. into Organization. Case statements or *themata* are a special case of *dicta* found only in, and closely tied to the organization of, Part II.]

The *Decretum* in the form that circulated after around 1150 (the Köln Dombibliothek manuscripts 127 and 128 are good early examples) has three parts. The first part consists of 101 distinctions. The division into distinctions is traditionally thought to have been made by Paucapalea (fl. 1140s) rather than by Gratian. The division seems to have taken some time to reach its final form: Pennington has pointed out that there were 99-distinction versions of the first recension, and Rufinus writing in the mid-1160s presumably about the second recension describes Part I as “having been completed in on hundred distinctions.” Groups of distinctions form treatises within the work: the first twenty distinctions, for example, form a treatise on laws (*tractatus de*

*legibus*) that explores the sources of law. Other topics covered include ecclesiastical hierarchy and clerical discipline. The second part consists of 36 cases, which are further subdivided into 168 questions. Four of the questions, C.11 q.2, C.17 q.3, C.22 q.3, and C.29 q.1 are empty, and one, C.33 q.3, encapsulates a treatise on penance (*tractatus de penitentia*, abbreviated *de Pen.*) divided into seven distinctions. The third part of the *Decretum* is a treatise on sacraments (*tractatus de consecratione*, abbreviated *de Cons.*) divided into five distinctions.<sup>1</sup> Contemporaries like Rufinus and Stephen of Tournai understood the parts to treat respectively of ministries, business, and sacraments: “*primam ministeriis, secundam negotiis, tertiam ecclesiasticis deputat sacramentis.*”<sup>2</sup>

---

<sup>1</sup> Anders Winroth, *The Making of Gratian's Decretum* (Cambridge: Cambridge University Press, 2000), 5.

<sup>2</sup> Rufinus and Stephen of Tournai describe the tripartite division in exactly the same words: “librum suum in tribus partibus distinguit, quarum primam ministeriis, secundam negotiis, tertiam ecclesiasticis deputat sacramentis.” Heinrich Singer, *Summa Decretorum* (Aalen : Paderborn: Scientia Verlag ; F. Schöningh, 1963), 5. “He divides his book into three parts, the first of which he devotes to ecclesiastical ministries, the second to [ecclesiastical] business, and the third to ecclesiastical sacraments.” Robert Somerville and Bruce Clark Brasington, eds., *Prefaces to Canon Law Books in Latin Christianity: Selected Translations, 500-1245* (New Haven, Conn: Yale University Press, 1998), 192–93. (Rufinus) “Harum primam ministeriis, secundam negotiis, tertiam ecclesiasticis deputat sacramentis.” Johann Frierich Schulte, ed., *Die Summa über Das Decretum Gratiani* (Aalen: Scientia Verlag, 1965), 6. “The first of these parts is devoted to ecclesiastical ministries, the second to ecclesiastical problems, the third to ecclesiastical sacraments.” Somerville and Brasington, *Prefaces to Canon Law Books in Latin Christianity*, 201. (Stephen of Tournai) (Brasington translated the preface of Rufinus, Somerville that of Stephen of Tournai. Somerville



The *dictum ante canonem* (abbreviated d.a.c.) and *dictum initiale* (abbreviated d.init.) are special cases of *dicta* and, like the *dictum post canonem*, were written by Gratian (i.e., the person performing the author-function at that place in the text). My most important finding is that the author of the first-recension *dicta* did not write the case statements. Gratian intended the case statements to be memorable, not necessarily realistic.

The hypothetical case statements or *themata* that introduce each of the thirty-six *causae* (cases) that constitute the second part of the *Decretum* reflect at least in some places an outlook and a set of concerns that university-educated urban professionals of the present day would not find it difficult to identify with. At a time when considerable attention is being paid in the realm of popular discourse (if not yet public policy) to the issue of “opportunity hoarding” on behalf of upper- and upper-middle class children

---

and Brasington, *Prefaces to Canon Law Books in Latin Christianity*, vii. There is no indication as to why Somerville chose to translate *negotiis* as “problems” rather than “business”).

Given the probable dates of the *summae* of Rufinus (1164) and Stephen of Tournai (1165/66), it is just possible that Stephen got the wording from the preface to the *Summa* of Rufinus. Kenneth Pennington, [Rufinus](#) and [Stephanus Tornacensis](#), *Bio-Bibliographical Guide to Medieval and Early Modern Jurists*. Stephen, however, had been a student of Rufinus in Bologna, and it seems more likely that this is something he had heard Rufinus say in lecture. **Kenneth Pennington**, “*Étienne of Tournai*”.



by their university-educated, urban professional parents, the intense concern of “a certain man” in the first case statement to secure a good ecclesiastical career for his son (and his willingness on more than one occasion to resort to bribery in order to realize that objective) has a distinctly contemporary feel to it.<sup>3</sup>

In other cases, the hypothetical case statements are quite remote from our own concerns. In case 28, for example, a married infidel, by which Gratian means a Muslim, converts to Christianity. As a result, his wife separates from him out of hatred for his new faith. The convert then takes a new Christian wife, and after she dies, he is made a priest. Finally, “on account of the merit of his life and learning,” he is elected bishop. Gratian extracts questions from this fantastical fact-pattern, however, that are of considerable theoretical, though perhaps not practical, interest: whether there can be marriage between infidels, whether there is ever a situation in which a man can take another wife if his previous wife is still alive; and whether someone who had one wife

---

<sup>3</sup> “Quidam habens filium …” C.1, d.init., edF. 1.357. Although Gratian does not say so explicitly, I think that we can reasonably infer that the father is an urban merchant or professional, rather than a member of the landed aristocracy, since his wealth, the source of several bribes, is so readily convertible into cash.



before he was baptized and another wife after he was baptized should be considered a bigamist.

The third part, *de Consecratione*, is problematic in several ways. That *de Cons.* is a very late addition is not in doubt—it is absent from all first-recension manuscripts, nor was it included in the earliest manuscripts of the more widely-circulated second-recension version of the *Decretum*, as indicated by glosses found in the Gt, Pf, and Tr manuscripts describing Gratian's work as “principally divided into two parts.”<sup>4</sup>

---

<sup>4</sup> Need footnote on Gt, Pf, and Tr gloss



## Chapter 2c

### *Correctores Romani* edition (1582)

The *Decretum* never received the kind of formal ecclesiastical (papal or conciliar) recognition that was conferred on decretal collections like the *Liber Extra*, *Liber Sextus*, and *Clementinae*. Nevertheless, it remained foundational both to the *Corpus Iuris Canonici*, the body of canon law that governed the Roman Catholic Church right up until 1917, as well as to the academic study of canon law throughout the medieval and well into the early modern periods. As a result, the text of the *Decretum* came in for close attention as part of the reform program initiated at the Council of Trent (1545-1563). A commission led by the Spanish archbishop Antonio Agustín (1517-1586), the *Correctores Romani*, was charged with producing an approved version of the text of the *Decretum* (along with the rest of the *Corpus Iuris Canonici*) for the age of the printing press.

The *Correctores* proceeded from the Renaissance Humanist assumption that their job was to identify and include the most textually correct version of the material source for each of Gratian's authorities. In this sense the *Correctores* were less concerned with

getting at the text of the collection Gratian actually compiled, and more concerned with getting at the text that they thought Gratian should have compiled, if his concerns had been those of a Renaissance philological scholar.<sup>1</sup>

The 1582 printed edition of the *Decretum* includes the ordinary gloss, the standardized commentary by [Johannes Teutonicus](#) (c.1170–1245) in 1216, which was revised and updated by [Bartholomeus Brixiensis](#) (fl.1234–1258) around 1240-1245. UCLA's online version of the [Corpus Iuris Canonici \(1582\)](#) is a convenient way for modern scholars to consult the ordinary gloss, which is an important tool for understanding the reception of Gratian's text by medieval jurists.

---

<sup>1</sup> The idea that the *Correctores* were trying to reproduce the text of the *Decretum* as they thought Gratian should have written it, as opposed to the text as he actually wrote it, goes back at least as far as Friedberg: "Vides non id in animo habuisse correctores Romanos, ut restitueretur decretum, quale a Gratiano compositum esset, sed quale a Gratiano componi debuisset." ("You see it in the mind that the *Correctores Romani* had, as the *Decretum* was restored, not how it was composed by Gratian, but how it ought to have been composed by Gratian." trans. PLE) (edF 1.lxxviii, see also Anders Winroth, "[Emil Friedberg's Edition of Gratian's Decretum: An Evaluation](#)".)



# Chapter 3

The third order of business then is to recount the history of efforts, from the *Summa Parisiensis* on, of legal scholars and historians to find the man behind the name Gratian. Here, you need to lay out the traditional evidence used for solving this puzzle. You are lucky that Noonan, Pennington etc. have laid out the evidence and played with it for you.

– Stan Chodorow

## Gratian

The starting point for all modern scholarship on Gratian himself is John T. Noonan Jr.'s classic 1979 *Traditio* article "Gratian slept here: The changing identity of the father of the systematic study of canon law". Noonan's approach is apophasic, demonstrating that there is no sound historical basis for most of the affirmative statements, for example that Gratian was a Camaldolese monk at the monastery of SS. Felix and Nabor in Bologna, that many mid-twentieth century handbooks and reference works presented as commonplace.<sup>1</sup>

---

<sup>1</sup> John T. Noonan, "Gratian Slept Here: The Changing Identity of the Father of the Systematic Study of Canon Law," *Traditio* 35 (January 1979): 145–72.

Very little is known about the historical Gratian, and almost everything that previous generations thought was known about Gratian has proved, on closer examination, to be myth.<sup>2</sup> On the evidence found in the *Decretum*, Gratian seems to have been less interested in contemporary ecclesiastical or secular politics in general, and in the Investiture Controversy (1076-1122) in particular, than some modern scholars think he should have been. Gratian similarly seems, based on the relatively limited use he makes in the *Decretum* of material from the *Codex* and *Digest* of Justinian, to have been either less interested in, or less aware of, the revival of the study of Roman law in the early twelfth century than we think he should have been. On the other hand, he seems to have been more aware of, and more interested in, the emergence of scholastic theology in northern France than we can easily explain.<sup>3</sup>

---

<sup>2</sup> Noonan.

<sup>3</sup> See Stephan Kuttner, "The Father of the Science of Canon Law," *Jurist* 1 (1941): 2-19. Kuttner argues is that Gratian applied Ivo's program as it had been developed in the intervening forty years by scholastic theology, most notably by Abelard, not in its original form: scholastic theology got the program from Ivo, Gratian got it from scholastic theology. See also D. E. Luscombe, "Abelard and the *Decretum* of Gratian," in *The School of Peter Abelard: The Influence of Abelard's Thought in the Early Scholastic Period*, Cambridge Studies in Medieval Life and Thought, new series, v. 14 (London: Cambridge U.P., 1969), 214-23.



You do not need to exhaustively catalog all of the Gratian myths that Noonan debunked, one is enough. As an example of his approach and conclusions, show how Noonan debunked one of the most prominent elements of the mythology surrounding Gratian, the claim that Gratian was a Camaldoles monk.

Odofredus (+1265), a professor of law at Bologna, claimed that Gratian “was staying”

(stabat) at the monastery of SS. Felice and Nabor while working as a rubricator.<sup>4</sup>

Odofredus did not indicate the basis for his statement that Gratian had been a resident

of SS. Felice and Nabor, nor did he say that Gratian had been a monk. In 1758,

Camaldoles scholars Mittarelli and Costadono discovered papal privileges conferred

in 1113 and 1153 identifying SS. Felice and Nabor as Camaldoles, and inferred from

them that the monastery – and by extension Gratian himself – had been Camaldoles

---

<sup>4</sup> Noonan identifies Odofredus as the source for associating Gratian with SS. Felix and Nabor (Noonan, “Gratian Slept Here.”, 148), but cf. R. W. Southern, *Scholastic Humanism and the Unification of Europe*, vol. 1 (Oxford, UK ; Cambridge, Mass., USA: Blackwell, 1995), 286n5. Southern draws attention to an inscription found in a Geneva manuscript of the *Decretum*, Bibliothèque publique et universitaire, MS Lat. 60: *Anno domini MCL a Gratiano S. Feliciani Bononiensis monacho editum*. Pennington does not identify the manuscript as one known to date from the twelfth century (see Kenneth Pennington, *Gratianus, Bio-Bibliographical Guide to Medieval and Early Modern Jurists*), and of course there is no way of knowing how long after the manuscript was produced the inscription was added. The value of the inscription as independent evidence attesting to a connection between Gratian and SS. Felix and Nabor is therefore unknown, but it is possible that it predates Odofredus.



during the intervening years. The claim that Gratian had a connection to the Camaldolesse order attracted criticism in the eighteenth century, and fell out of favor in the nineteenth century, but was put back into circulation in the mid-twentieth century by Van Hove.<sup>5</sup>

Noonan (†2017) was a distinguished professor of law and an appellate judge who applied the standards of evidence of the law school and the courtroom to the historical record. **Noonan relied heavily on the legal hearsay standard as understood and applied by twentieth-century American courts of law as the basis for his debunking of much Gratian mythology.** By Noonan's standards, the statement of Odofredus – hearsay, not attributed to a source, a century after the fact – is nothing more than legend. **Winroth, however, has questioned whether this is an appropriate standard for evaluating historical evidence.**

---

<sup>5</sup> Cite Noonan, "Gratian Slept Here.", but also cite Alphonse van Hove, *Prolegomena Ad Codicem Iuris Canonici*, Editio artera auctior et emendatior, *Commentarium Lovaniense in Codicem Iuris Canonici*, v. 1, t. 1 (Mechliniae: H. Dessain, 1945). Van Hove's *Prolegomena* had an outsized impact on the trajectory of medieval canon law studies in twentieth century North America because of its having been used by the 1964 Yale seminar led by Peter Landau in the absence of Stephan Kuttner, who was undergoing treatment for cancer at the time. Seminar participants including Charles Donahue and Robert Somerville went on to have unusually influential academic careers in the field (Donahue at Harvard, and Somerville at Columbia where he was Winroth's dissertation advisor).



To conclude, we have reason to believe that Gratian composed and commented upon a substantial portion of the Concordia. In such composition and commentary he revealed himself to be a teacher with theological knowledge and interests and a lawyer's point of view. He worked in Bologna in the 1130s and 1140s. Beyond these conclusions, we have unverified hearsay, palpable legend, and the silent figure in the shadows of S. Marco.<sup>6</sup>

**Transition: Noonan wrote before the discover of the first recension; recent (post-1996) discussion of the identity of Gratian necessarily focus on accounting for the existence of the first and second recensions.**

### **Winroth's argument**

(Pennington's argument is largely a response to Winroth's, discuss Winroth's argument first, followed by Pennington's critique of and response to it.)

"In the interests of simplicity and clarity, I have therefore chosen to call the author of the first recension Gratian 1 and the author of the second recension Gratian 2. These

---

<sup>6</sup> Noonan, "Gratian Slept Here," 172.



labels are not intended to suggest that Gratian 1 and Gratian 2 could not have been the same person.”<sup>7</sup>

Winroth’s arguments in chapters 1-5 of *The Making of Gratian’s Decretum* are agnostic as to whether Gratian 1 is the same person as Gratian 2. The question of authorial identity is not relevant to the primary argument that Winroth advances in those chapters, which is that the Admont, Barcelona, and Florence manuscripts represent a coherent first recension of the *Decretum* rather than an abbreviation. In chapter 6 and in his conclusion, however, Winroth observes that the second recension reflects a considerably more sophisticated legal culture than that reflected in the first recension, notably in the reception of concepts and terminology from the study of Roman law, and argues that it is unlikely that one author underwent such an intellectual evolution in a single generation.

Fourteenth International Congress of Medieval Canon Law, Toronto, 2012.<sup>8</sup>

---

<sup>7</sup> Anders Winroth, *The Making of Gratian’s Decretum* (Cambridge: Cambridge University Press, 2000), 122.



Winroth interprets an entry in a Siena necrology (Siena, Biblioteca comunale degli Intronati F12, f.5r) as evidence that Gratian died as bishop of Chiusi on the feast day of St Lawrence (August 10) in either 1144 or 1145.

## **Pennington's argument**

### **No transition, and very incomplete**

Pennington, depending on Eichbauer, argues that because the number of additional canons found in the margins of Bc, and in both the margins and appendices of Fd and Aa, fall well short of the number that would be expected if they had been copied from a Vulgate *Decretum* (87 are missing from Aa, and 62 are missing from Fd), these manuscripts represent an intermediate “next-to-last” stage in the development of the *Decretum* text intermediate the pre-Vulgate and Vulgate versions. Furthermore, the fact that the Lateran II canons *are* present in this “next-to-last” stage undermines the argument going back to Fransen and supported by Lenherr that the Lateran II canons

---

<sup>8</sup> Anders Winroth, “Where Gratian Slept: The Life and Death of the Father of Canon Law,” *Zeitschrift Der Savigny-Stiftung Für Rechtsgeschichte: Kanonistische Abteilung* 99, no. 1 (August 2013): 105–28.



are a last-minute addition, and poorly if at all integrated into the argument of the *Decretum*.<sup>9</sup>

## **Gratian's teaching career**

Evidence concerning Gratian's career as a teacher of canon law is contradictory. Most medieval and modern readers of the *Decretum* have seen it as a teaching text, both in the sense that its intended use (or at least one of its intended uses) was as an introductory textbook, and also in the sense that it reflected the author's experience as a classroom teacher. On the one hand, the early manuscript tradition for the *Decretum* displays enough variation, even among the first-recension manuscripts Aa, Bc, Fd, and P, to suggest that the text developed continuously over the span of, and in parallel with, a relatively long teaching career — perhaps a decade or more. On the other hand, Gratian's personal obscurity is so profound — Stephan Kuttner compared Gratian to

---

<sup>9</sup> Kenneth Pennington, "The Biography of Gratian, the Father of Canon Law," *Villanova Law Review* 59 (2014): 686.



the anonymous builders of Gothic cathedrals<sup>10</sup> — as to suggest a short, or even no, teaching career.<sup>11</sup>

Considering the fact that Gratian was universally credited as the author of the foundational textbook for the study of canon law in the classical period, one of the truly stellar intellectual achievements of the renaissance of the twelfth century, his complete

---

<sup>10</sup> "He is one of those deeply admirable men of pious modesty, like the builders of the medieval cathedrals, whose person is hidden and effaced completely behind their immortal works." Kuttner, "The Father of the Science of Canon Law," 4.

<sup>11</sup> **Expand: Winroth argues for a short teaching career, Southern argues for no teaching career.** R.W. Southern is the extreme outlier in this regard, citing as late as 1995 Gratian's supposed association with the Camaldoleses order to advance an argument that he had never been a teacher: "Reliable writers after his death affirmed that Gratian was a Benedictine monk of the Camaldolese Order, a small Italian branch of the Benedictines, which combined austerity of life with considerable freedom to contribute to the work of the Church in a wide variety of ways. This is what Gratian did. He engaged in a very un-monastic branch of study, and produced a law book which became the most widely distributed and probably most generally used of all the masterpieces of twelfth-century learning." "the book, and not the teaching of pupils, was the offering Gratian's monastic life brought to the Church". Southern, *Scholastic Humanism and the Unification of Europe*, 1:286–87, 304. Southern also asserts that no one claims to have been Gratian's student — "none of the earliest users of his work claims to have been taught by him". "although other scholars made some very early alterations and additions to the text of the *Decretum*, none of them claimed to have been taught by him." Southern, 1:287, 304. **Does Southern provide footnotes to support these statements?**



absence from the kinds of academic genealogies and origin stories that Odofredus (+1265) handed down to his own Roman law students is striking indeed.<sup>12</sup>

We might be better able to assess Gratian's teaching career and how it influenced the compilation and transmission of the *Decretum* were we in a position to positively identify at least a few of his students. Here again, however, the evidence is mixed at best. Three twelfth-century canonist have, with varying degrees of plausibility, been identified as having been students of Gratian: Paucapalea (fl. 1140s), Simon of Bisignano (fl. 1174–1179), Omnebene (d. 1185).

### **Paucapalea**

Paucapalea was easily the most prominent figure from the early period of the Bolognese school of decretists who followed in Gratian's immediate footsteps. Pennington flatly

---

<sup>12</sup> **Expand:** The tradition of Roman law teaching in Bologna preserved an elaborate genealogy of masters and their students: Pepo preceded Wernerius [Irnerius], who taught the Four Doctors. Odofredus is an important source for this tradition, and his remarks about Irnerius can be found in his commentary on *Dig. 1.1.6.* Odofredus, *Lectura super Digesto veteri*, (2 volumes, Lyon: 1550-1552; repr. Bologna 1967-1968). For a short excerpt from the Latin text with accompanying English translation, see Charles Radding, *The Origins of Medieval Jurisprudence: Pavia and Bologna, 850- 1150* (New Haven: Yale University Press, 1988), 159–60. **See also Pennington “Big Bang”.** Note that Odofredus is also the source for associating Gratian with the monastery of SS. Felix and Nabor. Noonan, “Gratian Slept Here,” 148.



states that Paucapalea was “[a] pupil of Gratian and the author of the first commentary on the Decretum.”<sup>13</sup> But in the preface to that commentary, Paucapalea does not refer to Gratian by name: “*Magistri autem hoc opus condentis ipsa decreta sunt materia.* The decrees themselves, however, are the subject matter of the master producing this work.”<sup>14</sup>

The omission is potentially significant. Prefaces to academic books in the twelfth century formed a recognized genre with well-established conventions. In an age resistant to any attempt by the human present to appropriate the authority of the divinely-inspired past, a commentator like Paucapalea would have been expected to offer an explicit statement in his preface affirming the authority of both the work commented on and its author. If Paucapalea had had something to say about the authority of Gratian, his readers would have expected him to do so in his preface. That

<sup>13</sup> Kenneth Pennington, [Paucapalea](#), *Bio-Bibliographical Guide to Medieval and Early Modern Jurists*.

<sup>14</sup> Johann Friedrich von Schulte, ed., *Die Summa Des Paucapalea über Das Decretum Gratiani* (Giessen: E. Roth, 1890), 3. Robert Somerville and Bruce Clark Brasington, eds., *Prefaces to Canon Law Books in Latin Christianity: Selected Translations, 500-1245* (New Haven, Conn: Yale University Press, 1998), 184. Indeed, based on a search through the OCR text layer of a PDF file scanned from the Schulte edition and provided to me by Ken Pennington, Paucapalea appears not to refer to Gratian by name anywhere in his *summa*.



he did not suggests that he may not have known enough about Gratian to credibly do so.<sup>15</sup>

Furthermore, Paucapalea in the preface to his *summa* clearly shows both that he is aware of the third part of the *Decretum, de Consecratione*, and that he uncritically accepts it as having been compiled by the same Master responsible for the rest of the work: "*In extremis de ecclesiarum dedicatione et corpore et sanguine domini atque baptismo nec non et confirmatione pleniter tractat, et in his suum tractatum terminat.* In the end, he treats fully the dedication of churches, the body and blood of the Lord, baptism, and also confirmation, and with these things concludes his treatise."<sup>16</sup>

If Rambaud is correct that *de Consecratione* was not part of the original *Decretum*, the fact that Paucapalea attributed it to the same unnamed master responsible for rest of the text

---

<sup>15</sup> Going to need a very heavyweight footnote here referencing Minnis. A. J. Minnis, *Medieval Theory of Authorship: Scholastic Literary Attitudes in the Later Middle Ages*, 2nd ed, Middle Ages Series (Philadelphia: University of Pennsylvania Press, 1988). Richard William Hunt, "The Introductions to the 'Artes' in the Twelfth Century," in *The History of Grammar in the Middle Ages: Collected Papers*, ed. G. L. Bursill-Hall, Amsterdam Studies in the Theory and History of Linguistic Science. Series 3: Studies in the History of Linguistics v.5 (Amsterdam: J. Benjamins, 1980), 117–44

<sup>16</sup> Schulte, *Die Summa Des Paucapalea iiber Das Decretum Gratiani*, 3. Somerville and Brasington, *Prefaces to Canon Law Books in Latin Christianity*, 184–85



on which he was commenting strongly suggests that he did not have direct knowledge of Gratian or his plan of work.

### **Omnibonus or Omnebene**

Omnibonus or Omnebene (d.1185) composed an abbreviation (*summa?* abridged version?) of the *Decretum* in the 1150s.<sup>17</sup> The basis for the claim that Omnebene was a student of Gratian is an entry for the year 1130 in the chronicle of Norman monk Robert of Torigni (†1186)

*Gratianus episcopus Clusinus coadunavit decreta valde utilia ex decretis, canonibus, doctoribus, legibus Romanis, sufficientia ad omnes ecclesiasticas causas decidendas, que frequentantur in curia Romana et in aliis curiis ecclesiasticis. Hec postmodum abbreviavit magister Omnebonum episcopus Veronensis, qui fuerat eius discipulus.* (Gratian bishop of Chiusi brought together very useful decrees out of decrees, canons, doctors, Roman laws, sufficient for deciding all ecclesiastical cases, which are resorted to at the Roman curia and in other ecclesiastical courts. Master Omnebonum

<sup>17</sup> Kenneth Pennington, [Omnibonus](#), *Bio-Bibliographical Guide to Medieval and Early Modern Jurists*.



bishop of Verona, who had been his student, afterwards abbreviated them.)<sup>18</sup>

Although the evidence of Robert's chronicle is by Noonan's standard hearsay, Robert's reputation as a well-informed source concerning Gratian has improved in recent years since the Siena necrology has made the identification of Gratian as bishop of Chiusi more probable.

The dates do not preclude Omnebene's having been a student of Gratian. Although Omnebene lived into the 1180s, the abbreviation itself seems to date from the 1150s, perhaps a decade after the end of whatever teaching career Gratian may have had, and at precisely the moment that continent-wide interest in the text was increasing most rapidly.

**Is there any internal evidence in Omnebene's abbreviation that Gratian was his teacher? Cite recent editions of Omnebene's abbreviation.**

---

<sup>18</sup> Kuttner, *Repertorium*, 453. MGH Scriptores (SS) 6 (1844), 490. *Chronique de Robert de Torigny*, ed. Léopold Delisle, 183-184 (see BSB online version). *Chronicles of the reigns of Stephen, Henry II, and Richard I*, ed. Richard Howlett (1884). *Chronica Roberti de Torigneio*, ed. Richard Howlett, RS 82 4.118.



## Simon of Bisignano

Simon of Bisignano (fl. 1174–1179) was the author of a *Summa decretorum* dated c.1177–79, and a number of glosses on the *Decretum* are attributed to him.<sup>19</sup> The evidence that Simon of Bisignano was a student of Gratian's comes from references within his writings.<sup>20</sup> Simon refers to himself as *discipulus* and to Gratian as *magister* in at least two

---

<sup>19</sup> Kenneth Pennington, [Simon of Bisignano](#), *Bio-Bibliographical Guide to Medieval and Early Modern Jurists*. Pennington does not indicate that Simon of Bisignano was a student of Gratian. Glosses attributed to Simon of Bisignano are indicated by the sigla s., si., and sy. Rudolf Weigand, "The Development of the Glossa Ordinaria to Gratian's *Decretum*," in *The History of Medieval Canon Law in the Classical Period, 1140–1234: From Gratian to the Decretals of Pope Gregory IX*, ed. Wilfried Hartmann and Kenneth Pennington, *History of Medieval Canon Law* (Washington, D.C: Catholic University of America Press, 2008), 97. **Simon's glosses are dated earlier than his *Summa* (cite Hartmann and Pennington).**

<sup>20</sup> Stephan Kuttner, "Research on Gratian: Acta and Agenda," in *Studies in the History of Medieval Canon Law*, Collected Studies CS325 (Aldershot, Hampshire, Great Britain : Brookfield, Vt., USA: Variorum ; Gower, 1990), 7. Noonan, "Gratian Slept Here.", 157, rejects this, but Kuttner disagrees. Simon of Bisignano refers to Gratian as *magister* and to himself as *discipulus* in his comments on C.27 q.2, Simon of Bisignano, *Summa in Decretum Simonis Bisignianensis*, ed. Pier Virginio Aimone-Braida, *Monumenta Iuris Canonici. Series A, Corpus Glossatorum* ; Vol. 8 (Città del Vaticano: Biblioteca Apostolica Vaticana, 2014), 412. Simon again refers to Gratian as *magister* in his comments on D.4 c.31 of *de Consecratione* Simon of Bisignano, 542. **Page numbers may not coincide with print edition cited: the version of the Aimone edition I have access to is a PDF dated Fribourg 2007. BibTex records in merged.bib are inconsistent as to whether to use the author's or editor's name as the basis for the tag.**

<Q.II> § *Sequitur secunda questio*

Sacram et authenticam pie recordationis Gratianus de matrimonio distinctionem composuit, quam quidam inanis glorie aucupes fecere bifrontem et simplicium potibus inuidentes more seuorum animalium aquas limpidas turbauerunt. Gratianus uero, canonum habita notitia et mente perspecta, dixit



places in his *Summa* (**Landau says “frequently”**). There is, however a hiatus of almost four decades between the latest plausible date for the end of Gratian’s teaching career, and the documentary evidence we have for Simon’s professional career as a canonist. Although it is not chronologically impossible for Simon to have been Gratian’s student, it seems more likely that he described himself as *discipulus* in something like the same

---

inter sponsum et sponsam esse matrimonium, sed initiatum; inter uirum et uxorem consummatum. Sed quoniam non est discipulus super magistrum nec debet securis contra eum qui secat gloriari timemus os nostrum in celum mittere et eum reprehendere cuius non sumus digni soluere corrigiam calceamenti.

Note that the phrase “sed inanis glorie aucupes fecere bifrontem” also occurs in Rufinus’s treatment of the same question, C.27 q.2. Rufinus, *Summa Decretorum*, ed. Heinrich Singer (Aalen : Paderborn: Scientia Verlag ; F. Schöningh, 1963), 440.

Gratian of pious remembrance composed a sacred and *authenticam* distinction concerning marriage, *quam quidam inanis glorie aucupes fecere bifrontem* and not seeing with the drinks of the simple, [they] troubled clear waters in the manner of *seuorum* animals. But Gratian *habita notitia* of the canon and *mente perspecta*, said that matrimony is between spouse and spouse, but having been consecrated (*initiatum*), is consummated between husband and wife. But since the disciple is not above his master (Matthew 10:24, Luke 6:40), and the axe (*securis*) ought not boast (*gloriari*) against him who cuts (*secat*) (Isaiah 10:15), we fear to send our mouth into heaven (Psalm 72:9?) and to find fault with him the string (*corrigiam*) of whose shoes we are not worthy to loosen (John 1:27) (PLE).

### 31. Solet queri

Quod quamuis Magister noster docuerit, nos tamen in hoc nolumus eum imitari.

Although our master taught that, we nevertheless do not wish to imitate him in this (PLE).



sense that we might describe contemporary scholars like Chodorow, Pennington, and Winroth as “students of Gratian”.

## Epilogue

By the beginning of the thirteenth century, the systematic study of canon law had evolved in a direction of which its father Gratian would probably not have approved, and which he may even have had difficulty recognizing as the same practical or applied branch of theology his own work presupposed it to be. After around 1190, the attention of both academic and practicing canonists decisively turned away from theological jurisprudence in the tradition of Gratian, based on a rich collection of scriptural, patristic, conciliar, and papal source texts, and towards a much more narrowly circumscribed technical jurisprudence focused exclusively on contemporary papal decretals.



## Chapter 4

### Stylometry

The contemporary audience for Gratian's *Decretum* seems to have been reasonably satisfied that a single author was responsible for the collection as they knew it, even if they knew so little about the historical Gratian that they felt the need to retroactively provide a largely fictional backstory for the eponymous author. Modern students of Gratian, on the other hand, have been willing to entertain the possibility that the *Decretum*, at least in its final most widely-circulated form, was the product of collective authorship. Stephan Kuttner, as the first item on his 1984 agenda for Gratian studies, asked:

was [the *Concordia discordantium canonum*] drafted and completed in one grandiose thrust, or did the original version go through successive redactions? And if the latter, was it Gratian himself, or Gratian with his disciples, or an early generation of canonists after him, who completed the

final recension which from the mid-twelfth century on was used in the schools and in adjudging cases?<sup>1</sup>

Anders Winroth's 1996 discovery of the first recension of the *Decretum* provided a convincing answer to the first part of Kuttner's question: the original *did* go through successive redactions. Winroth used the name Gratian 1 to refer to the compiler or compilers of the first recension, and Gratian 2 to refer to the compiler or compilers of the second recension. He then reformulated the second part of Kuttner's question by asking whether Gratian 1 was the same person as Gratian 2. Winroth's answer to the question posed in this form, that Gratian 2 was *not* the same person as Gratian 1, has been the focus of vigorous but so far inconclusive scholarly debate over the last twenty years. Much of the debate has focused on competing evaluations of the Sankt Gallen Stiftsbibliothek 673 (Sg) manuscript of the *Decretum*, and more recently on secondary evidence such as a twelfth-century Siena necrology in which the name Gratian appears.

---

<sup>1</sup> Stephan Kuttner, "Research on Gratian: Acta and Agenda," in *Studies in the History of Medieval Canon Law*, Collected Studies CS325 (Aldershot, Hampshire, Great Britain : Brookfield, Vt., USA: Variorum ; Gower, 1990), 10.



It does not appear that the debate over the authorship of the *Decretum* can be settled on the basis of the currently available evidence. The goal of my dissertation project has been to find new evidence relevant to the question of the authorship the *Decretum* as it has been posed by both Kuttner and Winroth, by using computational stylometric methods to analyze the authorship of the *dicta* traditionally attributed personally to Gratian.

## Preliminary observations

Stylometry is the measurement of style. "Style is a property of texts constituted by an ensemble of formal features which can be observed quantitatively or qualitatively."<sup>2</sup> While style has both qualitative and quantitative aspects, stylometry is concerned only with quantitative aspects of style. One well-established use of stylometry is to attribute authorship. And for the purpose of authorship attribution, the formal linguistic features that stylometry measures are the frequencies of occurrence of common words.

---

<sup>2</sup> J. Berenike Herrmann, Karina van Dalen-Oskam, and Christof Schöch, "Revisiting Style, a Key Concept in Literary Studies," *Journal of Literary Theory* 9, no. 1 (2015): 44.



Linguists draw a distinction between function words and content words. The more frequently a word occurs in a language, the more likely it is to be a function word, and the less likely it is to be a content word. Function words are words like prepositions and conjunctions. Content words are words like adjectives, nouns, and verbs. Function words convey meaning by their use in grammatical structure. The Latin conjunction “*sed*” does not mean anything by itself, but rather it places two words or grammatical constructs into an adversative relationship with each other.

Another way of thinking about the distinction is to note that in a given language, function words constitute a closed class, while content words constitute an open class. Language-speaking communities can and do make up new adjectives, nouns, verbs all the time; content words are therefore an open class that can be added to at will. But new prepositions and conjunctions are almost never added to a language, and their usage changes very slowly over time, if it changes at all, and function words are therefore, for all practical purposes, a closed, finite, class.

Evidence from experimental psychology suggests that both authors and readers process function words at a subconscious level.<sup>3</sup> The frequency with which a given author uses particular function words is therefore considered to be more or less invariant, making it a reliable authorial signature.

**Some stylometric techniques use only function words, while others use common words generally, without regard to whether they are function or content words, but in either case function words are disproportionately important to stylometric analysis because of their observed frequency.**

### **The Federalist (Hamilton and Madison)**

Stylometric analysis of the frequencies of common words for the purpose of attributing authorship has had a number of notable successes. The validity of this approach for textual scholarship was firmly established by the work of Frederick Mosteller and David L. Wallace on the *Federalist Papers*. The authorship of 12 of the *Federalist Papers*,

---

<sup>3</sup> Mike Kestemont, "Function Words in Authorship Attribution from Black Magic to Theory?" in *Proceedings of the 3rd Workshop on Computational Linguistics for Literature (CLfL)* (Gothenburg, Sweden: Association for Computational Linguistics, 2014), 59–66.



49–57 and 62–63, had been disputed since the early 19th century, with competing claims advanced on behalf of Alexander Hamilton and James Madison.<sup>4</sup> In 1944, Douglass Adair, using traditional scholarly methods, settled the dispute largely to the satisfaction of early American historians, determining that Madison was the author of all 12 of the disputed numbers.<sup>5</sup> In 1964, Mosteller and Wallace confirmed Adair's findings by conducting a stylometric analysis of the frequencies of 70 function words to compare the 12 disputed numbers with numbers securely attributed to Hamilton and Madison.<sup>6</sup>

---

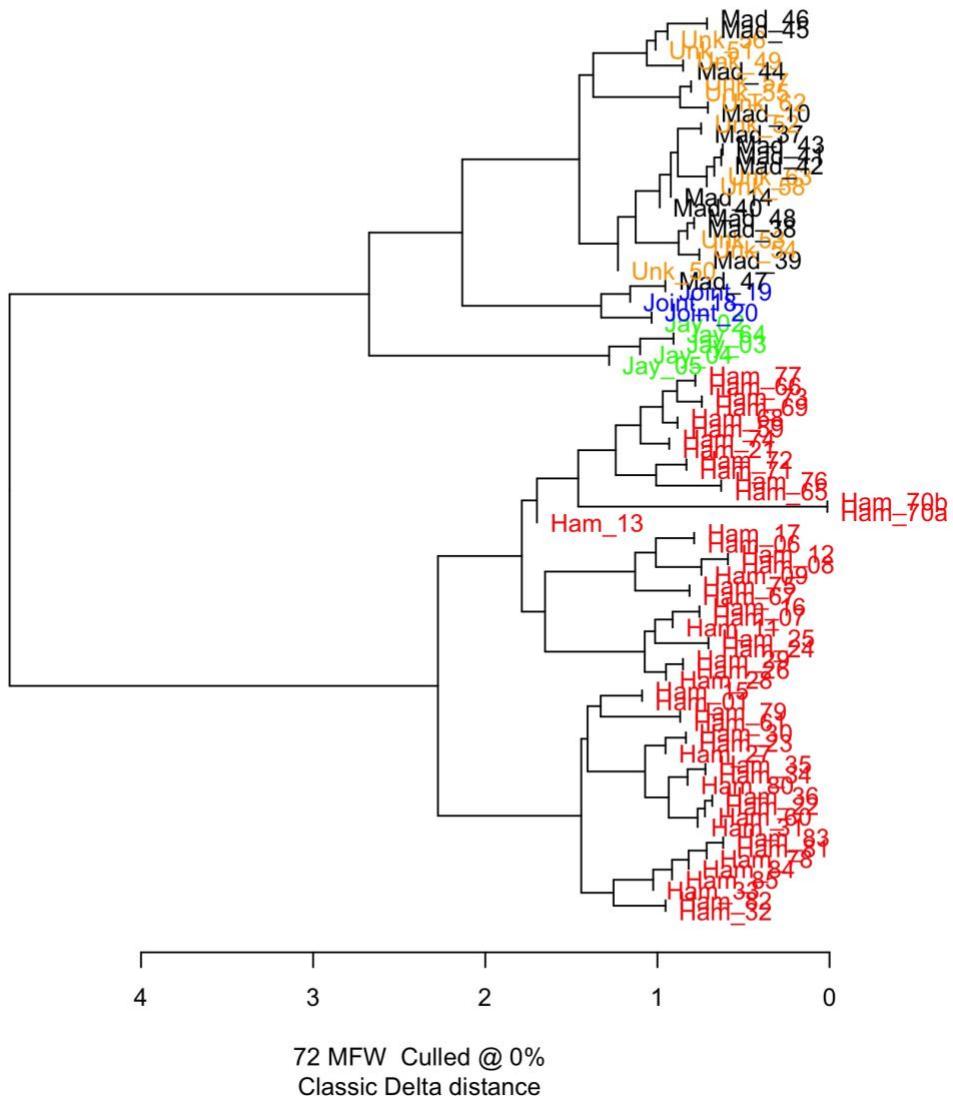
<sup>4</sup> Frederick Mosteller and David L. Wallace, *Inference and Disputed Authorship: The Federalist*, Addison-Wesley Series in Behavioral Science: Quantitative Methods (Reading, Mass: Addison-Wesley Pub. Co, 1964), 14. See also Douglass Adair, "The Authorship of the Disputed Federalist Papers," *The William and Mary Quarterly* 1, no. 2 (1944): 98–122, 104.

<sup>5</sup> Adair and Douglass Adair, "The Authorship of the Disputed Federalist Papers: Part II," *The William and Mary Quarterly* 1, no. 3 (1944): 235–64.

<sup>6</sup> Mosteller and Wallace, *Inference and Disputed Authorship*.



### Federalist Cluster Analysis



*Federalist*

## Definition of first- and second-recension *dicta*

**Abstract:** In the absence of good modern critical editions for the first and second recensions of the *Decretum*, a proxy for the first recension is created by applying the variants recorded in Winroth's appendix as transformations to Friedberg's text. The proxy for the first recension is then subtracted from Friedberg's text, leaving text from the second recension as the difference. The text samples from the first and second recensions of the *Decretum* that provide the basis for authorship attribution are built up by iteratively appending short units of text (the individual first- and second-recension *dicta*) that are non-contiguous in the original context of the *Decretum*.

Many of the examples in this chapter will distinguish between first- and second-recension *dicta*, so this is an appropriate point at which to introduce an explicit definition for the way in which those terms will be used in the following discussion. Because almost every word in the first-recension *dicta* corresponds to a word in the second-recension *dicta*, we could consider the first-recension *dicta* to be a subset of the second-recension *dicta*, and conversely, we could consider the second-recension *dicta* to



be a superset of the first-recension *dicta*. While true enough from a commonsensical point of view, this is not a useful definition for the kinds of questions we would like to be able ask and answer, such as whether Gratian 1, the author of the first-recension *dicta*, is the same person as Gratian 2, the author of the second-recension *dicta*.

Instead, for the purpose of the following analyses, the second-recension *dicta* are defined as the ordered set of every word from the *dicta* in the text of Friedberg's 1879 edition of Gratian's *Decretum* for which there is not a one-to-one correspondence to a word in the first-recension *dicta* as defined by Anders Winroth's appendix "The Contents of the First Recension of Gratian's *Decretum*.<sup>7</sup>" An alternative restatement would be to define the second-recension *dicta* as the difference left by subtracting all of the words of the first-recension *dicta* as defined by Winroth's appendix from the text in the Friedberg edition. This definition is implemented by passing sequentially through the *dicta* and applying the following three rules:

---

<sup>7</sup> Anders Winroth, *The Making of Gratian's Decretum* (Cambridge: Cambridge University Press, 2000), 197–227.



- If a *dictum* is listed in Winroth's appendix as being in the first recension of the *Decretum*, and as not having been added to or changed in the second recension, the text for that *dictum* is included in the first recension sample. This rule is applied on a per-*dictum* basis.
- If a *dictum* is in the text of the Friedberg edition, and is not listed in Winroth's appendix as being in the first recension, in either unmodified or modified form, the text for that *dictum* is included in the second recension sample. This rule is applied on a per-*dictum* basis.
- If a *dictum* is listed Winroth's appendix as being in the first recension, but as having been added to or changed in the second recension, those words indicated by the appendix are included in the first recension sample, while those words in the text of Friedberg not corresponding to the words indicated by the appendix are included in the second recension sample. This rule is applied on a word-by-word basis.

Take D.54 d.p.c.23 as an example. The complete text of the *dictum* as it appears in the Friedberg edition (column 214) is:



*Ecce, quomodo serui ad clericatum ualeant assumi, uel quomodo non admittantur. Liberti quoque non sunt promouendi ad clerum, nisi ab obsequiis sui patroni fuerint absoluti. Unde in Concilio Eliberitano:*

Winroth's appendix indicates that only the first sentence of the *dictum* appears in the first recension:

d.p.c. 23: 1 *Ecce quomodo serui – 2 quomodo non admittantur.*<sup>8</sup>

Therefore, “*Ecce, quomodo serui ad clericatum ualeant assumi, uel quomodo non admittantur.*” is included in the first recension text sample, and “*Liberti quoque non sunt promouendi ad clerum, nisi ab obsequiis sui patroni fuerint absoluti. Unde in Concilio Eliberitano:*” is included in the second recension text sample.

Note that the individual *dicta* are too short for direct analysis by the techniques discussed in this chapter.<sup>9</sup> The smallest unit of Latin prose for which computational

<sup>8</sup> Winroth, *The Making of Gratian's Decretum*, 201. The numbers 1 and 2 refer to line numbers relative to the first line of the *dictum*, as opposed to the first line of the column, in the print version of the Friedberg edition.

<sup>9</sup> This paragraph may have to be moved to a separate section explaining the rationale for the various roll-ups of the *dicta* used in the following analyses, e.g., *Gratian0*, *Gratian1*, *dePen*, *Gratian2*, etc.

styliometry works is about 2,500 words.<sup>10</sup> The longest first-recension *dictum* (*de Pen.* D.1 d.p.c.87) is 1,591 words, and the longest second-recension *dictum* (C.7 q.1 d.p.c.48) is 692 words. As a result, first and second recension samples long enough for analysis have to be created by rolling up or concatenating the first and second recension *dicta* as they occur sequentially but discontinuously throughout the *Decretum*.

Both the text of Friedberg's 1879 edition of Gratian's *Decretum* and Winroth's appendix are potentially open to criticism. Although modern scholars admire Friedberg's learning and energy — the 1879 edition of the *Decretum* was only one of many such projects that he undertook — his editorial standards were those of 140 years ago. In particular, Friedberg's selection of eight unrepresentative German manuscripts as the basis for his edition, and his particular reliance on two of them — Köln Erzbischöfliche Diözesan- und Dombibliothek 127 (Ka) and 128 (Kb) — are seen today as serious deficiencies.<sup>11</sup>

<sup>10</sup> Maciej Eder, "Does Size Matter? Authorship Attribution, Small Samples, Big Problem," *Literary and Linguistic Computing* 30, no. 2 (June 2015): 171.

<sup>11</sup> Winroth, *The Making of Gratian's Decretum*, 9-11. See also Stephan Kuttner, "De Gratiani Opere Noviter Edendo," *Apollinaris* 21 (1948): 118–28 (Latin), and "Research on Gratian.", 10, 21-22, which mentions the deficiency of Friedberg's edition without offering a detailed critique.



Winroth himself acknowledged the provisional nature of his appendix.<sup>12</sup>

Furthermore, Pennington has pointed out that although Winroth's appendix includes D.100 d.a.c.1, D.100 c.1, and D.101 d.p.c.1, in the Paris (P), Florence (Fd), and Barcelona (Bc) manuscripts, the text of the first recension ends with D.99 c.1.<sup>13</sup>

Nevertheless, in the absence of a critical edition for the first recension, applying the variants recorded in Winroth's appendix as a set of transformations to the text of Friedberg's edition to generate a stand-in or proxy for the text of the first recension is a workable approach.<sup>14</sup> This method is well-theorized in a Digital Humanities context as an example of "deformance." The term, proposed by Lisa Samuels and Jerome McGann in "Deformance and Interpretation" (1999), conflates the words "deformation" and "performance", and describes a process through which a text is transformed by the

---

<sup>12</sup> "The list is based on a collation of *incipits* and *explicits* of every canon and *dictum* in the first recension. Differences within the texts may very well have been overlooked, and minor differences have not normally been registered." Winroth, *The Making of Gratian's Decretum*, 197.

<sup>13</sup> Kenneth Pennington, "The Biography of Gratian, the Father of Canon Law," *Villanova Law Review* 59 (2014): 685.

<sup>14</sup> **Anders, my understanding is that you took a similar approach when you created the baseline text for the edition in progress of the first recension (although you adopted a set of orthographic conventions different from Friedberg's in the resulting text).**



application of a series of deformances to generate a “paratext”.<sup>15</sup> The paratext is different from the original text, but defined by the deformances through which it was generated from it, and can be analyzed for otherwise unavailable insights into the original text.

The methodology for producing the text samples used in this project involves multiple stages of deformation. Starting with Friedberg’s 1879 edition of Gratian’s *Decretum* as the text, Winroth’s appendix, which compactly encodes first-recension variants with respect to Friedberg, is used as a program (literally, as will be seen in the section on corpus preparation below) for deforming Friedberg’s text to produce the first paratext, the proxy first recension *dicta*. The first paratext is then used as the basis for a second deformation, by which the first paratext is subtracted from Friedberg’s text to create the second paratext representing second recension additions and changes to the *dicta*.

The approach of deriving all of the text samples used in this study using only Friedberg’s text and the first-recension variants recorded in Winroth’s appendix as

---

<sup>15</sup> Lisa Samuels and Jerome McGann, “Deformance and Interpretation,” *New Literary History* 30, no. 1 (1999): 25–56.



sources has one final argument in its favor, which is that it enables reproducibility.

Reliance on publicly available data means that those who wish to reproduce these results are not dependent on private decisions about the content of the text samples.

Depending on the nature of the analysis we wish to conduct, we may choose to either include or exclude the *dicta* from *de Penitentia*. Including the *dicta* from *de Penitentia*, there are 897 *dicta* represented in the first-recension text sample and 419 represented in the second-recension sample. Of those, 65 *dicta* are represented in both the first- and second-recension samples. Excluding the *dicta* from *de Penitentia*, there are 836 first-recension and 398 second-recension *dicta*, of which 61 *dicta* are represented in both samples.

## Corpus preparation

The most important and time-consuming aspect of any digital humanities project is corpus preparation. The availability of a suitable corpus of electronic texts is a baseline requirement for carrying out stylometric analysis. The ideal textual basis for a project of this nature would be a set of electronic texts of good modern critical editions of both the first and second recensions of Gratian's *Decretum*, following consistent orthographic



conventions, and adhering to a widely-accepted encoding standard such as the XML Text Encoding Initiative (TEI P5) format. The Mellon Foundation-supported effort directed by Anders Winroth to edit the first recension is ongoing, but work on Winroth's edition in progress had not reached a sufficiently advanced state of completion for it to be used within the time frame available for my project.<sup>16</sup>

As a result, this investigation depends for both the first and the second recension on the electronic text of the Friedberg edition that Timothy Reuter and Gabriel Silagi used to produce the *Wortkonkordanz zum Decretum Gratiani* for the MGH.<sup>17</sup> Anders Winroth and Lou Burnard of the Oxford Text Archive (OTA) provided copies of the MGH e-text separately. The copies differed significantly, and the e-text had to be reconstructed

---

<sup>16</sup> As of the most recent, 22 April 2019, version, eight case statements (for cases 1-3, 9, 15, 24, 30, and 34) appear to have a complete critical apparatus. An addition six case statements (for cases 4-7, 10, and 11) have an incomplete critical apparatus that records variant readings from Fd only. The critical apparatus for the case statement for case 35 records a single variant reading from Aa. The remaining 21 case statements (for cases 8, 12-14, 16-23, 25-29, 31-33, and 36) have no critical apparatus at all.

**Update for 5 October 2019 version of the edition in progress. The case statement are used as the example here because they are the focus of my dissertation.**

<sup>17</sup> Timothy Reuter and Gabriel Silagi, eds., *Wortkonkordanz Zum Decretum Gratiani*, Monumenta Germaniae Historica. Hilfsmittel 10 (München: Monumenta Germaniae Historica, 1990).



through an editorial process quite similar to preparing a critical edition to restore it to a state as close as possible to what Reuter and Silagi presumably intended.

In 1946, Father Roberto Busa, SJ (d. 2011), began work on what ultimately became the *Index Thomisticus*, a concordance of the works of St Thomas Aquinas. In 1949, Father Busa secured crucial support from Thomas J. Watson of IBM, allowing concordance generation to be carried out by means of electro-mechanical and later electronic computers operating on punch-card data. The *Index Thomisticus* is recognized today as the first important humanities computing project, and figures prominently in origin stories for digital humanities as a discipline.<sup>18</sup> The success of Father Busa's project inspired a number of imitators, as well as the development of specialized software and data formats to support such efforts. Reuter and Silagi's *Wortkonkordanz* was probably

---

<sup>18</sup> Susan Hockey, "The History of Humanities Computing," in *A Companion to Digital Humanities*, ed. Susan Schreibman, Raymond George Siemens, and John Unsworth, Blackwell Companions to Literature and Culture 26 (Malden, MA: Blackwell Pub, 2004), 4–6. The highest honor in the field of Digital Humanities is the Roberto Busa Prize, awarded by the Alliance of Digital Humanities Organizations (ADHO). A notable past recipient of the Busa Prize is John Burrows, who first introduced the fundamental stylometric technique now known as Burrows's Delta in a lecture he delivered on the occasion of receiving the award in 2001.



the last major Busa-style concordance. For this reason, the MGH e-text of the Friedberg edition was encoded in the obsolete Oxford Concordance Program (OCP) format.

The MGH e-text introduced a small number of textual errors in addition to those it inherited from printed version of Friedberg's edition. The table below lists all currently known errors in the MGH e-text:<sup>19</sup>

Citation	Column	Error	Correction
D.6	11	quantam ad moralem	quantum ad moralem
d.p.c.3		intelligentiam	intelligentiam
D.23 c.2	79		carnis passione, mortuus uera corporis sui morte, resurrexit uera
D.54	214	Quid autem serui	Quod autem serui ecclesiarum

<sup>19</sup> Data current as of 23 May 2020. For more recent error reports, see the list I maintain for the Stephan Kuttner Institute on [GitHub](#). Thanks to Anders Winroth for reporting the errors in D.6 d.p.c.3 (6 October 2019) and D.23 c.2 (23 August 2019).



d.p.c.22		ecclesiarum
C.2 q.6	483	contra sententiam
		contra sententiam restitutionem
c.41		restitutionem inpetraverit
		inpetrauerit
C.4 d.init.	536	a communione ait
		a communione sit remouendus
		remouendus
C.4 d.init.	536	in episcoporum indicio
		in episcoporum iudicio
C.7 d.init.	566	Quidam longa
		Quidam longa inualetudine
		inualetudinem grauatus
		grauatus episcopus
		episcopus

The error in D.23 c.2 is particularly noteworthy. Contrary to the widespread belief that the MGH e-text was created by scanning a physical copy of the Friedberg edition using optical character recognition (OCR) technology and then correcting the results, it is the product of keyboard transcription. Conclusively, the e-text contains at least one instance of homeoteleuton. The following fours lines from the text of D.23 c.2 in the Friedberg edition (column 79):

*Patre et Spiritu sancto omnium creaturarum; qui passus  
sit pro salute nostra uera carnis **passione, mortuus uera***



*corporis sui morte, resurrexit uera carnis suae receptione  
et uera animae resumptione, in qua ueniat iudicare*

were transcribed as the following three lines in the e-text:

*Patre et Spiritu sancto omnium creaturarum; qui passus  
sit pro salute nostra uera carnis suae receptione  
et uera animae resumptione, in qua ueniat iudicare*

skipping over the words “*passione, mortuus uera corporis sui morte, resurrexit uera carnis.*”<sup>20</sup>

Notwithstanding its textual flaws and the highly specialized and outdated requirements that constrained the choice of file format, the MGH e-text remains a useful tool for the study of Gratian’s *Decretum*.<sup>21</sup>

<sup>20</sup> Thanks to Anders Winroth for bringing the instance of homeoteleuton at D.23 c.2 in the MGH e-text to my attention (August 23, 2019). Clemens Radl of the MGH confirmed to Winroth that the e-text was typed.

<sup>21</sup> Anders Winroth, “Uncovering Gratian’s Original Decretum with the Help of Electronic Resources,” *Columbia Library Columns* 46, no. 1 (1997): 26–31.



The deformance algorithm used to generate the paratexts described in the previous section on the definition of the first- and second-recension *dicta* was implemented in the form of a 201-line Python program. The program reads the MGH e-text of the Friedberg edition, and parses it to extract the *dicta*. The Oxford Concordance Program (OCP) format in which the e-text is encoded is extremely difficult to parse because it is not tree-structured—it has start tags for textual elements such as canons and *dicta*, cases and distinctions, but not (unlike XML) end tags.<sup>22</sup> The extraction engine captures every element of text between a *dictum* start tag (`<T A>` or `<T P>`) and the start tag for the next element that can possibly follow a *dictum*:

```
import re

f = open('edF.txt', 'r')
file = f.read()

# (?<=...) positive Lookbehind assertion.

dicta = re.findall('(?:<T [AP]\>|(?=<T [AP]\>))(.*?)" # dictum starts
with dictum ante or dictum post tag.
```

---

<sup>22</sup> Hockey, “The History of Humanities Computing.”



```

'(?:'                      # non-capturing group.

'\\<1 [CD][CP]?\\>|'    # dictum ends with major division,
'\\<2 \\d{1,3}\\>|'      # or number of major division,
'\\<3 \\d{1,2}\\>|'      # or number of question,
'\\<4 \\d{1,3}\\>|'      # or number of canon,
'\\<P 1\\>|'            # or Palea,
'\\<T [AIPRT]\\>'       # or inscription or text tag.

')', file, re.S)        # re.S (re.DOTALL) makes '.' special character
match any character including newline.

```

The extracted *dicta* require considerable scrubbing before they can be used. Here, for example, is what D.54 d.p.c.23 looks like in its raw state:

```
[ ' -Gratian.+ Ecce, quomodo serui ad clericatum ualeant assumi,\n
uel quomodo non admittantur. Liberti quoque non sunt promouendi\n
ad clerum, nisi ab obsequiis sui patroni fuerint absoluti.\n
Unde in Concilio Eliberitano: -[c. 80.]+'\n]
```

Each *dictum* is then processed into an item (key-value pair) in a Python dictionary:

```
{'D.54 d.p.c.23': 'Ecce, quomodo serui ad clericatum ualeant assumi, uel
quomodo non admittantur. Liberti quoque non sunt promouendi ad clerum, nisi
ab obsequiis sui patroni fuerint absoluti. Unde in Concilio Eliberitano:'}
```



The first recension variants from the Friedberg edition recorded in Winroth's appendix are then encoded as a list of dictionaries in which the 'pattern' item is the variant represented as a Python regular expression:

```
[{'key': 'D.54 d.p.c.23', 'pattern': '(Ecce, quomodo serui.*?quomodo non  
admittantur\.)'}]
```

Finally, the deformation engine uses the variants encoded as regular expression patterns to generate the first and second paratexts corresponding the first- and second-recension *dicta*. For each *dictum*, the text matching the pattern is inserted into a dictionary representing the first recension paratext; then the text resulting when the text matching the pattern is replaced by the null string '' is inserted into a dictionary representing the second recension paratext:

```
import re

dictionary_1r = {} # first recension paratext
dictionary_2r = {} # second recension paratext
dictionary_Fr = {'D.54 d.p.c.23': 'Ecce, quomodo serui ad clericatum ualeant  
assumi, uel quomodo non admittantur. Liberti quoque non sunt promouendi ad
```



```

clerum, nisi ab obsequiis sui patroni fuerint absoluti. Unde in Concilio
Eliberitano:'}

keysandpatterns = [{key': 'D.54 d.p.c.23', 'pattern': '(Ecce, quomodo
serui.*?quomodo non admittantur\.)'}]

for i in range (len(keysandpatterns)):

    key = keysandpatterns[i][key']

    pattern = keysandpatterns[i][pattern']

    result = re.search(pattern, dictionary_Fr[key'])

    dictionary_1r[key] = result.group(1)

    dictionary_2r[key] = re.sub(pattern, '', dictionary_Fr[key])

```

Here is the resulting first recension paratext:

```
{'D.54 d.p.c.23': 'Ecce, quomodo serui ad clericatum ualeant assumi, uel
quomodo non admittantur.'}
```

and the corresponding second recension paratext:

```
{'D.54 d.p.c.23': 'Liberti quoque non sunt promouendi ad clerum, nisi ab
obsequiis sui patroni fuerint absoluti. Unde in Concilio Eliberitano:'}
```



## Simplified two-dimensional visualization

Visualizing data from the *dicta* in a simplified two-dimensional form is a useful first step toward understanding how stylometric analysis works in practice. For the purpose of the following discussion, the *dicta* will be divided into four samples:

- the hypothetical case statements (*dicta initiales* or *themata*) that introduce the thirty-six cases in Part II of the *Decretum*,
- the first-recension *dicta* from Parts I and II of the *Decretum*,
- the first- and second-recension *dicta* from *de Penitentia*, and
- the second-recension *dicta* from Parts I and II of the *Decretum*.

In the following code, tables, and plots, these four samples will be labelled Gratian0, Gratian1, dePen, and Gratian2. The Gratian1 and Gratian2 samples, representing the first- and second-recension *dicta* from Parts I and II of the *Decretum*, were generated according to the procedure detailed in the preceding section on corpus preparation.

Although it is theoretically possible to split the text of the hypothetical case statements (*themata*) and the *dicta* from *de Penitentia* into separate first- and second-recension



samples following the same procedure used to produce the Gratian1 and Gratian2 samples, as a practical matter it is not useful to do so. The only case statement (*thema*) for which Winroth's appendix notes a textual difference is C.19 d.init.<sup>23</sup> The first-recension version of the text of C.19 d.init. omits a 13-word clause added to the second recension version, seemingly for the purpose of piling up descriptive detail. (*unus relicta propria ecclesia eo inuitio, alter dimissa regulari canonica cenobio se contulit*). Gratian0, the sample containing the cases statements, is therefore made up of 99.6% first-recension text. Similarly, the number of words added to the *dicta* in *de Penitentia* between the first and second recensions is relatively minimal, 556 words out of a total of 10,081 words; dePen, the sample containing the *dicta* from *de Penitentia*, is therefore made up of 94.5% first-recension text.

In addition to plotting a two-dimensional visualization of word frequency data from the *dicta*, this section will set the stage for a subsequent one, which introduces an authorship attribution technique known as Burrows's Delta. Burrows's algorithm

---

<sup>23</sup> Winroth, *The Making of Gratian's Decretum*, 216.



calculates a metric for the distance between a sample of unknown authorship with a corpus of samples of known authorship. In that discussion, Gratian0, the sample containing the hypothetical case statements (*themata*), will be treated as the sample of unknown authorship. Gratian1, dePen, and Gratian2, the samples containing the first-recension *dicta*, the first- and second-recension *dicta* from *de Penitentia*, and the second-recension *dicta*, will be treated as the corpus of samples of known authorship. Therefore, the values for means and standard deviations that provide the basis of comparison between the unattributed sample and the attributed corpus have to be calculated without taking the values from Gratian0 into account.

The demonstration of Burrows's Delta will make a point of the fact that the technique can be used at a higher number of dimensions ( $n > 3$ ) than can be visualized in graphical form. Word frequency data for the four most frequent words (MFWs) will therefore be collected from the start, even though the data for the third- and fourth-most frequent words will not be used in this section.

First, identify the four most frequent words in the comparison text samples, Gratian1, dePen, and Gratian2:



```

import re

def get_tokens(filename):
    '''open text file and return list of tokens'''

    # text = open(filename, 'r').read().lower()

    f = open(filename, 'r') # open file
    text = f.read() # read file
    text = text.lower() # lower-case text

    tokens = [word for word in re.split('\W', text) if word != ''] # remove
    punctuation

    return tokens

def get_features(samples):
    tokens = []
    for sample in samples:
        tokens += get_tokens('../corpus/' + sample + '.txt')
    types = list(set(tokens)) # create unordered list of unique words
    tmp = dict.fromkeys(types, 0) # create temporary dictionary, initialize
    counts to 0
    for token in tokens: tmp[token] += 1 # count words
    # re-order words in temporary dictionary numerically by descending

```



```

frequency

# re-order words with same frequency alphabetically

features = {

    key: value for key, value in sorted(tmp.items(),
        key = lambda item: (-item[1], item[0]))}

}

return features

samples = ['Gratian1', 'dePen', 'Gratian2']

features = get_features(samples)

list(features.keys())[:4] # 4 most frequent words (MFWs)

```

The four most frequent words in the three comparison samples Gratian1, dePen, and Gratian2—the samples treated as being of known authorship—are *in*, *non*, *et*, and *est*. The selection of samples makes a difference to the order. Were Gratian0, the sample treated as being of unknown authorship, to be included, the four most frequent words would be *in*, *et*, *non*, and *est*. (The rank reversal between the second- and third-most frequent words is a result of the fact that *non* occurs quite infrequently in Gratian0; see the table below.) After identifying the four most frequent words in the three



comparison samples, next, count the numbers of occurrences of those words in each of the samples:<sup>24</sup>

```
import pandas as pd
```

```
def get_counts(features, samples):
    columns = []
    for sample in samples:
        columns[sample] = []
        tmp = get_features([sample])
        for feature in features:
            columns[sample].append(tmp.get(feature, 0))
    return pd.DataFrame(columns, index = features)
```

---

```
mfws = list(features.keys())[4:] # 4 most frequent words (MFWs)
```

<sup>24</sup> Much of the analysis from this point forward will take advantage of the specialized capabilities of a Python software library called pandas. The name pandas is not a reference to the charismatic animal, but an acronym derived from the term “panel data.” The package is widely used in the field of data science, and provides a dataframe abstraction that represents two-dimensional numerical word-frequency data in a much more natural way than native Python data structure like lists and dictionaries do. The pandas dataframe abstraction can be thought of as a close analog to the Excel spreadsheets that were such a ubiquitous feature of John Burrows’s and David Hoover’s early experiments in stylometry.



```
unknown = 'Gratian0'

counts = get_counts(mfw, [unknown] + samples)
```

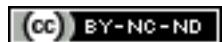
Gratian0 Gratian1 dePen Gratian2

	in	74	1450	252	411
non		24	1360	270	306
et		70	1293	260	345
est		13	965	182	167

After determining the number of occurrences of the MFWs, next, determine the length (total word count) for each of the samples:

```
def get_lengths(samples):
    filenames = ['../corpus/' + sample + '.txt' for sample in samples]
    lengths = {}
    for i in range(len(samples)):
        lengths[samples[i]] = len(get_tokens(filenames[i]))
    return pd.DataFrame(lengths, index = ['words'])

lengths = get_lengths([unknown] + samples)
```



Gratian0	Gratian1	dePen	Gratian2
----------	----------	-------	----------

words	3605	56713	10081
-------	------	-------	-------

Finally, divide the number of occurrences of the MFWs in the samples by the sample length and multiply the quotient by 1,000 to determine the normalized frequency of occurrence per 1,000 words for each of the MFWs in each of the samples:

```
frequencies = (counts / lengths.values) * 1000
```

Gratian0	Gratian1	dePen	Gratian2
----------	----------	-------	----------

in	20.5270	25.5673	24.9975
----	---------	---------	---------

non	6.6574	23.9804	26.7831
-----	--------	---------	---------

et	19.4175	22.7990	25.7911
----	---------	---------	---------

est	3.6061	17.0155	18.0538
-----	--------	---------	---------

*In* is the most frequently occurring word in the *dicta*. There are 1,450 occurrences of *in* out of 56,713 words in the first-recension *dicta* (25.5673 occurrences per 1,000 words), 252 occurrences of *in* out of 10,081 words in the *dicta* from *de Penitentia* (24.9975 per 1,000), and 411 occurrences of *in* out of 14,255 words in the second-recension *dicta* (28.8320 per 1,000). It is more convenient to characterize word frequencies in units of



occurrences per 1,000 words than percentage, since at that scale most of the values we are concerned with are greater than 1.0. The overall mean frequency of occurrence of *in* for the combined *dicta* from the first recension, *de Penitentia*, and the second recension is therefore 2,113 occurrences out of 81,049 words or 26.0706 per 1,000. It is common, however, for authorship attribution algorithms to use the mean of the mean frequencies of occurrence for each of the samples rather than the overall mean frequency of occurrence in the corpus made up of all of the samples. The motivation for using the mean of means for the individual samples rather than the overall mean frequency of occurrence is to ensure that largest sample does not dominate the result. The value for the mean frequency of occurrence of *in* that will be required at subsequent stages of this demonstration, then, is the mean of 25.5673, 24.9975, and 28.8320, or 26.4656 occurrences per 1,000.

*Non* is the second most frequently occurring word in the *dicta*. There are 1,360 occurrences of *non* out of 56,713 words in the first-recension *dicta* (23.9804 occurrences per 1,000 words), 270 occurrences of *non* out of 10,081 words in the *dicta* from *de Penitentia* (26.7831 per 1,000), and 306 occurrences of *non* out of 14,255 words in the second-recension *dicta* (21.4662 per 1,000). The overall mean frequency of occurrence of



*non* for the combined *dicta* from the first recension, *de Penitentia*, and the second recension is therefore 1,936 occurrences out of 81,049 words or 23.8868 per 1,000. The mean of the mean frequencies of occurrence of *non* for each of the samples is the mean of 23.9804, 26.7831, and 21.4662, or 24.0765 occurrences per 1,000.

Comparing the frequencies of occurrence of *in* and *non* in the two outlying samples, dePen and Gratian2, reveals unexpectedly large variations for such common words. (Large variations would be less surprising with uncommon words for which small differences in number could result in a large difference in percentage.) *In* occurs 15.3% more frequently in Gratian2 than in dePen, and 13.3% less frequently in dePen than in Gratian2. *Non* occurs 24.8% more frequently in dePen than in Gratian2, and 19.9% less frequently in Gratian2 than in dePen. Even compared to the mean, *non*, for example, occurs 11.2% more frequently in dePen and 10.8% less frequently in Gratian2 than the mean of means. It is clear then that against an overall background of “orderliness” (Zipf) in the word-frequency distribution, individual samples can display striking and potentially significant levels of variation.

Word count and sample length data were collected and used to calculate frequencies for Gratian0 above, but those values will not be used in this section. Disregard the Gratian0 column, and use only the columns corresponding to the three comparison samples, Gratian1, dePen, and Gratian2, to calculate the means for the values in each of the rows in the frequency table representing the four most frequent words:

```
means = frequencies[samples].mean(axis = 1).to_frame('mean')
```

	Gratian1	dePen	Gratian2	mean
in	25.5673	24.9975	28.8320	26.4656
non	23.9804	26.7831	21.4662	24.0765
et	22.7990	25.7911	24.2020	24.2640
est	17.0155	18.0538	11.7152	15.5948

We can graph the number of occurrences of *in* and *non* per 1,000 words in the *dicta*, with the frequency of *in* plotted along the horizontal x-axis, and the frequency of *non* plotted along the vertical y-axis, to produce a simplified visualization of the total variation between the three samples. Means are provided for context: the vertical dashed line



represents the mean of means for the horizontal (*in*) axis, and the horizontal dashed line represents the mean of means for the vertical (*non*) axis.

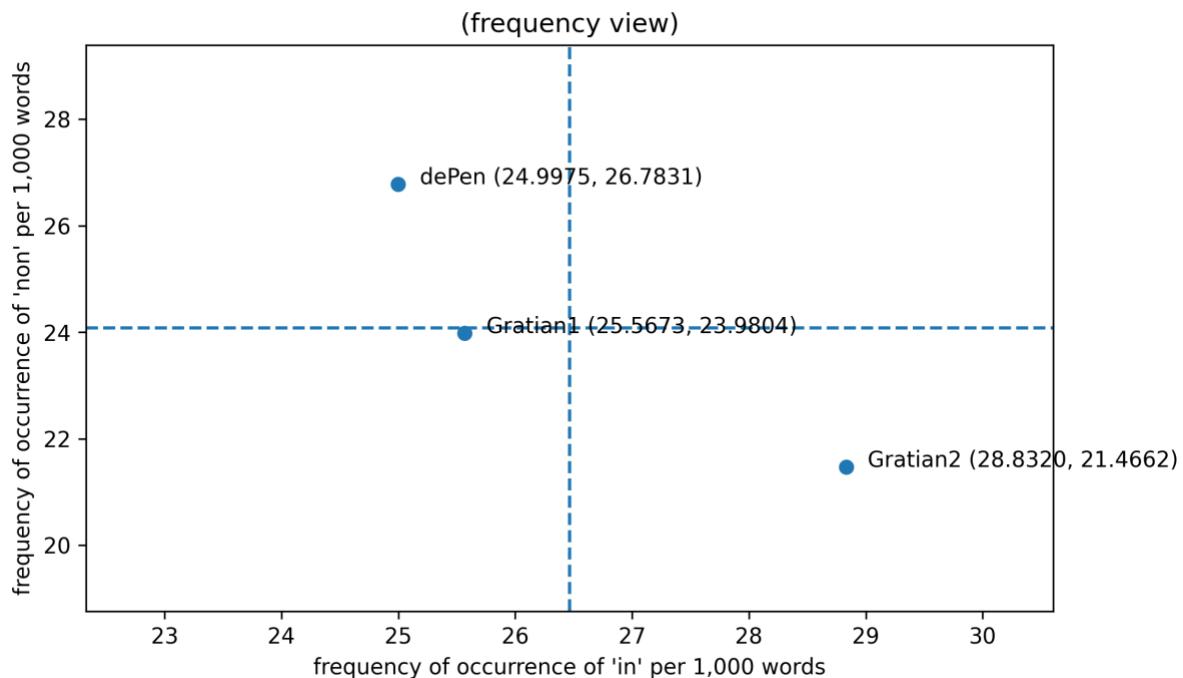


Figure 0a updated 14 May 2020<sup>25</sup>

---

<sup>25</sup> The actual generation of Figure 0a was deferred until after the sample standard deviations for *in* and *non* per 1,000 words had been calculated below. Framing the dimensions of the plot to twice the standard deviation from the mean along both axes improves graphical layout and readability.

Figure 0a introduces several conventions common to two-dimensional graphical representations of word frequency data that readers will encounter repeatedly throughout this chapter. The feature (in this case the frequency of occurrence of the word *in*) that explains more of the variation between the samples is plotted along the horizontal x-axis, while the feature (in this case the frequency of occurrence of the word *non*) that explains less of the variation between the samples is plotted along the vertical y-axis. Although this plot, produced by the Matplotlib Python two-dimensional plotting library, is rectangular and the axes are approximately to scale, many of the figures in this chapter were generated using stylo, an R package for stylometric analysis, which outputs square plots. Regardless of appearance, however, readers should bear in mind that the area plotted is always wider than it is tall, that is, that it displays greater variation between samples horizontally along the x-axis than it does vertically along the y-axis.

Figure 0a plots the values for the first-recension *dicta*, the second-recension *dicta*, the *dicta* from *de Penitentia*, the second-recension *dicta*, and (labelled Gratian1, dePen, and Gratian2 respectively), as well as the means (indicated by the dashed lines), for the frequencies of *in* and *non* per 1,000 words. It is more statistically meaningful, however,



to measure and plot the differences between values and means in units of standard deviations rather than frequency per 1,000 words. The difference of a value from the mean divided by standard deviation is referred to as the value's z-score. A value that has a difference of one standard deviation from the mean is said to have a z-score of 1.0 or -1.0 depending on whether the value is greater or lesser than the mean. The formula used to calculate the sample standard deviation is:

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$$

The formula is slightly daunting notationally, but it is not difficult to use it to calculate the desired results. The example immediately below shows all of the intermediate steps involved in using the formula to compute the sample standard deviation of the frequency of occurrence of the word *in* in the Gratian1, dePen, and Gratian2 samples. (The motivated reader can use a calculator to repeat the process for the frequency of occurrence of the word *non*.) Remember that for the purpose of calculating sample standard deviation, the value of the mean ( $\bar{x}$ ) is **not** the overall mean frequency of

occurrence of the word *in* across all of the samples, but the mean of the frequencies for each of the samples individually (the mean of means).

First, calculate the squared deviations from the mean of means for the frequency of *in* in the first-recension *dicta* (Gratian1):

$$(x_1 - \bar{x})^2 = (25.5673 - 26.4656)^2 = (-0.8983)^2 = 0.8069,$$

for the frequency of *in* in the *dicta* from *de Penitentia* (dePen):

$$(x_2 - \bar{x})^2 = (24.9975 - 26.4656)^2 = (-1.4681)^2 = 2.1553,$$

and for the frequency of *in* in the second-recension *dicta* (Gratian2):

$$(x_3 - \bar{x})^2 = (28.8320 - 26.4656)^2 = (2.3664)^2 = 5.5998.$$

Then, as indicated by the summation operator  $\Sigma$ , sum the three squared deviations from the mean of means, divide the sum by their number ( $N = 3$ ) minus one, and take the square root of the quotient:

$$s = \sqrt{\frac{1}{2}(0.8069 + 2.1553 + 5.5998)} = \sqrt{\frac{1}{2}(8.5620)} = \sqrt{4.2810} = 2.0691$$

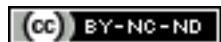
The units of s are the same as those used to calculate the mean, in this case, the frequency of occurrence of a word per 1,000 words.

Once again disregarding the Gratian0 column of the frequencies table, calculate the standard deviations for the rest of the rows representing the four most frequent words, using only the values in the columns corresponding to the three comparison samples, and the means computed from them:

```
standard_deviations = frequencies[samples].std(axis = 1).to_frame('std')
```

	Gratian1	dePen	Gratian2	mean	std
in	25.5673	24.9975	28.8320	26.4656	2.0691
non	23.9804	26.7831	21.4662	24.0765	2.6598
et	22.7990	25.7911	24.2020	24.2640	1.4970
est	17.0155	18.0538	11.7152	15.5948	3.3997

As noted above, the definition of a value's z-score is the difference of that value from the mean divided by the standard deviation. A z-score can be calculated for a value even if that value was not used to determine the mean and standard deviation to be used. That means that z-scores can be calculated for word frequencies in the



unattributed sample Gratian0 using the means and standard deviations calculated using the corresponding word frequencies in the attributed samples Gratian1, dePen, and Gratian2. Just as word frequencies were calculated for Gratian0 above, z-scores will be calculated for Gratian0 here, which will be used in the next section to determine the value of Burrows's Delta. The formula used to calculate the z-score is:

$$z = \frac{x - \bar{x}}{s}$$

For the frequency of *in* in the case statements or *themata* (Gratian0):

$$z = \frac{x - \bar{x}}{s} = \frac{20.5270 - 26.4656}{2.0691} = \frac{-5.9386}{2.0691} = -2.8702,$$

for the frequency of *in* in the first-recension *dicta* (Gratian1):

$$z = \frac{x - \bar{x}}{s} = \frac{25.5673 - 26.4656}{2.0691} = \frac{-0.8983}{2.0691} = -0.4342,$$

for the frequency of *in* in the *dicta* from *de Penitentia* (dePen):

$$z = \frac{x - \bar{x}}{s} = \frac{24.9975 - 26.4656}{2.0691} = \frac{-1.4681}{2.0691} = -0.7095,$$

and for the frequency of *in* in the second-recension *dicta* (Gratian2):



$$z = \frac{x - \bar{x}}{s} = \frac{28.8320 - 26.4656}{2.0691} = \frac{2.3664}{2.0691} = 1.1437.$$

(Because both the numerator and the denominator of the formula for calculating z-scores have units of frequency of occurrence per 1,000 words, z is a dimensionless number.)

Calculate the z-scores for the remaining most frequent words, and then plot the coordinates of the attributed samples Gratian1, dePen, and Gratian2:

```
z_scores = (frequencies - means.values) / standard_deviations.values
```

	Gratian0	Gratian1	dePen	Gratian2
in	-2.8702	-0.4342	-0.7095	1.1437
non	-6.5491	-0.0361	1.0176	-0.9814
et	-3.2375	-0.9786	1.0201	-0.0414
est	-3.5264	0.4179	0.7233	-1.1412

in	-2.8702	-0.4342	-0.7095	1.1437
non	-6.5491	-0.0361	1.0176	-0.9814
et	-3.2375	-0.9786	1.0201	-0.0414
est	-3.5264	0.4179	0.7233	-1.1412

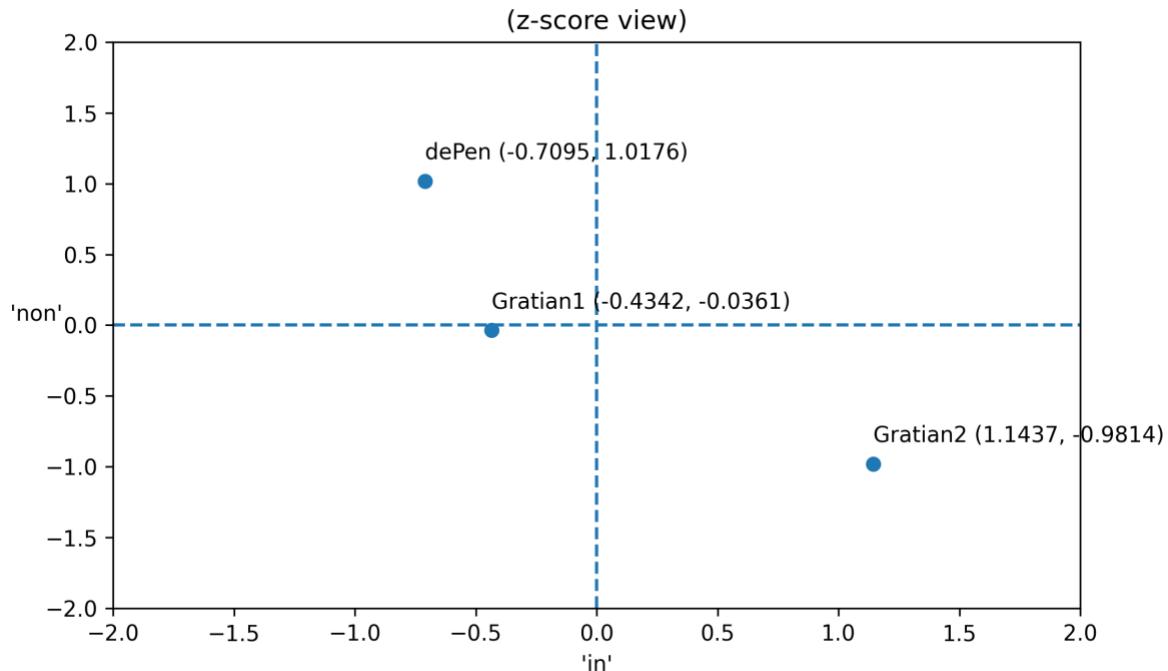


Figure 0b updated 15 May 2020

Labels on the axes of Figure 0b refer to standard deviations (values of z) away from the mean of means (represented by the dashed lines).

Figures 0a and 0b represents the axes as orthogonal (perpendicular) to one another.

Although doing so is acceptable as a first-order approximation in a simplified representation of this kind, plotting the values along orthogonal axes invokes an implicit assumption that the word frequencies (in this case, of *in* and *non*) are

completely independent of one another, i.e., that there is no correlation or covariance relationship between the words' frequency of occurrence in the samples. This is not necessarily the case, and an advanced technique introduced below, principal component analysis (PCA), handles this problem in a more mathematically sophisticated way.

Now, we are obviously not going to make an attribution of authorship based on the frequencies of only two function words. Increasing the number of function words for which one collects frequency data increases the accuracy of stylometric analysis, up to a point. There is, however, a limit to the marginal value of each additional word included in an analysis, for two reasons. The first reason is that the frequency of each word in a corpus of text tails off inversely as the word's rank, a relationship (approximately) described by the equation  $r \times f = C$ . (The rank of the most frequent word is defined as 1, that of the second most frequent word as 2, and so on.) As a consequence, assuming words are weighted in proportion to their frequency of occurrence in the corpus, every marginal word has less value as evidence than the word before it. Some stylometric techniques, such as Burrows's Delta, weight each word from a corpus of texts equally, but such choices need to be made with an informed awareness of the trade-offs



involved. The second reason is that every marginal word is less likely to be a function word, and more likely to be a content word. Twenty-four out of the thirty most frequent words from Gratian's *dicta* are function words potentially suitable for use in stylometric analysis, but only 64 out of the 250 most frequent words.

## Zipf's law

The observed empirical relationship that word frequencies in a corpus of text tail off inversely as the rank is known as Zipf's law, after American quantitative linguist George Kingsley Zipf (d.1950). Though he does not seem to have claimed discovery, Zipf published the first extended discussions of the phenomenon.<sup>26</sup> **The rank-frequency distribution of words in a corpus of any language is not merely a curiosity, but rather provides a general background of what Zipf characterized as “orderliness” against which variations in frequency of occurrence of individual words are both detectable**

---

<sup>26</sup> George Kingsley Zipf, *The Psycho-Biology of Language: An Introduction to Dynamic Philology* (Boston: Houghton Mifflin Company, 1935), 39-48; and George Kingsley Zipf, *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology* (Cambridge, Mass: Addison-Wesley Press, 1949), 73-131. Zipf referred to the relationship as “the law of diminishing returns of words” (1949). He expressed the relationship as  $ab^2 = k$  (1935) and  $r \times f = C$  (1949).



and meaningful. Thus it is worthwhile to discuss theoretical predictions about the rank-frequency distribution of words in a corpus, and to demonstrate that Gratian's *dicta* display the expected orderliness of word distribution.

If the words in a corpus of text are rank-ordered from most to least frequent, Zipf's law posits that, as a first-order approximation, the frequency of the Nth word will be  $1/N$  times that of the most frequent word.<sup>27</sup> In other words, the theoretical Zipf distribution predicts that the frequency of the second most frequent word in a corpus of text should be one half that of the most frequent word, the frequency of the third most frequent word should be one third that of the most frequent word, and so on. (See Figure Za below.)

---

<sup>27</sup> In Digital Humanities courses,  $1/N$  is typically presented as Zipf's law. The reductionist  $1/N$  representation of the rank-frequency relationship is misleading insofar as it ignores scaling considerations, and elides the discrete rather than continuous nature of the variables representing rank and frequency.



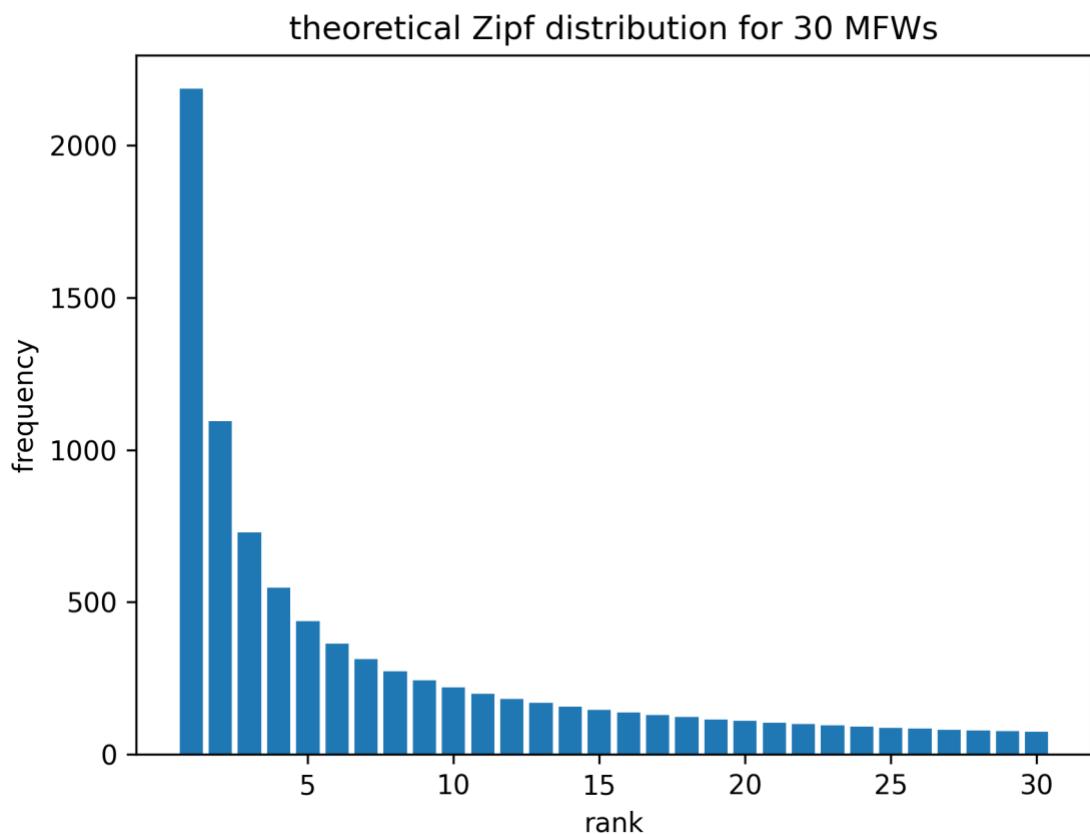


Figure Za updated 27 May 2020

Zipf tried several different approaches to the quantitative analysis of the distribution of words in corpora. His initial attempt, in 1935, sought to characterize the orderliness of word distributions by the relationship between the number of occurrences of a word,  $b$ , and the number of words  $a$ , a relationship Zipf expressed as  $ab^2 = k$ . For example, in

Gratian's *dicta* there are 8,028 words (*a*) for which there is 1 (*b*) occurrence, 2,462 words (*a*) for which there are 2 (*b*) occurrences, 1,135 words (*a*) for which there are 3 (*b*) occurrences, and 660 words (*a*) for which there are 4 (*b*) occurrences. Plugging these values for *a* and *b* into the formula  $ab^2 = k$  yields 8,028, 9,848, 10,215, and 10,560 respectively as values for the "constant" *k*. As the number of occurrences *b* increases, values of *k* for Gratian's *dicta* remain fairly stable, mostly falling between 10,000 and 11,000.

b	a	$b^2$	k
1	8028	1	8028
2	2462	4	9848
3	1135	9	10215
4	660	16	10560
5	423	25	10575
6	290	36	10440
7	225	49	11025
8	173	64	11072

As this example suggests, the explanatory power of the formula  $ab^2 = k$  to characterize a word distribution is greatest for words that have a low number of occurrences. Zipf ultimately judged the formula  $ab^2 = k$  to be a less than satisfactory model for the full spectrum of word distribution in a corpus, in part because it implied fractional values of  $a$  for the most frequent words.<sup>28</sup> Zipf had a vivid awareness, ahead of its time for the pre-digital age in which he lived, of the fact that the variables in the formulas by which he sought to express these relationships represent discrete rather than continuous quantities.<sup>29</sup> **Zipf also noted that value of the exponent of b is likely to differ from 2 depending on the size of the corpus<sup>30</sup>**

---

<sup>28</sup> "Hence the  $ab^2 = k$  relationship is valid only for the less frequently occurring words which, however, represent the greater part of the vocabulary in use, though not always a great majority of the occurrences." ... "It is perhaps worth pointing out that the  $ab^2 = k$  relationship which appears valid for the frequency distribution of the less frequent words would demand fractional words when applied to the speech-elements of highest occurrence, such as *the* in English." Zipf, *The Psycho-Biology of Language*, 42-43.

<sup>29</sup> Zipf used the term "integrality" to describe the discrete, discontinuous, nature of frequency and rank. Zipf, *Human Behavior and the Principle of Least Effort*, 31, 35.

<sup>30</sup> Zipf, *The Psycho-Biology of Language*, 43.



Zipf's final attempt, in 1949, to give a quantitative account of the distribution of words in a corpus characterized the distribution in terms of rank and frequency as  $r \times f = C$ . Values for the constant  $C$  differ between corpora, depending, among other things, on corpus size. As an example, the theoretical Zipf distribution plotted in Figures Za and Zb has been scaled to facilitate direct comparison with actual data from Gratian's *dicta* plotted in Figures Zc and Zd. In all four plots, the first data point has a rank of 1 and a frequency of 2187, corresponding to the 2,187 occurrences of the most frequent word *in* in the *dicta*. This scaling is equivalent to setting the value of  $C$  to 2187, and letting  $r \times f = 2187$ .

Both the  $ab^2 = k$  and  $r \times f = C$  interpretations of Zipf's law can be restated with greater mathematical generality by noting that if the variables are plotted logarithmically, their relationships are linear, with the slope determined by the power (exponent) relationships involved. In the equation  $ab^2 = k$ , the number of occurrences  $b$  varies inversely as the square root of the number of words  $a$  (as  $\frac{1}{\sqrt{a}}$ , or equivalently as  $a^{-1/2}$ ).

The relationship of the logarithm of the number of occurrences to the number of words is linear, with a slope of -1/2 corresponding to the exponent (-1/2) of the number of words. (See Figure Zy in the appendix at the end of this section for actual data from



Gratian's *dicta* for values of  $b$  [Number of Occurrences] from 1 to 30. The actual slope is -0.5097, quite close to the predicted theoretical value of -0.5.)

Similarly, in the equation  $r \times f = C$ , the frequency  $f$  varies inversely as the rank  $r$  (as  $1/r$ , or equivalently as  $r^{-1}$ ). The relationship of the logarithm of frequency to the logarithm of rank is linear, with a slope of -1 corresponding to the exponent (-1) of the rank. (See Figure Zb below.) Note that the base of the logarithms does not matter, as long as the bases are the same for both axes. Regardless of whether we take base  $e$  (natural) or base 10 logarithms of rank and frequency, for example, the slopes will be the same: -1.0 for the theoretical Zipf distribution of word frequencies in Figure Zb, and -0.6518 for the actual frequencies of the thirty most frequent words in Gratian's *dicta* in Figure Zd.

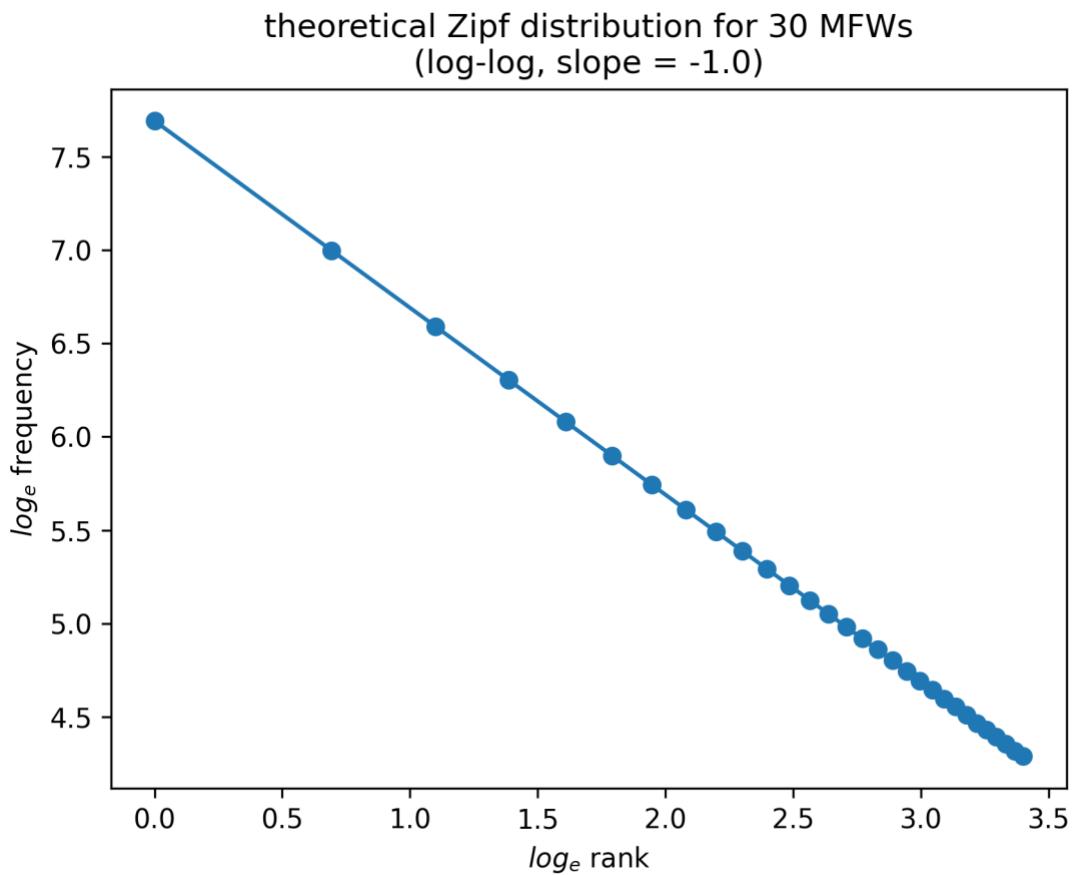


Figure Zb updated 27 May 2020

Figure Zc below plots the actual rank-frequency distribution of the thirty most frequent words (MFWs) in Gratian's *dicta*: *in* (2187), *et* (1968), *non* (1960), *est* (1327), *de* (925), *quod* (888), *ad* (832), *qui* (812), *sed* (736), *unde* (732), *uel* (705), *si* (669), *ut* (641), *cum* (589), *a*

(588), *autem* (582), *ex* (501), *sunt* (428), *enim* (424), *que* (423), *uero* (411), *etiam* (405), *ab* (391), *ait* (349), *esse* (339), *ergo* (338), *quia* (336), *item* (327), *per* (304), *nec* (293).

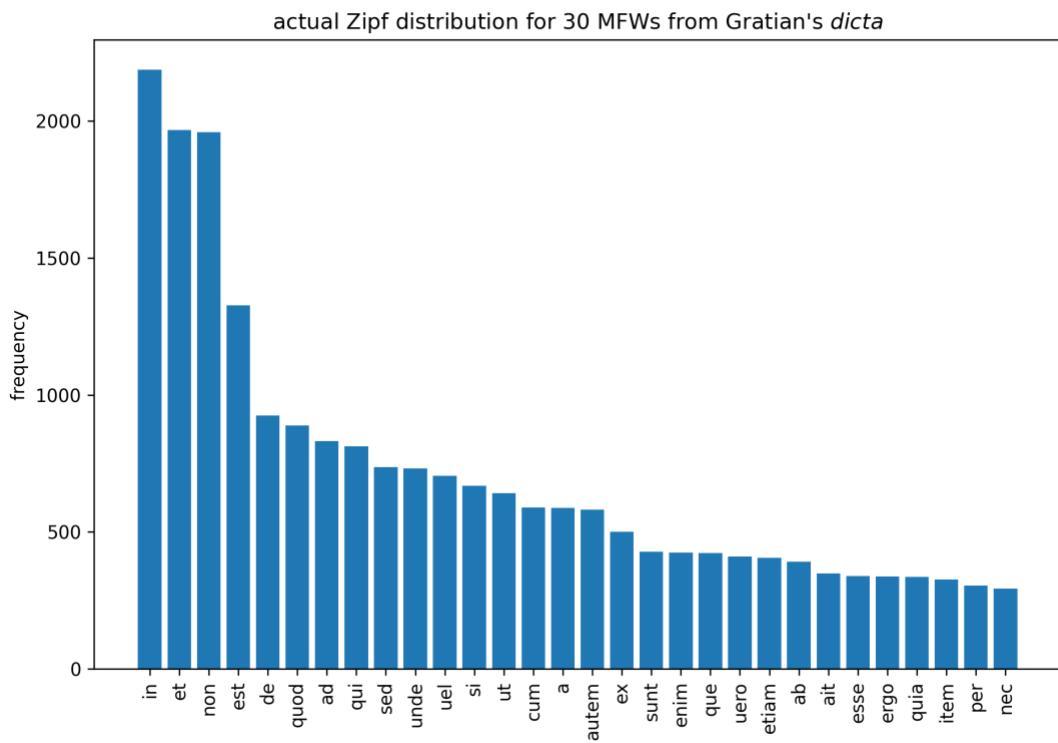


Figure Zc updated 27 May 2020

Zipf used word frequencies hand-tabulated from James Joyce's *Ulysses* as the data set for his exploration of the rank-frequency relationship, and it turns out that for English, the  $r \times f = C$  formulation holds up reasonably well.<sup>31</sup> The rank-frequency relationship does not on first inspection appear to hold up as well for Gratian's Latin as it does for Joyce's English, since the frequencies for the thirty most frequent words of the *dicta* do not drop off quite as sharply as the  $r \times f = C$  formulation of Zipf's law would predict. The frequency of *et*, the second most frequent word in Gratian's *dicta* is 0.8999 times that of *in*, the most frequent word, rather than 0.5 as Zipf's law would predict; and the frequency of *non*, the third most frequent word, is 0.8962 rather than 0.3333.

Plotting the data from Figure Zc on logarithmic axes and performing least-squares linear regression analysis lets us calculate the slope, -0.6518, for the rank-frequency tail-off of the thirty most frequent words from Gratian's *dicta*.<sup>32</sup> (See Figure Zd below.)

---

<sup>31</sup> "we have found a clearcut correlation between the number of different words in the *Ulysses* and the frequency of their usage, in the sense that they approximate the simple equation of an equilateral hyperbola:  $r \times f = C$  in which  $r$  refers to the word's rank in the *Ulysses* and  $f$  to its frequency of occurrence (as we ignore for the present the size of  $C$ )." Zipf, *Human Behavior and the Principle of Least Effort*, 24. See Zipf, 23-52, for Zipf's extended discussion of the rank-frequency distribution of words in Joyce's *Ulysses*.

<sup>32</sup>  $m = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$  or:



Transposing that result back into the linear (as opposed to logarithmic) frame of reference used in Figure Zc, the expression  $1/r^{0.6518}$  yields a better (though not perfect) fit to the actual rank-frequency data.

---

```
def regression_slope(data_points):
    n = len(data_points)
    x_values, y_values = zip(*data_points)
    x_bar = statistics.mean(x_values)
    y_bar = statistics.mean(y_values)
    xy_sum = 0
    x_squared_sum = 0
    for i in range(n):
        xy_sum += x_values[i] * y_values[i]
        x_squared_sum += x_values[i] ** 2
    return (xy_sum - n * x_bar * y_bar) / (x_squared_sum - n * x_bar ** 2)
```



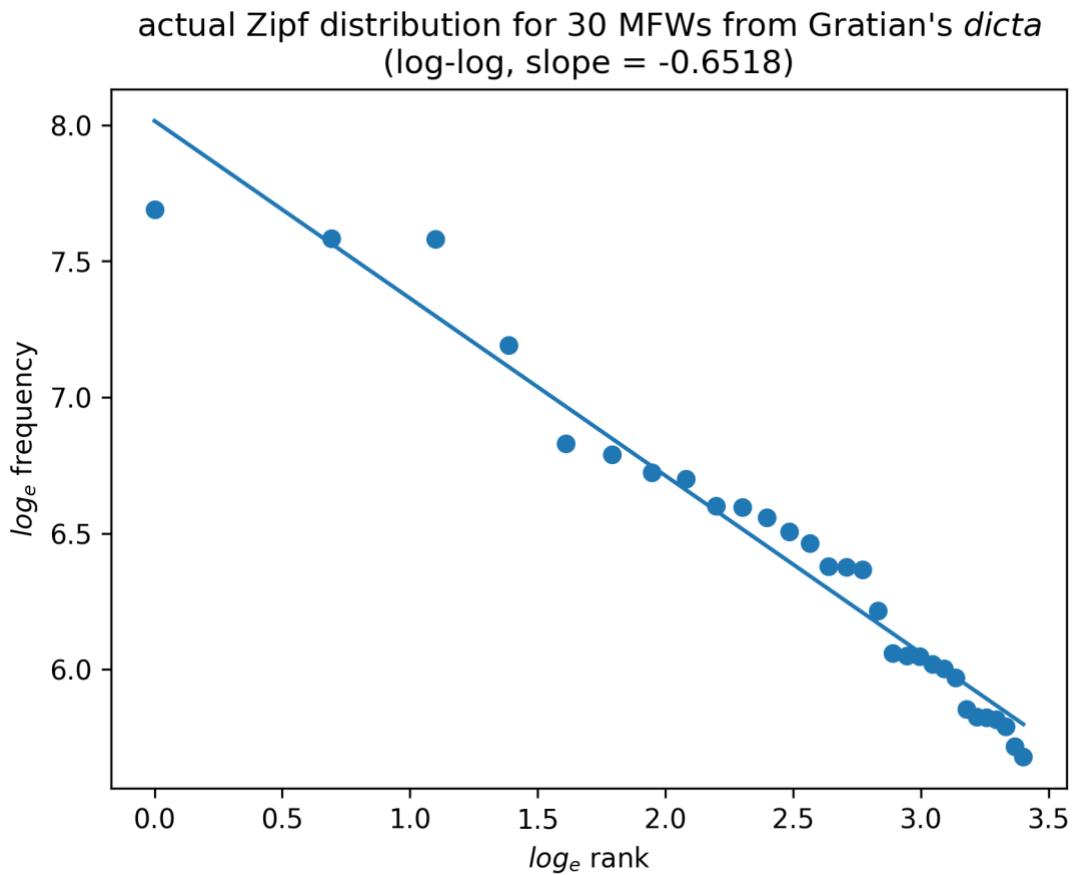


Figure Zd updated 27 May 2020

## Zipf's law (appendix)

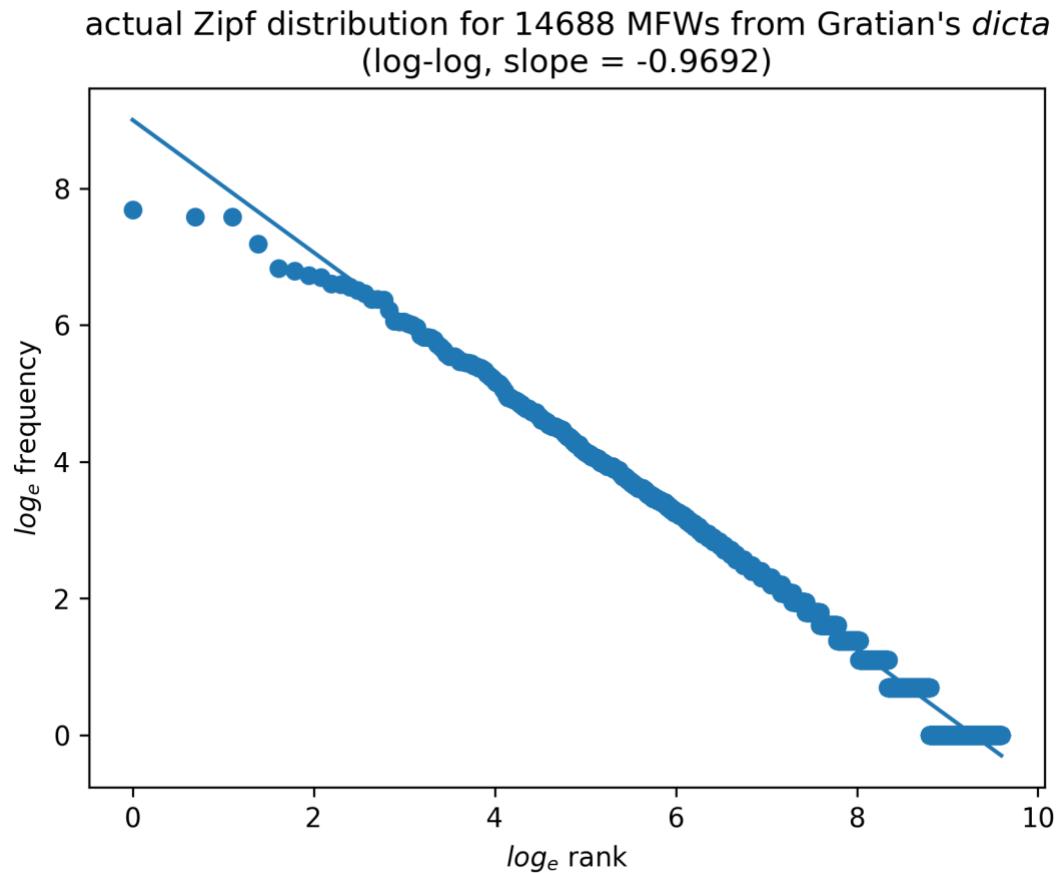


Figure Ze updated 27 May 2020

actual Zipf 1935 distribution ( $ab^2 = k$ ) for *dicta* from Gratian's *Decretum*  
number of occurrences between 1 and 30  
(log-log, slope = -0.5097)

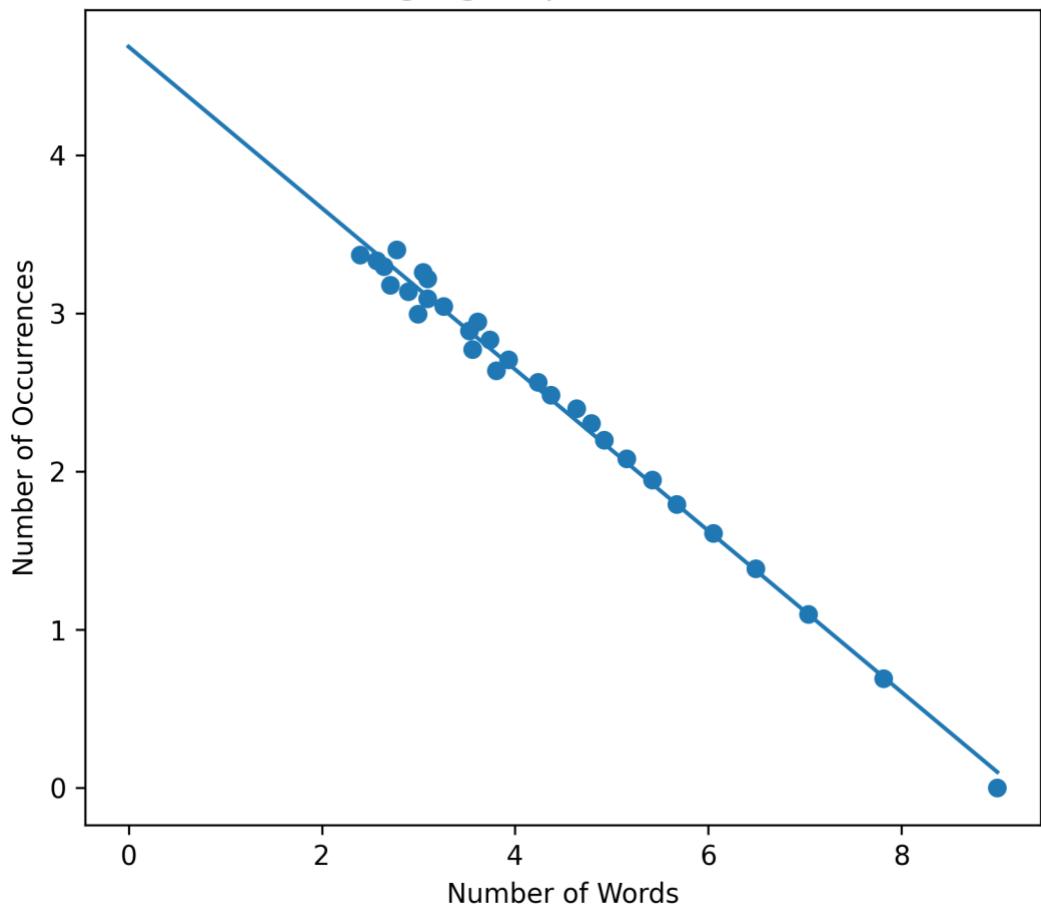


Figure Zy updated 27 May 2020

actual Zipf 1935 distribution ( $ab^2 = k$ ) for *dicta* from Gratian's *Decretum*  
 number of occurrences between 1 and 173

(log-log, slope = -0.6507)

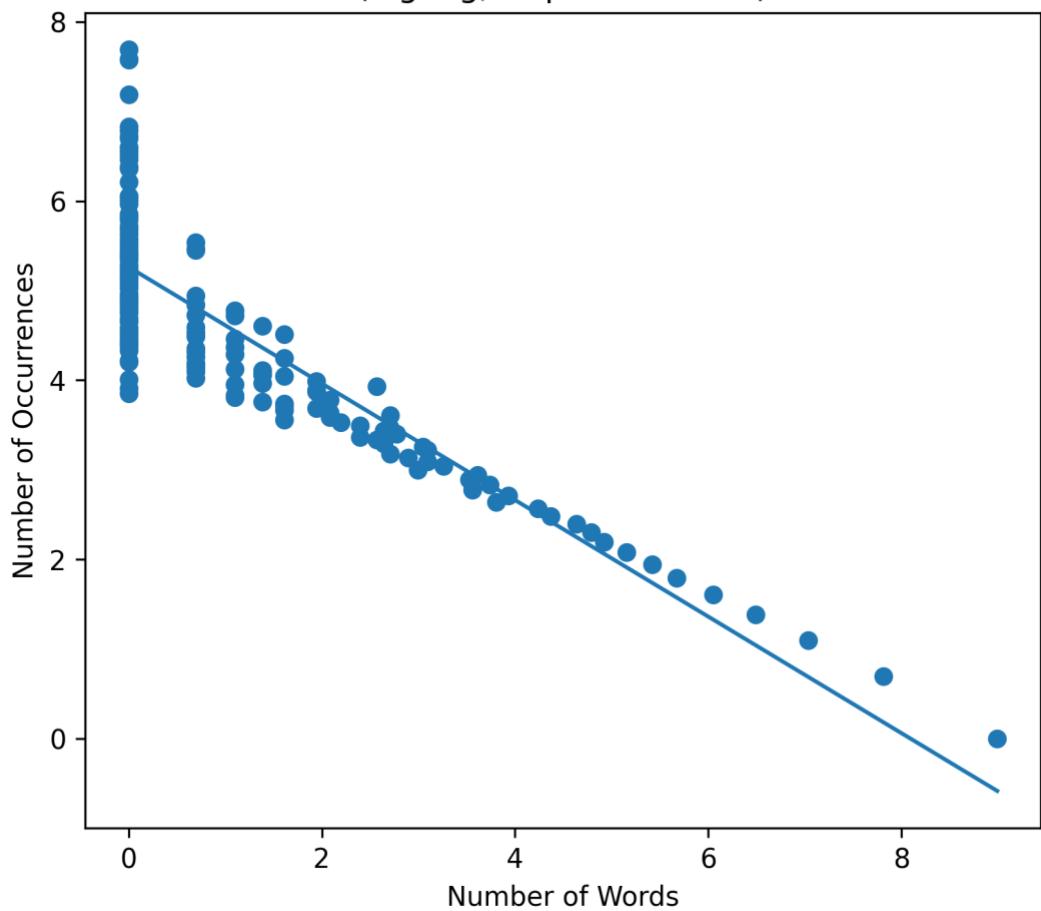


Figure Zz updated 27 May 2020<sup>33</sup>

---

<sup>33</sup> The vertical bands toward the upper-left hand corner of the plot—which also appear in Zipf's 1935 figures—require some interpretation. There can be a range of values for numbers of occurrence (b) for



## Burrows's Delta

The examples presented in the previous section are suggestive of ways in which differences between the frequencies of occurrence of common words in samples from a corpus of texts can be quantified in statistically meaningful units (standard deviations or values of z) and combined to represent the distance between those samples. This technique is, however, of limited value so long as we are restricted to the two, or at most three, dimensions the human mind is capable of visualizing. In 2001, John F. Burrows (d.2019) of the University of Newcastle, Australia, proposed a generalization that gets around the limitation on the number of features to two or three by averaging z-score distance measurements of word frequency data for any number of features. This has the effect of collapsing distance measurements in an arbitrary number of dimensions into a single metric. Burrows called this metric the Delta, and it is now

---

which only one value for number of words (a) has that number of occurrence. There are 62 cases in the *dicta* from Gratian's *Decretum* in which only a single word (a = 1) has a given number of occurrences (b). For example, only one word, *potest*, has 143 occurrences, and only one word, *in*, has 2,187 occurrences. The upper-left-most vertical band represents all 62 of the words for which only a single word (a = 1) has a particular value for the number of occurrences (b). Similarly, the next vertical band to the right represents all of the words for which only two words (a = 2) have particular values for the number of occurrences (b).



generally referred to as Burrows's Delta ( $\Delta_B$ ).<sup>34</sup> Expositions of Burrows's Delta sometime fail to make a clear enough distinction between the metric  $\Delta_B$  and the authorship attribution methodology in which Burrows applied it. The metric is not the methodology.

Attempts to attribute authorship are typically undertaken in scenarios where there is a large (enough) number of texts securely attributable to a known author, and a text, or at most a small number of texts, of unknown authorship. The attempt is then made to attribute the unknown text to one of the known authors, or to rule out such an attribution. Take the *Federalist* as an example. There are numbers of the *Federalist* of disputed or unknown attribution, a small and well-defined number of candidates for authorship—Hamilton, Jay, Madison—to whom those numbers might be attributed, and securely attributed samples from each of the candidates, conveniently enough from the same work.

---

<sup>34</sup> John Burrows, "Questions of Authorship: Attribution and Beyond: A Lecture Delivered on the Occasion of the Roberto Busa Award ACH-ALLC 2001, New York," *Computers and the Humanities* 37, no. 1 (February 2003): 5–32; and "'Delta': A Measure of Stylistic Difference and a Guide to Likely Authorship," *Literary and Linguistic Computing* 17, no. 3 (September 2002): 267–87.



Burrows's methodology assumes just such a scenario. He began by identifying the most frequent words (MFWs) in the corpus of comparison texts securely attributed to known authors. In Burrows's published descriptions of his methodology, he typically used all of the 30 most frequent words in the corpus of attributed comparison texts without distinguishing between function and content words. He then tabulated the number of occurrences of the most frequent words in each of the sample texts in the comparison corpus and normalized their frequency of occurrence as a percentage. Burrows then used the frequency data collected from the comparison texts to calculate a mean frequency of occurrence and sample standard deviation for each of the MFWs or features.

It is important to emphasize that the mean frequency of feature occurrence calculated at this stage of Burrows's algorithm and subsequently used to calculate the sample standard deviation for each feature is **not** the overall mean frequency across the corpus of attributed comparison text samples. Instead, the comparison corpus feature mean is calculated by averaging the normalized (percentage) frequency for each feature across all of the text samples in the attributed comparison corpus, without concern for differences in size (word count) between the samples. To refer back to the example



presented in the previous section as part of the two-dimensional visualization demonstration, we did **not** use the **overall** mean frequency of *in* across the three samples Gratian1, dePen, and Gratian2, (2,113 occurrences out of 81,049 words or 26.0706 per 1,000), but rather the mean of the normalized frequencies of *in* for each of the samples (the mean of 25.5673, 24.9975, and 28.8320, or 26.4656 occurrences per 1,000). (**mean of means**)

After calculating the mean of means and sample standard deviation for each of the features (MFWs), Burrows then converted the normalized (percentage) frequencies of occurrence for each feature in each sample in the comparison corpus to z-scores by subtracting the mean of means from the frequency and dividing the positive or negative difference by the standard deviation for the feature. At this point, Burrows turned his attention to the unattributed text, tabulating all occurrences of the 30 MFWs for which data had been collected from the comparison texts, then normalizing the word counts by converting them to percentage frequencies of occurrence. Burrows then converted the normalized frequencies for each feature in the unattributed test sample to z-scores based on the values for the mean of means and sample standard deviation derived from the feature frequencies in the attributed comparison corpus samples.



With these preliminaries out of the way, Burrows then calculated the value of the Delta by taking the average (arithmetic mean) of the absolute value of the differences between the z-score for a given feature (MFW) for the unattributed test sample and each of the comparison samples in the corpus of attributed texts. In Burrows's interpretation, the comparison test sample from the attributed corpus with the lowest Delta with respect to the unattributed test sample was most likely to share a common author with it.

It is not possible to apply Burrow's methodology in the case of the *dicta* from Gratian's *Decretum* without modification. As the survey in Chapter 3 above indicated, near-contemporaries knew next to nothing about Gratian. Perhaps most notably, although Gratian was thought to have been a teacher, no one in the generation following made an unambiguous claim to have been his student. There are no other writings securely, or even insecurely, attributed to him. Fortunately, Burrows's Delta can be readily adapted to the particular situation in which we find ourselves, where there are no other texts attributed to Gratian with which we can compare, for example, the hypothetical case statements (*themata*) or second-recension *dicta*.



Although other distance methods of authorship attribution have been proposed since,<sup>35</sup> Burrows's Delta is widely accepted in the scholarly literature of the field of computational linguistics, and it will therefore be used as the basis for the demonstrations in this section.

The first experiment will be a comparison of four subcorpora, Gratian0 (the hypothetical case statements or *themata*), Gratian1 (the first-recension *dicta* excluding the *dicta* from *de Penitentia*), dePen (first- and second-recension *dicta* from *de Penitentia*), and Gratian2 (the second-recension *dicta* excluding the *dicta* from *de Penitentia*), using the frequencies of occurrence of the four most frequent words (MFWs) in Gratian's *dicta* as the basis for comparison. We will hypothesize that the subcorpus containing the hypothetical case statements (*themata*) is the work of an unknown author, and will treat the other three subcorpora as making up a corpus of works by a known author. Using four subcorpora and four features, where every feature analyzed is represented in a different dimension, demonstrates that z-score distance methods can be extended to

---

<sup>35</sup> Stefan Evert et al., "Understanding and Explaining Delta Measures for Authorship Attribution," *Digital Scholarship in the Humanities* 32, no. suppl\_2 (December 2017): ii4–ii16.



cases in which the number of dimensions is greater than three. It also has the advantage of making the solution compact enough to allow readers to follow along and reassure themselves of the mathematical validity of all of the intermediate steps leading to the final result.

The first experiment resumes directly where the two-dimensional visualization demonstration left off, so all of the function definitions and variable values in force at the conclusion of that demonstration are still valid. In particular, this experiment inherits the z-scores for all of the four most frequent words (MFWs). While we disregarded the data for the third and fourth most frequent words (*et* and *est*) for the purpose of the visualization demonstration, they will be fully taken into account here. (Remember that the values for mean and standard deviations used to derive the z-scores were calculated without reference to the Gratian0 sample here being treated as the unknown).

First, split the z-scores into two new dataframes, one for the test sample Gratian0, assumed for the purpose of this experiment to be the work of an unknown author:

```
test = z_scores[[unknown]]
```



### Gratian0

in	-2.8702
----	---------

non	-6.5491
-----	---------

et	-3.2375
----	---------

est	-3.5264
-----	---------

the other for the comparison samples *Gratian1*, *dePen*, and *Gratian2*, assumed for the purpose of this experiment to represent the work of known authors:

```
corpus = z_scores[samples]
```

### Gratian1 dePen Gratian2

in	-0.4342	-0.7095	1.1437
----	---------	---------	--------

non	-0.0361	1.0176	-0.9814
-----	---------	--------	---------

et	-0.9786	1.0201	-0.0414
----	---------	--------	---------

est	0.4179	0.7233	-1.1412
-----	--------	--------	---------

The formula used to calculate Burrows's Delta is:

$$\Delta_B = \frac{1}{N} \sum_{i=1}^N |z_i(t) - z_i(c)|$$



It is easiest to deal with the formula in two steps, first evaluating the expression  $|z_i(t) - z_i(c)|$ . Note that because we take the absolute value of the result, the order of operands on either side of the subtraction operator '-' does not matter. For each of the three columns (Gratian1, dePen, and Gratian2) in the *corpus* dataframe, subtract the z-score in each row from the z-score in the same row of the *test* (Gratian0) dataframe, take the absolute value, and record the result in the corresponding column and row of the *differences* dataframe. For example, the z-score for *non* in *test* (Gratian0) is -6.5491, the z-score for *non* in the Gratian1 column of *corpus* is -0.0361, so the absolute value of the difference recorded in the *non* row of the Gratian1 column of *differences* would be 6.5130.

```
differences = (test.values - corpus).abs()
```

	Gratian1	dePen	Gratian2
in	2.436	2.1606	4.0139
non	6.513	7.5667	5.5677
et	2.2589	4.2576	3.1961
est	3.9443	4.2497	2.3852

	Gratian1	dePen	Gratian2
in	2.436	2.1606	4.0139
non	6.513	7.5667	5.5677
et	2.2589	4.2576	3.1961
est	3.9443	4.2497	2.3852

Given the layout of the *differences* dataframe in which we have stored the intermediate results, the part of the formula we deferred dealing with  $(\frac{1}{N} \sum_{i=1}^N)$  is simply a notationally exact way of indicating that we are to take the average (arithmetic mean) of the values in each of the columns, and record the resulting value of  $\Delta_B$  in the corresponding column of the *deltas* dataframe.

The seemingly simple act of taking the arithmetic mean (average) of the z-score distances between the samples for each feature has an interesting and non-intuitive implication. It was mentioned in passing in the previous section on visualization that plotting the z-score coordinates of word frequencies invokes the tacit assumption that the axes are in fact perpendicular to one another, an assumption that is at least potentially open to challenge. Burrows's Delta generalizes this assumption into an arbitrary number of dimension. The scholarly literature on authorship attribution methods describes distance metrics such as Burrows's Delta as measuring 'Manhattan Distance'. The analogy is to walking or driving from a starting to an ending point through a space in which the streets have been laid out at right angles to one another, like Manhattan.



```
row = (differences.mean(axis = 0)).to_frame(unknown).transpose()
```

Gratian1 dePen Gratian2

---

	Gratian0	3.788	4.5586	3.7907
--	----------	-------	--------	--------

The Gratian1 subcorpus is just slightly closer than the Gratian2 subcorpus to the unknown Gratian0 test case, with values of Delta for both rounding to 3.79. A candidate is defined as being *closest* to the unknown when it has the lowest mean of the absolute values of the differences between the z-scores for the unknown and the candidate. But as Burrows pointed out, one candidate will always have the lowest  $\Delta_B$ , so that in itself is not enough to make or to rule out an attribution of authorship. We will need further information before we can provide any kind of interpretation for the result. The most we can say based on this result is that the hypothetical case statements are less likely to have been written by the author of the *dicta* in *de Penitentia* than by the authors of either the first- or second-recension *dicta*.

The second experiment is a variation on the first, in which a 3881-word sample made up of seven extended passages from the pseudo-Augustinian *De vera et falsa penitentia* quoted by Gratian in *de Penitentia* are substituted for the 3605-word sample containing



the hypothetical case statements.<sup>36</sup> As noted in Chapter 0 above, Gratian can be said with a high degree of confidence **not** to be the author of *De vera et falsa penitentia*. The authors are strongly distinguished by their choice of post-positive conjunctions: Gratian has a preference for *autem*, while pseudo-Augustine has an even stronger preference for *enim*. Substituting the pseudo-Augustinian sample in place of the case statements demonstrates the kinds of results to be expected from Burrows's Delta in a situation in which an attribution of authorship can reasonably be ruled out.

Gratian1 dePen Gratian2

---

	psAug	2.6456	1.7373	3.4318
--	-------	--------	--------	--------

The third experiment extends the first by treating each of the subcorpora, Gratian0, Gratian1, dePen, and Gratian2 sequentially as the work of an unknown author, and the other three subcorpora as constituting a corpus of works by a known author. This is an

---

<sup>36</sup> *de Penitentia* D.1 c.88 (R1), D.3 c.42 (R1), D.3 c.49 (R1), D.5 c.1 (R1), D.6 c.1 (R1), and D.7 c.6 (R1). These seven extended passages average 554.4 words in length. See edF 1.XXXV, for a complete list of passages from *De vera et falsa penitentia* quoted in the *Decretum*. Explain rationale for omitting certain passages: D.25 c.5 (R2 or Palea), *de Penitentia* D.3 c.4.5 (what Friedberg means by 4.5 in this context is unclear), D.3 c.45 (R2). Acknowledge Karen Teresa Wagner, *De vera et falsa penitentia : an edition and study*, 1995.



attempt to demonstrate the adaptation of Burrows's technique in a circumstance in which there are no securely attributed comparison texts outside of the corpus, and in which there is some reason to suspect that there are multiple authors at work within the corpus.

```
# author candidates, e.g. Gratian 1, the Master of Penance, Gratian 2, etc.

candidates = ['Gratian0', 'Gratian1', 'dePen', 'Gratian2']

deltas = pd.DataFrame(columns = candidates)

limit = 4 # 4 most frequent words (MFWs)

for candidate in candidates:

    unknown = candidate

    samples = candidates[:]

    samples.remove(unknown)

    features = get_features(samples)

    mfws = list(features.keys())[:limit]

    counts = get_counts(mfws, [unknown] + samples)

    lengths = get_lengths([unknown] + samples)

    frequencies = (counts / lengths.values) * 1000

    means = frequencies[samples].mean(axis = 1).to_frame('mean')

    standard_deviations = frequencies[samples].std(axis = 1).to_frame('std')

    z_scores = (frequencies - means.values) / standard_deviations.values
```



```

test = z_scores[[unknown]]
corpus = z_scores[samples]
differences = (test.values - corpus).abs()
row = (differences.mean(axis = 0)).to_frame(unknown).transpose()
deltas = deltas.append(row)

```

Gratian0 Gratian1 dePen Gratian2

---

Gratian0	nan	3.788	4.5586	3.7907
Gratian1	1.4361	nan	0.3628	0.5453
dePen	1.9873	0.4515	nan	0.7673
Gratian2	1.7185	0.6278	0.7905	nan

Considering the results of the first three experiments together, we can start to form some very preliminary conclusions. Based on the values for  $\Delta_B$  in the table above, the most likely attribution is that the first-recension *dicta* (Gratian1) and the *dicta* from *de Penitentia* (dePen) have the same author. It is less likely that the first-recension *dicta* (Gratian1) and the second-recension *dicta* (Gratian2) have the same author. It is less likely still that the *dicta* from *de Penitentia* and the second-recension *dicta* have the same author. It is much less likely that the case statements (Gratian0) have the same author as



either the first- (*Gratian1*) or second-recension (*Gratian2*) *dicta*. Finally, the least likely attribution is that the case statements (*Gratian0*) have the same author as the *dicta* from *de Penitentia*.

---

The fourth and final experiment will compare the thirty most frequent words (MFWs) across fourteen subcorpora: cases (C.1-36 d.init.), laws (D.1-20 R1 *dicta*), orders1 (D.21-80 R1 *dicta*), orders2 (D.81-101 R1 *dicta*), simony (C.1 R1 *dicta*), procedure (C.2-6 R1 *dicta*), other1 (C.7-10 R1 *dicta*), other2 (C.11-15 R1 *dicta*), monastic (C.16-20 R1 *dicta*), other3 (C.21-22 R1 *dicta*), heresy (C.23-26 R1 *dicta*), marriage (C.27-36 R1 *dicta*), penance (R1 and R2 *dicta* from *de Penitentia*), and second (all R2 *dicta*, excluding those from *de Penitentia*).<sup>37</sup> For each of the fourteen subcorpora, we will hypothesize each subcorpus in turn to be the work of an unknown author, and will treat the other thirteen subcorpora

---

<sup>37</sup> The division of the first-recension (R1) *dicta* into twelve sections follows the division of Gratian's *Decretum* proposed by Alfred Beyer in *Lokale Abbreviationen Des Decretum Gratiani: Analyse Und Vergleich Der Dekretabbreviationen "Omnes Leges Aut Divine" (Bamberg), "Humanum Genus Duobus Regitur" (Pommersfelden) Und "de His Qui Intra Clastra Monasterii Consistunt"* (Lichtenthal, Baden-Baden), Bamberger Theologische Studien ; Bd. 6 (Frankfurt am Main ; PLang, 1998), 17-18.



as composing a corpus of works by a known author. The scale of the fourth experiment is similar to that of the experiments carried out by John Burrows and David Hoover, the pioneers of the technique, but makes it impractical to show intermediate results at every step in the process.

	cases	laws	orders1	orders2	simony	procedure	other1
cases	nan	2.2765	1.9247	2.0252	1.9637	1.9545	1.5714
laws	2.141	nan	1.249	1.502	1.4633	1.3147	1.4223
orders1	1.6184	1.0949	nan	1.1223	0.9685	0.8843	1.0499
orders2	1.8982	1.5244	1.2686	nan	1.382	1.684	1.4149
simony	1.6667	1.3491	0.9772	1.2195	nan	0.8878	1.1304
procedure	1.6187	1.1991	0.892	1.5095	0.8789	nan	1.079
other1	1.3353	1.3	1.0619	1.2722	1.1383	1.0753	nan
other2	1.9416	1.3233	1.0913	1.6291	1.1386	1.109	1.2963
monastic	1.4555	1.0451	0.8554	1.2676	1.0114	0.7986	0.93
other3	2.0705	1.3388	1.289	1.5146	1.1997	1.1057	1.3497
heresy	1.5177	1.031	0.7772	1.2182	0.5544	0.595	0.9485

	marriage	1.5448	1.0263	0.9848	1.265	0.884	0.9667	1.0494
	penance	1.5371	1.4473	0.7478	1.4005	0.9024	0.8781	1.3077
	second	1.374	1.0852	0.7764	1.1717	1.0623	0.9634	0.7971
		other2	monastic	other3	heresy	marriage	penance	second
cases		2.2782	1.7622	2.3628	1.8717	1.8923	1.8589	1.6334
laws		1.4369	1.1931	1.4345	1.1875	1.1924	1.6218	1.2323
orders1		1.1109	0.8693	1.2397	0.8267	1.0124	0.7505	0.7777
orders2		1.6873	1.4492	1.6208	1.4198	1.4526	1.5523	1.3195
simony		1.1287	1.0413	1.1711	0.59	0.9166	0.9059	1.0863
procedure		1.1223	0.821	1.0726	0.6569	0.9993	0.8818	0.9852
other1		1.2792	0.9649	1.3054	0.996	1.0853	1.3272	0.8152
other2		nan	0.7979	1.0346	1.0592	0.654	0.8633	1.0961
monastic		0.7429	nan	1.0578	0.7602	0.6611	0.7999	0.7799
other3		0.9505	1.1229	nan	1.1209	0.7121	1.1521	1.3067
heresy		0.9839	0.7672	1.1395	nan	0.7783	0.6756	0.8484
marriage		0.6126	0.6577	0.6552	0.7672	nan	0.7974	0.8676

penance	0.9152	0.8101	1.0992	0.7609	0.8146	nan	0.9026
second	1.0674	0.7861	1.2408	0.877	0.8796	0.8927	nan

# Chapter 4

## Principal component analysis

Techniques such as Burrow's and Argamon's Delta (measuring Manhattan and Euclidean distance respectively), which collapse vector distance data for an arbitrary number of features or dimensions into a single scalar value interpreted as a nearest-neighbor classification metric, are one way of reducing feature distances to a tractable form. Principal component analysis (PCA) is an alternative to Delta metrics that projects vector distance information for numbers of features greater than three into a two- or three-dimensional space for convenient visualization. PCA therefore has the advantage that it entails less loss of information than the Delta class of techniques that reduce data for all dimensions to a single metric.<sup>1</sup> (**Updated 24 June 2020.**)

---

<sup>1</sup> Earlier versions of this section were presented as conference papers. "Can Stylometry Provide New Evidence about the Identity of Gratian 1 and Gratian 2?", was presented to the session on Canon Law in the Twelfth and Thirteenth Centuries at the *Rem non novam nec insolitam aggredimur* conference and grand opening of the Stephan Kuttner Institute of Medieval Canon Law at Yale Law School, May 21-22, 2015. [Greta Austin, Thomas Bisson, Uta-Renate Blumenthal, Bruce Brasington, Melodie Eichbauer, Richard Helmholz, Eric Knibbs, Peter Landau, Kenneth Pennington, Edward Peters (University of Pennsylvania), Robert Somerville, and Anders Winroth.] "New evidence for the authorship of case statements and *dicta* in Gratian's *Decretum*" was presented to the Classical Sources III session at the Fifteenth International

PCA first combines as many of the raw dimensions as possible into synthetic components on the basis of strong correlations, either positive or negative. For example, referring back to Figures 0a and 0b in the two-dimensional visualization section above, the two dimensions of the plot could be collapsed into a single axis or component that can be thought of as representing the frequency with which *in* does, and *non* does *not*, occur in a given sample. The effect would be to reconfigure the plots in such a way that the samples representing the *dicta* from *de Penitentia* (dePen), the first-recension *dicta* (Gratian1), and the second-recension *dicta* (Gratian2) would be placed from left to right along a single horizontal axis. PCA then displays the two components that contribute the most to the total variation between the samples, and graphically arranges the samples according to their probability relative to those two components.<sup>2</sup> (Updated 24 June 2020.)

---

Congress of Medieval Canon Law (ICMCL) at Université Paris II Panthéon-Assas, July 17-23, 2016. [Gero Dolezalek, Anders Winroth (session chair).]

<sup>2</sup> For a general introduction to the use of principal component analysis (PCA) in literary stylometric analysis, see Hugh Craig, "Stylistic Analysis and Authorship Studies," in *A Companion to Digital Humanities*, ed. Susan Schreibman, Raymond George Siemens, and John Unsworth, Blackwell Companions to Literature and Culture 26 (Malden, MA: Blackwell Pub, 2004), 273–88 and Chapter 6



For the case statements, I simply used the text from the vulgate *Decretum* as it appears in the Friedberg edition.<sup>3</sup>

Because stylometric analysis for authorship attribution depends on the frequencies of prepositions and conjunctions, it is important to include enclitics substituting for conjunctions. Every word in the samples with a *-que* ending that is actually an enclitic, and not just part of the word, has been mapped to the word plus the pseudo-conjunction *xque*.<sup>4</sup>

---

"Style" in Matthew Lee Jockers, *Macroanalysis: Digital Methods and Literary History*, Topics in the Digital Humanities (Urbana: University of Illinois Press, 2013).

<sup>3</sup> This is perhaps not entirely satisfactory. It would be more methodologically consistent with the way in which the samples of the first-recension *dicta* were prepared to apply the differences found in Winroth's appendix to the case statements as well, however the differences are quite minimal. The only case statement (*thema*) for which Winroth notes a textual difference is C.19 d.init. (Winroth, 216). The first-recension version of the text omits a 13-word clause added to the second recension version, seemingly for the purpose of piling up descriptive detail. (*unus relicta propria ecclesia eo inuitio, alter dimissa regulari canonica cenobio se contulit*). None of the wordlists used to perform the principal component analyses include any of these 13 words, so the use of the vulgate rather than a proxy first-recension version of the text of C.19 d.init. has no effect on the outcome of these tests.

<sup>4</sup> In the case statements, 1st-, and 2nd-recension *dicta* from Gratian's *Decretum*, there are 747 occurrences of 79 unique words ending in *-que*. (This does not count 423 occurrences of the word *que* itself.) Of those,



Now that we have the preliminaries out of the way, we can take a look at the results. I used the *stylo* R package to generate all of the plots that I'm going to show you today.<sup>5</sup> R is a statistical programming language.<sup>6</sup> Mike Kestemont, Maciej Eder, and Jan Rybicki

---

498 are occurrences of 19 unique words from Schinke's 54-word pass list, while 249 occurrences of 60 unique words are not. It is from these 249 words that, according to Schinke, the *-que* ending should be detached as an enclitic.

However, the 249 words include 72 occurrences of 17 unique words ending with the adverbial enclitics *-cumque* or *-cunque*, from which the *-que* ending should not be detached. The 249 words also include a further 149 occurrences of 21 unique false positives:

*cumque*, *eque* (*aeque*), *namque*, *pleraque*, *plerique*, *plerisque*, *plerumque*, *quinque*, *unamquamque*, *unaqueque*, *unicuique*, *uniusciusque*, *unumquemque*, *unusquisque*, *usquequaque*, *utramque*, *utraque*, *utriusque*, *utrisque*, *utriusque*, *utrumque*.

This leaves only 28 occurrences of 22 unique words from which the *-que* ending should actually be detached as an enclitic.

False positives over-represent the frequency of occurrence of the *-que* enclitic as a conjunction by an order of magnitude. Including all false positives makes *xque* the 37th most frequent word in the sample, while excluding them makes it the 376th most frequent word. There are 55 occurrences of the word *namque*, the most frequently occurring false positive. Detaching the *-que* ending from *namque* overstates the frequency of *nam*, making what is actually the 480th most frequent word appear to be the 130th, while making *namque*, which is actually the 176th most frequent word in the samples when false positives are excluded, disappear from the list altogether.

<sup>5</sup> Maciej Eder, Mike Kestemont, and Jan Rybicki, "Stylometry with R: A Suite of Tools," in *Digital Humanities 2013: Conference Abstracts* (Lincoln, NE: University of Nebraska–Lincoln, 2013), 487–89, <http://dh2013.unl.edu/abstracts/>.

<sup>6</sup> R Core Team, *R: A Language and Environment for Statistical Computing* (Vienna, Austria: R Foundation for Statistical Computing, 2015), <https://www.R-project.org/>.



of the Computational Stylistics Group developed the package, and Mike Kestemont in particular has been very generous in his technical advice for this project.



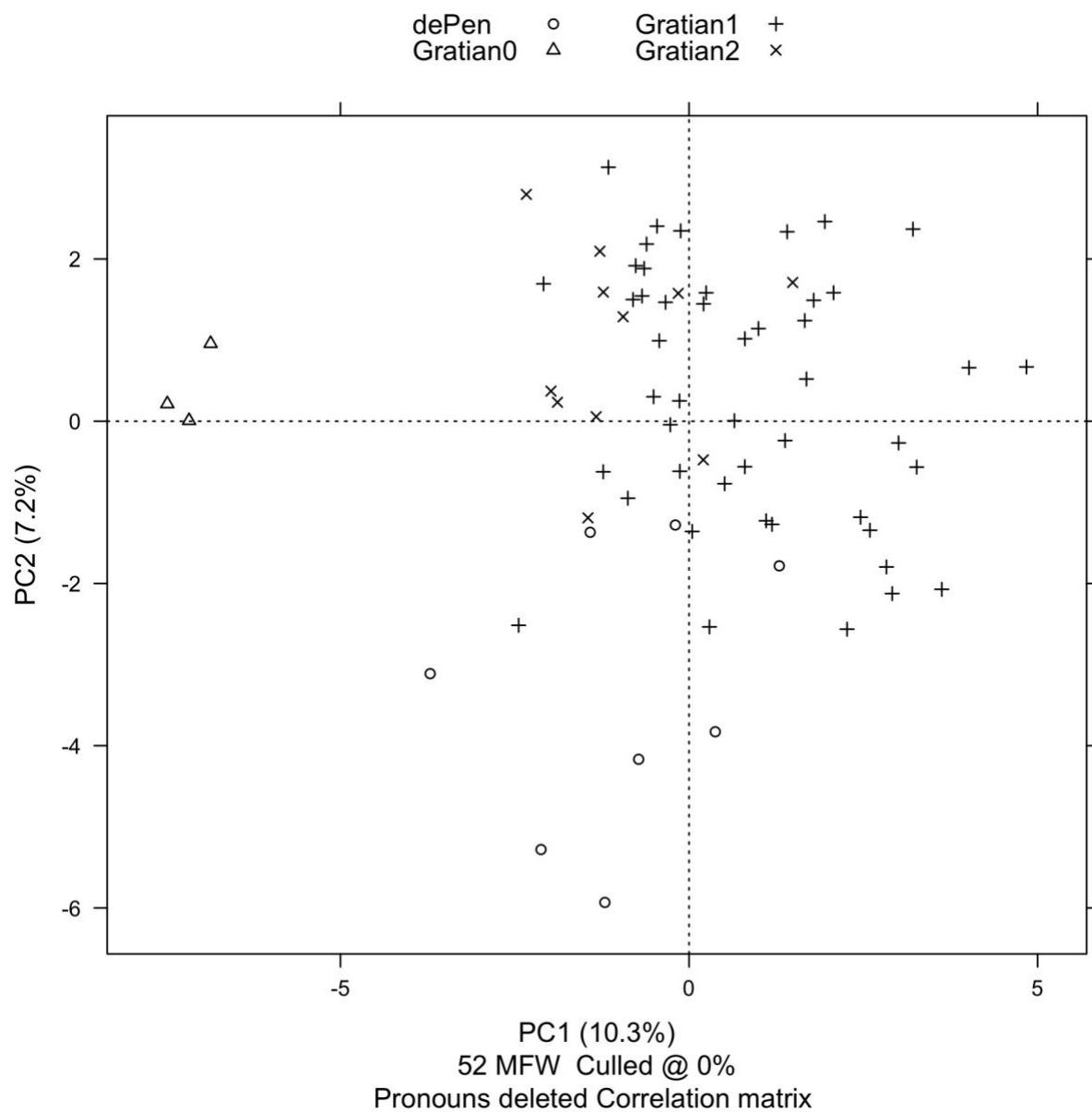


Figure 1 updated 28 May 2020

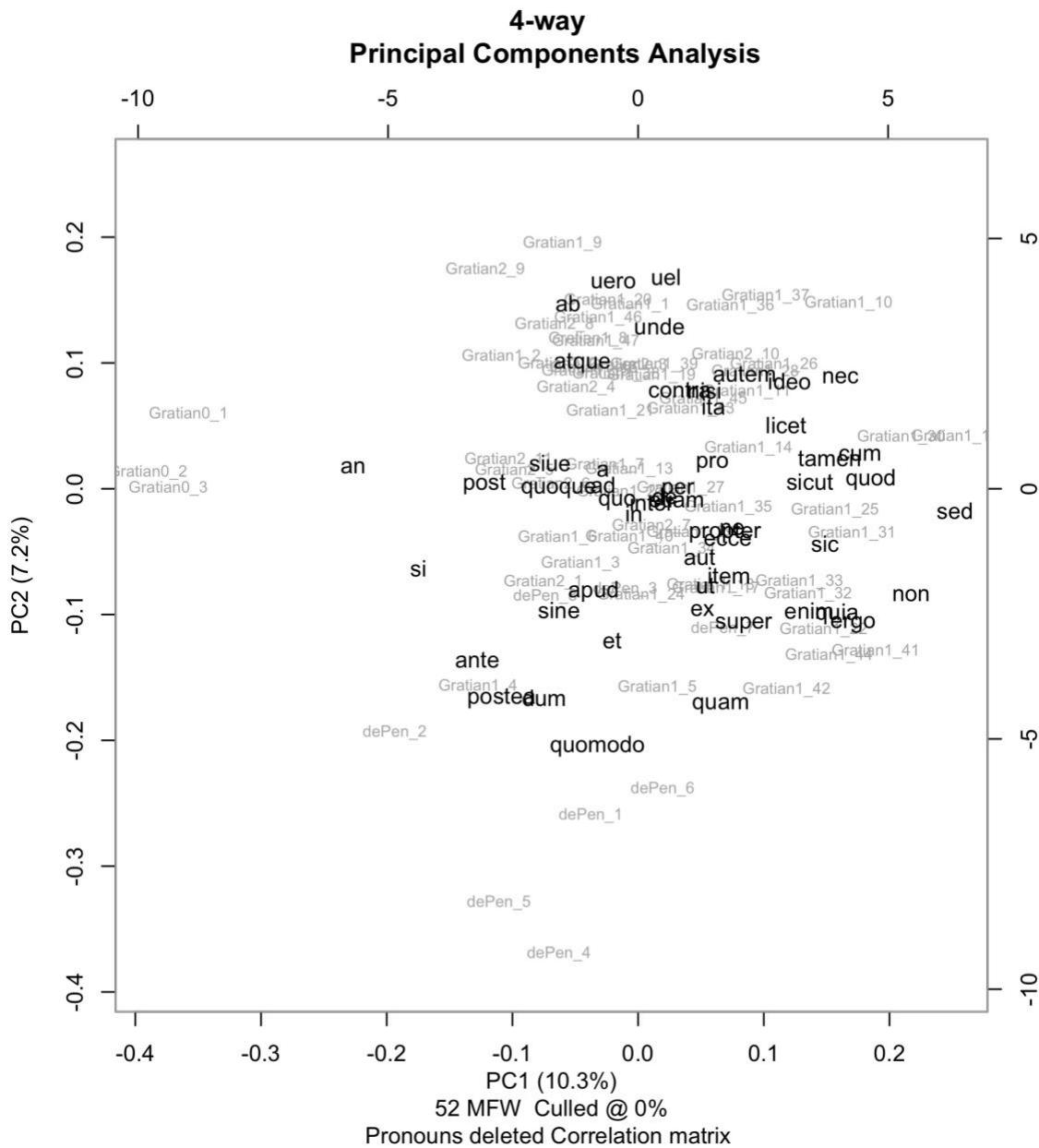
-----  
 features (e.g. frequent words) actually analyzed  
 -----

```
[1] in      et      non     de      quod    ad      sed      uel
[9] unde   si      ut      a       autem   cum     ex      enim
[17] uero   etiam   ab      ergo    quia    item   per      nec
[25] an     sicut   ita     nisi    tamen   pro    quam   sic
[33] quo    sine    aut    licet   post    contra siue   quoque
[41] ante   ne     inter   super   atque   dum    apud   postea
[49] ideo   propter ecce   quomodo
```

Figure 1 shows the PCA plot generated by a four-way comparison of the same samples used in the demonstration of Burrows's Delta in the previous section: the hypothetical case statements or *themata* (Gratian0), the first-recension *dicta* excluding the *dicta* from *de Penitentia* (Gratian1), first- and second-recension *dicta* from *de Penitentia* (dePen), and the second-recension *dicta* excluding the *dicta* from *de Penitentia* (Gratian2). The case statements are magenta, the first-recension *dicta* are green, the *dicta* from *de Penitentia* are blue, and the second-recension *dicta* are red. Each of the texts has been divided into 1200-words samples. Principal component 1 along the horizontal axis is 10.3%. Principal



component 2 along the vertical axis is 7.2%. That is, PC1 explains 10.3% of the total variation between the samples, and PC2 explains 7.2% of the total variation between the samples. This is good: as a general rule, we want to see a value for PC1 greater than 10% and we want to see a value for PC2 greater than 5%. The most visually striking aspect of this plot is the fact that the case statements are so far away from the *dicta*, and the next step is to take a look at which features producing that effect. (**Updated 24 June 2020.**)



*Figure 2 updated 28 May 2020*

Turning on the stylo feature loadings option lets us see how strongly particular words influence the placement of text samples along the PC1 and PC2 axes. This is called the feature's discriminative strength. For example, we see that *sed* and *non* are way out on the right of the PC1 axis, while *unde* is way down at the bottom of the PC2 axis.

Remember that in our first experiment with counting function words, *non*, the second most common word in the samples, was strongly associated with the first-recension *dicta*. Here we see *non* on the far right, and in fact the samples from the first-recension *dicta* (but not from the second-recension *dicta*) tend to spread out to the right. Note also that *in*, the most common word in the samples, is actually pretty close to the middle. So, it's not so much that the second-recension *dicta* have more occurrences of *in*, it's that the first-recension *dicta* have fewer.

What is really interesting here is that *an* and *si* cluster with the case statements, *an* very strongly, *si* somewhat less so. This makes sense because indirect questions dominate the language of the case statements. It is a question of genre. So the next step in the

stylistic analysis is to control for genre by removing the question words *an* and *si* from the list of function words.

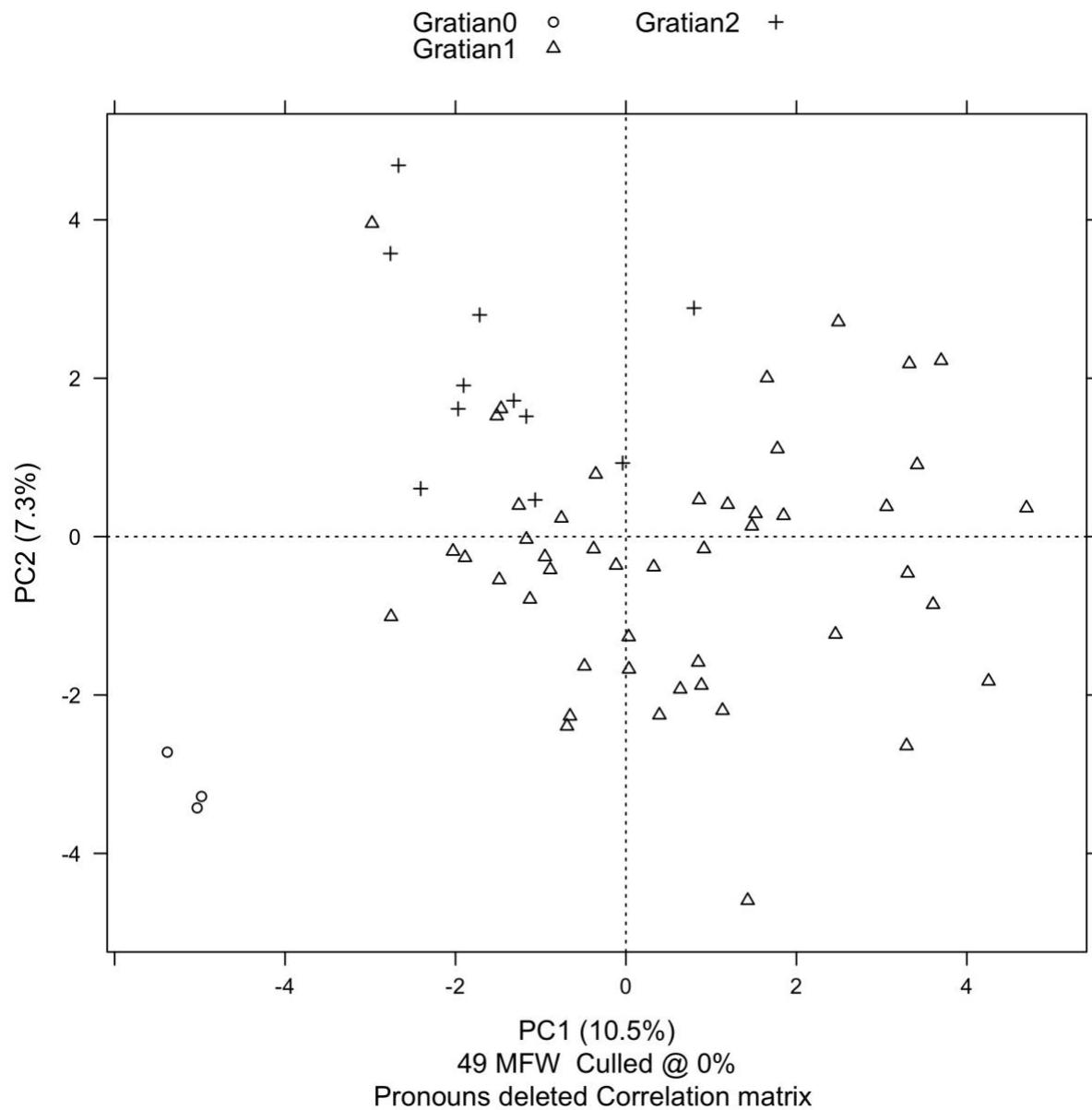


Figure 3 updated 19 Mar 2020

We've now reached the final stage of the three-way comparison between the case statements, the first-recension *dicta*, and the second-recension *dicta*. We are now using the 49 most frequent words on our function list instead of the 51 most frequent words, having commented out *an* and *si*. And even without *an* and *si*, PC1 still explains 10.5% of the total variation between the samples, down slightly from 11.2%. PC2 still explains 7.3% of the total variation between the samples. So, even controlling for genre, the distance between the case statements and the *dicta*—both first- and second-recension—is still quite striking.

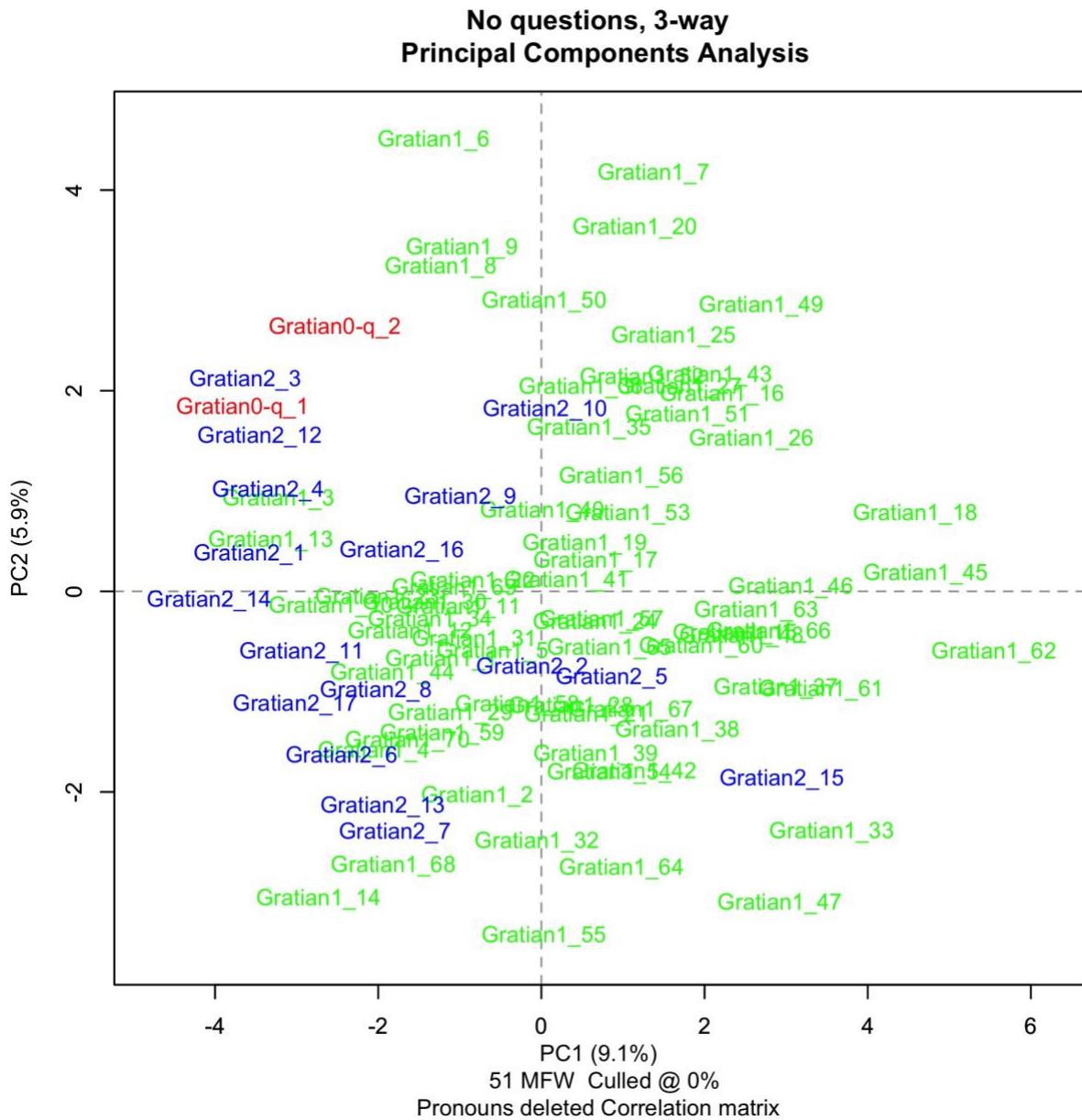


Figure 4 updated 25 Jun 2019

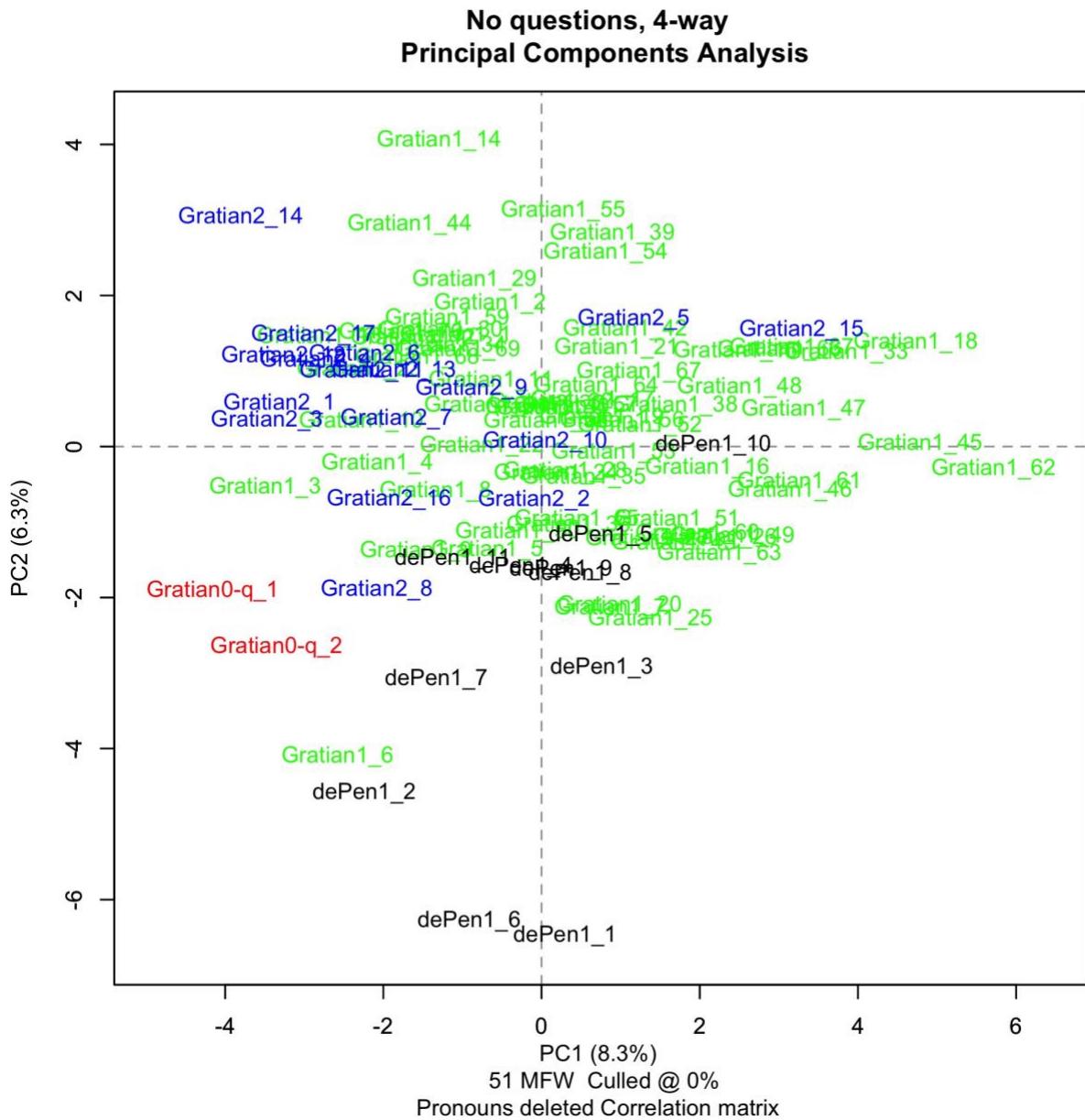


Figure 5 updated 25 Jun 2019

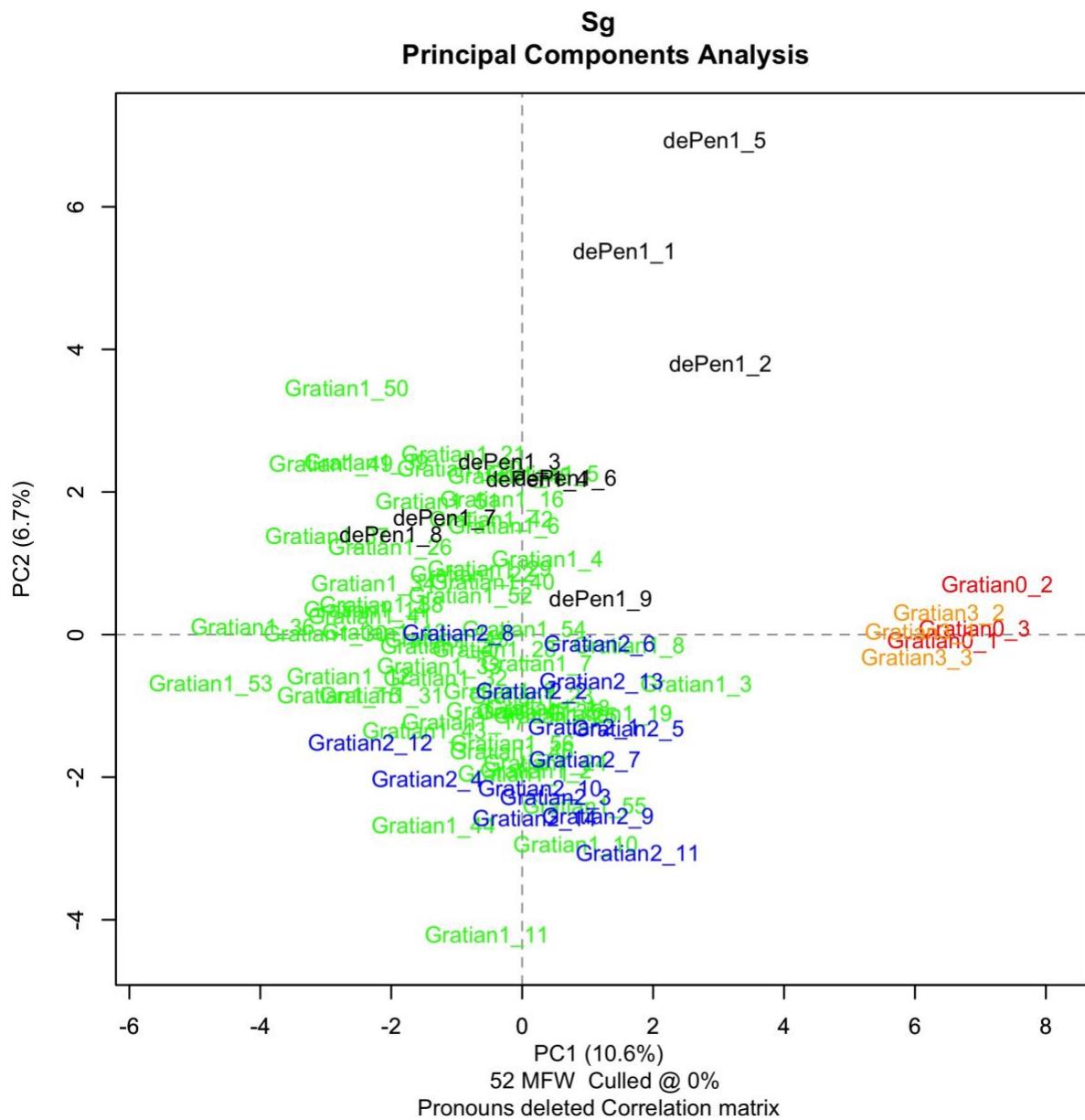


Figure 6 updated 21 Jan 2019

To turn to the other interesting aspect of the three-way comparison, you'll note that the second-recension *dicta* in blue cluster strongly to the upper-left quadrant. Now, Mike Witmore, a member of my dissertation committee who isn't an insider with respect to debates about Gratian's *Decretum*, but is very experienced in the use of stylometry with the plays of Shakespeare, was somewhat optimistic on the basis of this evidence that the first- and second-recension *dicta* might be statistically distinguishable.



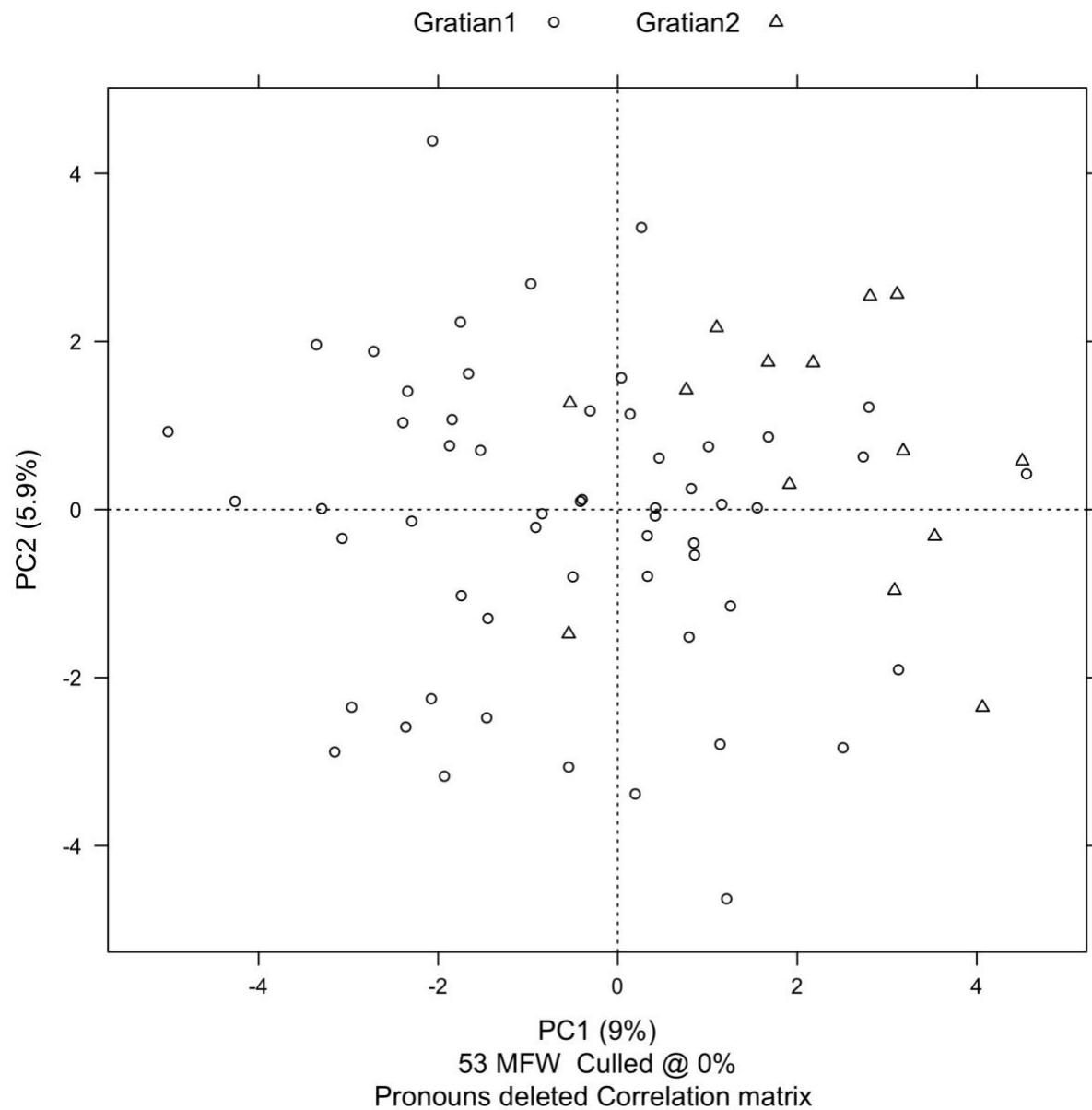


Figure 7 updated 19 Mar 2020

So, in an attempt to take a closer look at the *dicta* by themselves, I removed the case statements and ran a two-way comparison of 1000-word samples of just the first- and second-recension *dicta*, again, excluding the *dicta* from *de Penitentia*. (Stylo changes the color assignments depending on the number of samples, so in this plot the first-recension *dicta* are red and the second-recension *dicta* are green.) And the results are ambiguous. The PC1 axis is 9%, somewhat under the 10% threshold we would like to see. Also, although we see the second-recension *dicta* clustering mostly to the right of the PC1 axis, the two sets of samples are not separated as cleanly as we'd like to see, and certainly nowhere near as cleanly as the case statements were from the *dicta*.

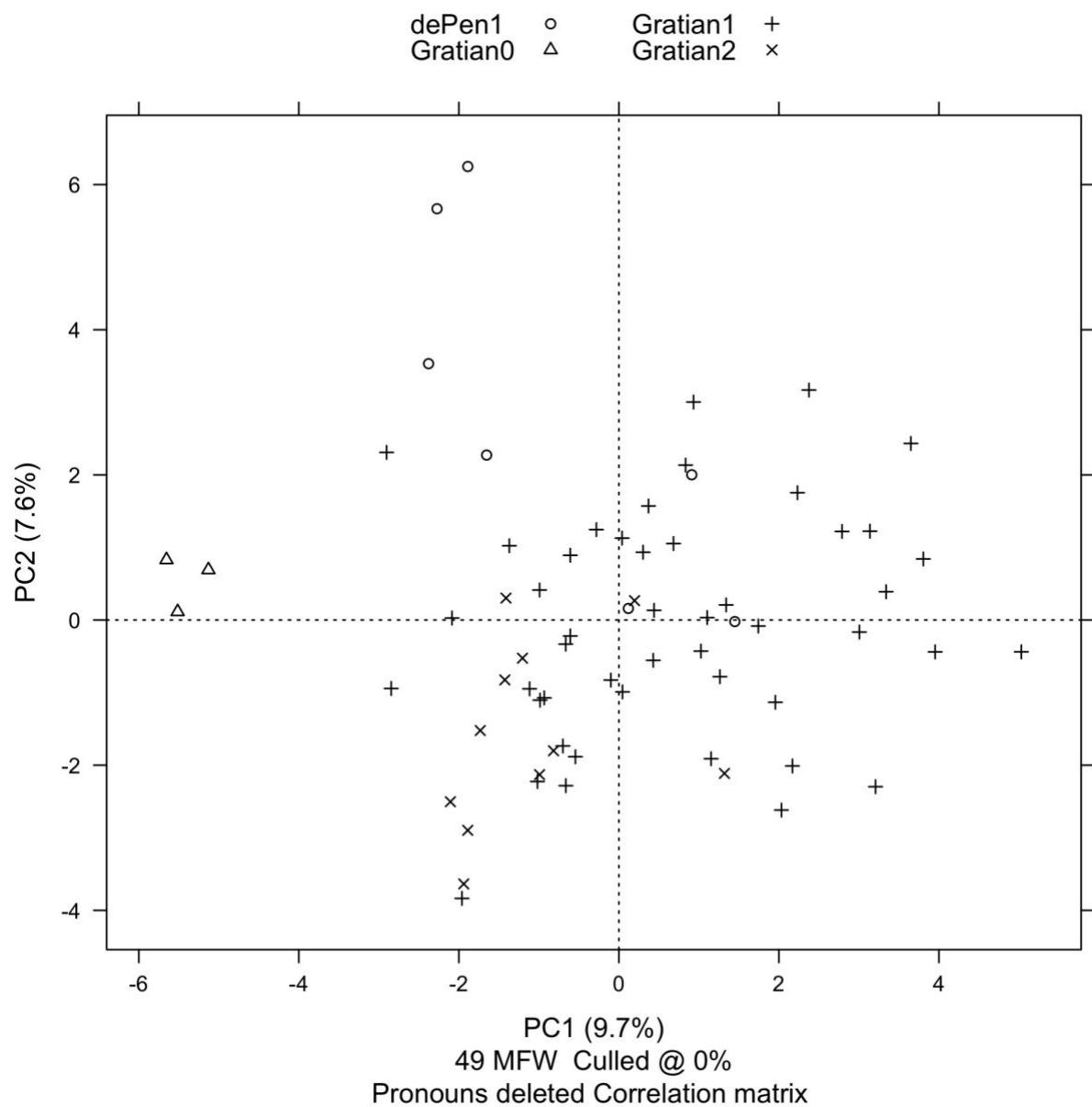


Figure 8 updated 19 Mar 2020

All of the slides we've seen so far exclude the *dicta* from *de Penitentia*, so before moving on to my conclusion, I do want to quickly show you what the results look like when we include the first-recension *dicta* from *de Pen.* (there are not enough words in the second-recension *dicta* in *de Pen.* to be statistically significant—9,525 vs. 556). Many scholars have observed that *dicta* and canons are poorly separated in *de Pen.* I believe that the unusual dispersion of the samples that you see in this plot is a result of that feature.

## Conclusion

Principal component analysis (PCA) of the frequencies of function words (prepositions and conjunctions) in the texts strongly suggests that the author of the case statements was not the same person as the authors of either the first- or second-recension *dicta*. PCA also suggests (less strongly) that the first- and second-recension *dicta* were not the work of either one or two authors, but are more likely to have been the product of collaborative authorship.

On Monday, Anders presented a sketch of what a stemma for the first recension might look like. It suggested that the textual transmission was far more complicated than we



may have imagined (or at least may have hoped for). The results I've presented here today suggest that the question of authorship is potentially as complicated as the question of transmission. I believe that there is enough evidence at least to question assumptions of monolithic authorships (of either the one Gratian or two Gratians variety). If we cannot satisfactorily answer the question "was there one Gratian or were there two?" it is probably because that is not the right question to ask.



# **Chapter 5**

## **Close Reading of the Case Statements**

[Introduction to comparative translations of the Sg and Friedberg (vulgate) texts of the case statements or *themata*.]

The rationale for using Friedberg as the baseline for comparison is that Winroth's edition in progress of the first recension is not yet complete enough for the purpose. As of the most recent, 22 April 2019, version, eight case statements (for cases 1-3, 9, 15, 24, 30, and 34) appear to have a complete critical apparatus. An addition six case statements (for cases 4-7, 10, and 11) have an incomplete critical apparatus that records variant readings from Fd only. The critical apparatus for the case statement for case 35 records a single variant reading from Aa. The remaining 21 case statements (for cases 8, 12-14, 16-23, 25-29, 31-33, and 36) have no critical apparatus at all.

Fortunately, the text of the case statements is fairly stable over the course of development of the *Decretum* between its first recension and vulgate versions. Winroth

in the appendix to *The Making of Gratian's Decretum*<sup>1</sup> reports only one significant variant reading for a case statement. The first-recension version of C.19 d.init. omits the phrase "unus relicta propria ecclesia eo inuito, alter dimissa regulari canonica cenobio se contulit." That the vulgate supplies the missing phrase is meaningful because it makes sense of the fact that there are two clerics in the story, an otherwise inexplicable feature of the first-recension and Sg versions of the case statement. Textual differences between the first-recension and vulgate versions of the other case statements are relatively minor. (e.g., "episcopus minime concurrit" AaBcFdP versus "episcopus minime occurrit" Fr. in C.4 d.init.)

This approach—applying the variants recorded in Winroth's appendix to Friedberg's vulgate text to generate a stand-in or proxy for the first recension—is consistent with the approach I have taken throughout my dissertation project. Winroth himself took a similar approach when he created the baseline text for the edition in progress of the first recension (although he imposed a set of orthographic conventions different from

---

<sup>1</sup> Anders Winroth, *The Making of Gratian's Decretum* (Cambridge: Cambridge University Press, 2000), 216.



Friedberg's on the resulting text). Regardless of how Winroth conceptualized what he was doing, it is an approach that is well-theorized in a Digital Humanities context as an example of "deformance." The term, proposed by Lisa Samuels and Jerome McGann in "Deformance and Interpretation" (1999), conflates the words "deformation" and "performance", and describes a process through which a text is transformed by the application of a series of deformances to generate a "paratext".<sup>2</sup> The paratext is different from the original text, but defined by the deformances through which it was generated from it, and can be analyzed for otherwise unavailable insights into the original text.

## Sg

The primary focus of my dissertation has been on the textual relationship between the case statements in Part II of Gratian's *Decretum* and the first-recension *dicta*. As I have discussed at some length in Chapter 4 above, the author of the case statements appears unlikely to have been the same person as the author of the first-recension *dicta*, the person we usually think of as Gratian. The text of the case statements is relatively stable

---

<sup>2</sup> Lisa Samuels and Jerome McGann, "Deformance and Interpretation," *New Literary History* 30, no. 1 (1999): 25–56.



over the course of the development of the *Decretum* from the first recension to the vulgate. There is, however, another version of the text of the case statements that needs to be considered, the one found in the Sankt Gallen Stiftsbibliothek 673 (Sg) manuscript of the *Decretum*.

The version of the *Decretum* preserved in Sg contains around a thousand canons, and it is therefore considerably shorter than either the first recension (1,860 canons) or the vulgate (3,945) versions of the text.<sup>3</sup>

Formally, Sg is not divided into parts; all of its content is presented in the form of cases. The first case, however, unique to Sg and referred to as Causa Prima to distinguish it from Causa I found in the first recension and vulgate versions, contains a subset of the canons found in Part I in other versions of the *Decretum*. Sg has 33 cases, adding Causa Prima and omitting cases 24, 25, 26, and 28.

---

<sup>3</sup> See Carlos Larraínzar, "El Borrador de La 'Concordia' de Graciano: Sankt Gallen, Stiftsbibliothek MS 673 (=Sg)," *Ius Ecclesiae: Rivista Internazionale Di Diritto Canonico* 11, no. 3 (1999): 593–666 for number of canons in Sg, emphasizing that it is approximate. Pennington has 1,050. See Winroth, *The Making of Gratian's Decretum* for number of canons in first recension. 3,945 is a conventional number.]



The Sg version of the text of the case statements has 3,226 words, shorter than but comparable to the first-recension and vulgate versions (3,605 words confirm). The text of the 32 directly corresponding case statements, removing C.Prima for Sg, and removing C.24, 25, 26, and 28 for the vulgate, is slightly more compact in Sg (3,173 words) than in the vulgate (3,312 words). **The Sg cases statements are not more “verbose” than their vulgate counterparts.** There are, however noticeable differences in vocabulary.

## **Comparative Translation of C.4 d.init.**

### **Sg text:**

(72b) In excommunicatione quidam erat constitutus et episcopum accusare disponit; adolescentem infra annos xiiii secum ad asserendam causam adducit. Qui dum ab accusatione prohibetur, facit se testem; et minorem accusatore minor vero accusatoris et testis personam gerere desiderat. Statuta die episcopus ad iudicium non venit suspenditur a communione. Renovato denuo iudicio accusator in accusatione culpabilis invenitur. Demum ad assertionem propriae causae procedit. Queritur ergo primo, an constitutus in excommunicatione possit alium accusare? Secundo, an infra xiiii annos in criminali (73a) causa quis testificari valeat?



Tercio, an ab accusatione quis prohibitus testificantis vicem possit assumere? Quarto, accusator et testis an possit idem esse? Quinto, an quis non veniens die praestituta excommunicari debeat id est a communicatione privari? Sexto, si culpabilis quis in episcopali iudicio inveniatur ad assertionem propriae causae sit amplius admittendus?

### Sg translation:

Someone was excommunicated and arranges to accuse a bishop; he brings in an adolescent under 14 years for the purpose of asserting the case with him. Who, when he is prohibited from accusation, makes himself a witness and the minor the accuser;<sup>4</sup> but the minor wishes to perform the role of [both] accuser and witness. On the appointed day, the bishop does not come to court [and] is suspended from communion. After the court reconvened again,<sup>5</sup> the accuser is found [to be] at fault in [bringing] the accusation. Finally, he proceeds to the assertion of his own case. It is asked therefore first whether an excommunicate can accuse another? Second, whether someone (*quis*) under 14 years is able to testify in a criminal case? Third, whether someone (*quis*)

<sup>4</sup> Reading here as *accusatorem*; presumably as a result of scribal error, Sg reads *accusatore*.

<sup>5</sup> lit. “the court having been renewed again”.



prohibited from accusation can assume the part of testifying? Fourth, whether the same person can be [both] accuser and witness? Fifth, whether someone (*quis*) not coming on the predetermined (*praestituta*) day ought to be excommunicated, that is, deprived of communion? Sixth, if someone (*quis*) found [to be] at fault in the episcopal court ought to be further (*amplius*) admitted to the assertion of his own case?

One of the apparent contrasts between the Sg and vulgate (Fr.) versions of C.4 d.init. is in their use of legal technical terms. The refusal of the accused bishop to appear before the court on the appointed day, presumably because he does not recognize its authority to try a case arising from an improper accusation, is described in the Sg version using everyday language: "*episcopus ad iudicium non venit* (the bishop does not come to court)." Similarly in the fifth question arising out of his refusal: "*an quis non veniens die praestituta excommunicari debeat* (whether someone not coming on the predetermined day ought to be excommunicated)." The vulgate (Fr.) version of C.4 d.init., however, uses *minime occurrit* in place of *non venit* and *occurrens* in the place of *veniens* in a way that presumes that *minime occurrit* is a legal technical term with almost precisely the

same meaning we use today when we say that a criminal defendant has been cited for failure to appear in court.<sup>67</sup>

Excommunication is discussed in the Sg version of C.4 d.init. in a way that indicates that the concept is not an entirely settled one. On the one hand, the phrase *in excommunione constitutus* seems to be used in the Sg text of the case statement as a legal technical term: someone, like the accuser, who is said to be *in excommunione constitutus* seems to have been permanently excommunicated, while the accused bishop is said to have been *suspenditur a communione* (suspended from communion), by implication only temporarily, for his failure to appear in court. On the other hand, the Sg version of the case statement takes care to define excommunication: “*id est a communicatione privari*

<sup>6</sup> There is also a literary precedent for translating the somewhat opaque phrase *minime occurrit* as “does not appear” in the commentary of Maurus Servius Honoratus on Virgil: “*aut audire quod non dicitur, videre quod minime occurrit* (or to hear what is not spoken, to see what does not appear)”. **Add reference to critical edition of Servius.**

<sup>7</sup> AaBcFdP reads *minime concurrit* rather than *minime occurrit*. It is not clear whether this variant is meaningful, or whether it simply perpetuates an error introduced into the hypothesized α hyparchetype from which all first-recension manuscripts are now thought to have been derived. **Add reference to Winroth article.**



(that is, to be deprived of communion)", something the composer of the vulgate (Fr.) version of the case statement does not necessary.

The Sg version of C.4 d.init. uses the present tense, possibly intended to be understood as historical present, consistently throughout the case statement.

"After **sī**, **nisi**, **num**, and **nē**, all the **ali-**'s drop away." Moreland and Fleischer, 217. See also Gildersleeve and Lodge, §107; Allen and Greenough, §310a.

### **Fr. text:**

(Fr. 536) Quidam in excommunicatione constitutus episcopum accusare disponit; adolescentem infra decimum et quartum etatis suaे annum ad assertionem suaे causae adducit; prohibitus ab accusatione adolescentem accusatorem et se testem facit; adolescens personam accusatoris et testis gerere cupit; die statuta ad electorum iudicium iudicium episcopus minime occurrit; a communione suspenditur; tandem renouato iudicio accusator culpabilis in accusatione inuenitur; demum ad assertionem propriae causae procedit. Hic primum queritur, an in excommunicatione constitutus alium accusare ualeat? Secundo, an infra decimum quartum annum in criminali causa testari quis possit? Tertio, utrum ab accusatione prohibitus personam testificantis possit assumere? Quarto, an idem possit



esse accusator et testis? Quinto, an die constituta non occurrens a communione **ait** remouendus? Sexto, si in episcoporum **indicio** accusatoris persona culpabilis inuenta fuerit, an ad assertionem propriae causae de cetero sit admittenda?

The MGH e-text edited by Reuter and Silagi introduced the OCR errors **ait** into the text of the fifth question (*Quinto, an die constituta non occurrens a communione ait remouendus?*) and **indicio** into the text of the sixth question (*Sexto, si in episcoporum indicio accusatoris persona culpabilis inuenta fuerit ...*). The printed version of Friedberg correctly reads *a communione sit remouendus* and *in episcoporum iudicio*.

### **Fr. translation:**

Someone permanently excommunicated (*in excommunicatione constitutus*) arranges to accuse a bishop; he brings in (*adducit*) an adolescent under fourteen years of age to the assertion (t.t.) of his case; prohibited from accusation, he make the adolescent the accuser and himself the witness; the adolescent wishes to conduct/perform (*gerere*) the role (*personam*) of [both] accuser and witness; on the appointed day, the bishop does not appear before (*lit. to*) the court of the judge-electors; he is suspended from communion; at last, after the court reconvened [*lit. the court having been*

renewed], the accuser is found [to be] at fault in [bringing] the accusation; finally he proceeds to the assertion (t.t.) of his own case.

**AaBcP text:**

Quidam in excommunicatione constitutus episcopum accusare disponit; adolescentem infra xiiiium etatis sue annum ad assertionem sue cause adducit. Prohibitus ab accusatione adolescentem accusatorem et se testem facit. Adolescens personam accusatoris et testis gerere cupit; die statuta ad iudicium electorum iudicum episcopus minime concurrit; a communione suspenditur; tandem renouato iudicio accusator culpabilis in accusatione inuenitur; demum ad assertionem proprie cause procedit. Hic primum queritur, an in excommunicatione constitutus alium accusare ualeat? Secundo, an infra xiiiium annum in criminali causa testari quis possit? Tercio, utrum ab accusatione prohibitus personam testificantis possit assumere? Quarto, an idem possit esse accusator et testis? Quinto, an die constituta non occurrens a communione sit remouendus? Sexto, si in episcoporum iudicio accusatoris persona culpabilis inuenta fuerit, an ad assertionem proprie cause de cetero sit admittenda?

**Quidam]** AaBc *om.* P **excommunicatione]** AaBc excommunicatione P **adolescentem]** AaP adulescentem Bc **etatis...annum]** BcP annum etatis sue Aa **die]** AaBc di P **iudicium]** AaP iuditium Bc **iudicum]** AaBc iudicium P **concurrit]** AaBcP occurrit Fr. **suspenditur]** AaBc suspendere P **iudicio]** AaP iuditio Bc **an]** AaBc autem P AaBc infra P **criminali]** AaBc criminari P **ab]** AaBc *om.* P **an...Quinto,]** AaBc *om.* P **an]** AaBc aut P **iudicio]** AaP iuditio Bc

## Comparative Translation of C.6 d.init.

### Sg text (93 words):

(77a) Fornicatores quidam et notati infamia episcopum quendam de symonia nituntur inpetere. Expetit reus iudicium alterius archiepiscopi provinciae, accusatores tandem in probatione deficiunt, ad suae innocentiae assertionem cogitur reus. Queritur ergo utrum crimine irretiti seu infamia notati, ad huiusmodi accusationem sint admittendi? Secundo, an episcopus si in eos accusationem retorquere voluit, an simplici assertioni suae cedendum sit? Tercio, si liceat sibi expetere iudicium archiepiscopi alterius provinciae? Quarto, cuius sibi iudicium sit expetendum, si circa suam sententiam conprovinciales episcopi discordes

extiterint. Quinto, si in probatione deficit accusator, an cogendus sit reus ad probatione suae innocentiae?

**Sg translation:**

Certain fornicators also noted with infamy strive to accuse a certain bishop concerning simony. The defendant seeks out the court of the archbishop of another province, [but] in the end the accusers fall short in proof, [and] the defendant is summoned to the assertion of his innocence. It is asked therefore whether (*utrum*) [those] caught in a net by crime or (*seu*) noted with infamy, ought to be admitted to an accusation of this kind? Second, whether the bishop if he wished to turn back the accusation against them, whether it ought to be yielded from his plain assertion. Third, if it is allowed to him (*sibi*) to seek out the court of the archbishop of another province? Fourth, whose court ought to be sought out for him (*sibi*), if co-provincial bishops *extiterint* discords about his sentence? Fifth, if the accuser falls short in proof, whether the defendant ought to be summoned to proof of his innocence?

### Fr. text (95 words):

(Fr. 553) Duo fornicatores et infamia notati quendam religiosum<sup>8</sup> episcopum de symonia accusare nituntur; reus alterius prouinciae archiepiscopi iudicium expetit; tandem in probatione deficit accusator;<sup>9</sup> reus cogitur ad innocentiae suaे assertionem. (Qu. I.) Primo queritur, an crimine irretiti uel infamia notati ad huiusmodi accusationem sint admittendi? (Qu. II.) Secundo, si episcopus in eos accusationem uertere uoluerit, an simplici assertioni suaे sit fides habenda? (Qu. III.) Tertio, si liceat sibi expetere iudicium archiepiscopi alterius prouinciae? (Qu. IV.) Quarto, cuius iudicium sibi sit expetendum, si circa suam sententiam episcopos suaे prouinciae discordare contigerit? (Qu. V.) Quinto, si in probatione deficit accusator, an reus sit cogendus ad probationem innocentiae suaे?

---

<sup>8</sup> Does the adjective *religiosum* indicate that the accused bishop is a member of the regular as opposed to the secular clergy? What difference if any does this make with respect to the legal issues raised by the case statement?

<sup>9</sup> It is not clear why *accusator* is singular in Fr., especially since the logic of the case statement leads the reader to expect a plural reading. Verified as singular in Fr. (553), Ka (124vb), and Kb (117va).



### **Fr. translation:**

Two fornicators also noted with infamy strive to accuse a religious bishop concerning simony; the defendant seeks out the court of the archbishop of another province; [but] in the end the accuser falls short in proof; [and] the defendant is summoned to the assertion of his innocence. First it is asked, whether [those] caught in a net by crime or (*uel*) noted with infamy ought to be admitted to an accusation of this kind? Second, if the bishop wished to turn the accusation against them, whether confidence ought to be held from his plain assertion? Third, if it is allowed to him (*sibi*) to seek out the court of the archbishop of another province? Fourth, whose court ought to be sought out for him (*sibi*), if it happened that the bishops of his province disagree about his sentence? Fifth, if the accuser falls short in proof, whether the defendant ought to be summoned to the proof of his innocence?

### **Comparative Translation of C.7 d.init.**

### **Sg text:**

(79b) Episcopus quidam infirmitate longa detentus rogavit alium sibi substitui, precibus cuius pontifex romanus annuit, et quod rogavit ei concessit; postea convaluit episcopus idem, et quod prius fecerat rescindi cupit. Movet questionem contra qui succasserat. Reposcit (80a) cathedram



suam tamquam sibi debitam. Queritur ergo an episcopo vivente aliis in eadem possit ecclesia ordinari? Secundo, an cathedram iste reposcere valeat, quam aliter intercessione propria intercepit?

### Sg translation:

A certain bishop held down (*detentus*) by a long infirmity asked that another be substituted (*substitui*) for himself, through whose prayers the Roman pontiff approved, and what [the bishop] asked [the pope] granted to him; afterwards the same bishop recovered, and what he had previously done, he wishes to be annulled (*rescindi*). He moves a question against [him] who had succeeded (*succasserat*). He demands back his see as if a debt to himself. It is asked therefore whether (*an*), the bishop [still] living, someone can be ordained into the same church? Second, whether he is able to demand back (*reposcere*) a see which *intercepit* otherwise through his own intercession?

**Fr. text:**

(Fr. 566) Quidam longa inualetudine<sup>10</sup> grauatus episcopus alium sibi substitui rogauit, cuius precibus summus Pontifex annuit, et quod rogauerat ei concessit. Postea uero conualuit idem episcopus, et quod prius fecerat cupid rescindi; aduersus eum, qui sibi accesserat, questionem mouet, suam cathedram tamquam sibi debitam reposcit. Hic primum queritur, utrum uiuente episcopo aliis possit in ecclesia eadem ordinari? Secundo, an iste ualeat reposcere cathedram, quam sua intercessione alter accepit?

**Fr. translation:**

A bishop burdened (*grauatus*) by a long sickness (*inualetudine*) asked that another be substituted (*substitui*) for himself, through whose prayers the supreme pontiff approved, and what [the bishop] had asked [the pope] granted to him. But afterwards the same bishop recovered, and what he had previously done, he wishes to be annulled (*rescindi*); against him, who had acceded (*accesserat*) to him, he moves a question [and] demands back

<sup>10</sup> *inualetudine*: Fr. (566), Aa (156r), Bc (150ra), and Fd (36rb) read *inualitudine* or *invalidudine*. There is an OCR error (*inualetudinem*) in the MGH e-text. Ka (127va), Kb (120rb), and Sg (79b) read *infirmitate*. Check P.



his see as if a debt to himself. Here it is first asked whether (*utrum*), the bishop [still] living, someone can be ordained into the same church? Second, whether he is able to demand back (*reposcere*) a see which another received through his intercession?

*quidam*: translate *quidam* as a demonstrative adjective (“a certain X”) when it follows the noun it modifies; translate *quidam* as an article (“an X”) when it precedes the noun it modifies, as is the convention in modern European languages.

*substitui* and *rescindi*: 3rd conjugation present passive infinitives.

*possit* and *valeat*: translate *possit* as “he can”; translate *valeat* as “he is able to”.

*Movet questionem, questionem mouet*: legal t.t., reappears in C.14 d.init.

Square brackets enclose words supplied in the translation.

There are eight noun-adjective pairs in the text of C.7 d.init. In the Sg version of the text, the adjective follows the noun in all but one instance: *Episcopus quidam, infirmitate longa, pontifex romanus, episcopus idem, cathedram suam, episcopo vivente, eadem ... ecclesia*, and *intercessione propria*. In the Fr. version of the text, the adjective precedes the noun in all



but one instance: *Quidam ... episcopus, longa in ualeudine, summus Pontifex, idem episcopus, suam cathedram, uiuente episcopo, ecclesia eadem, and sua intercessione.*

## **Comparative Translation of C.11 d.init.**

### **Sg text:**

(92b) Adversus clericum clericus de prediis questionem agitavit, quem ad civilem iudicem producere voluit. Reus non nisi ante ecclesiasticum iudicem stare voluit; actor vero iudicis civilis potentia illum a possessione sua deiecit. Ille (93a) contempta episcopi sui sententia offitium suum administravit, hoc conperito episcopus sine spe restitucionis in eum sententiam dedit. Queritur ergo primum, an clericus ante civilem iudicem producendus sit? Secundo, si producendus non est, an hec culpa digna sit suspensione? Tercio, si digna non fuerit, an contemptorem sententiae sui episcopi irreparabiliter oporteat deponi?

### **Sg translation:**

A cleric stirred up (*agitavit*) a question concerning estates against [another] cleric whom he wished to bring before a civil judge. The defendant did not wish to stand [trial] except before an ecclesiastical judge; but the plaintiff (*actor*) drove him out from his



possession through the power of a civil judge. He, disregarding the sentence of his bishop, administered his office, [and] the bishop, when he found out about this, gave final judgement against him without hope of reinstatement. It is asked therefore first, whether a cleric ought to be brought before a civil judge? Second, if he ought not to be brought forth, whether this fault is worthy of suspension. Third, if it was not worthy, whether it is necessary that the one in contempt be deposed irreparably by (lit. of) the sentence of his bishop?

actor: this is the only occurrence in any form of the word actor in the Sg version of the text of the case statements. (The word actor also occurs exactly once in the Fr. version of the text of the case statements, in the same place.) The initiator or initiators of a legal action are far more frequently described in the case statements using forms of the word accusator (10 occurrences in Sg, and 18 occurrences in Fr.)

Ille contempta episcopi sui sententia offitium suum administravit, hoc comperto episcopus sine spe restitucionis in eum sententiam dedit: lit. He, the sentence of his bishop having been disregarded, administered his office, this thing having been found out, the bishop gave sentence against him without hope of restitution. The Sg version of



C.11 d.init. omits a crucial intermediate step from the sequence of events in the narrative. In the Fr. version of the narrative, when the plaintiff's bishop learns of his having litigated his claim against a clerical defendant before a civil judge, he suspends the plaintiff from the exercise of his ecclesiastical duties. This is the sentence the plaintiff disregards.

**Fr. text:**

(Fr. 626) Clericus aduersus clericum questionem de prediis agitauit, quem ad ciuilem iudicem producere uoluit reus non nisi ante iudicem ecclesiasticum stare uolebat; actor uero potentia ciuilis iudicis illum a possessione sua deiecit. Quo audito episcopus eum ab offitio suspendit; ille contempta episcopi sui sententia offitium suum administravit. Hoc conperito episcopus sine spe restitutionis in eum sententiam dedit. (Qu. I.) Hic primum queritur, utrum clericus ante ciuilem iudicem sit producendus? (Qu. II.) Secundo, si producendus non est, an hec culpa sit digna suspensione? (Qu. III.) Tertio, si digna non fuerit, an contemptorem sententiae sui episcopi inreparabiliter oporteat deponi?

**Fr. translation:****Comparative Translation of C.19 d.init.****Sg text:**

(144b) Volunt duo clericci ad monasterium transire. Petiit uterque ab episcopo suo licentiam. Unde queritur, an debeat episcopus permittere, ut propria ecclesia relicta, clericus monachus fiat? Secundo, si episcopus dare licentiam noluerit, an eo invito monasterium adire possit? Tertio, si contigerit ipsos regulares canonicos fuisse, utrum monasterii ingressus eis esset concedendus?

**Sg translation:**

Two clerics want to transfer to a monastery. Each sought permission from his bishop. Whence it is asked whether a bishop ought to permit a cleric to become a monk after leaving his own church? Second, if the bishop did not wish to give permission, whether the cleric can enter a monastery against his wishes? Third, if it happened that they had been canons regular, whether entry to a monastery should have been granted to them.

(Sg)

**Fr. text:**

(Fr. 839) Duo clerici ad monasterium transire uolunt; uterque licentiam ab episcopo suo petiit; unus relict a propria ecclesia eo inuit o, alter dimissa regulari canonica cenobio se contulit. Modo queritur, si episcopus debeat permittere, ut relict a propria ecclesia clericus monasterium ingrediatur? Secundo queritur, si episcopus licentiam dare noluerit, an eo inuit o monasterium possit adire? Tertio, si contigerit ipsos regulares canonicos fuisse, utrum concedendus esset eis monasterii ingressus?

**Fr. translation:**

Two clerics want to transfer to a monastery; each sought permission from his bishop; one after leaving his own church against the wishes of his bishop, the other *dimissa regulari canonica* betook himself to a monastery. Now it is asked if a bishop ought to permit a cleric to enter a monastery after leaving his own church? Second it is asked, if the bishop did not wish to give permission, whether the cleric can enter a monastery against his wishes? Third, if it happened that they had been canons regular, whether entry to a monastery should have been granted to them?

## Incipits to first questions

### Sg

[Fully one-third—11 out of 33—of the incipits to first questions in the Sg version of the case statements are unique. The Friedberg (vulgate) and Sg versions of incipits to first questions are identical in only three out of 32 directly comparable case statements (those for C.30 d.init., C.32 d.init., and C.33 d.init.).]

Primum ergo queritur (4), Queritur ergo (4), Unde queritur (4), Queritur igitur (3), Unde primo queritur (3), Queritur igitur primum (2), Queritur (2), Hic primum queritur (1), Hinc queritur (1), Primum igitur queritur (1), Queritur ergo primo (1), Queritur ergo primum (1), Unde modo queritur (1), Unde primo dubitatur (1), Unde primum queritur (1), Unde prius queritur (1), Unde queri potuit (1), Unde questio primum proponitur (1)

### Fr.

Hic primum queritur (15), Queritur (8), Modo primum queritur (3), Nunc primum queritur (3), Primo queritur (2), Primum queritur (2), Hic primo queritur (1), Modo queritur (1), Queritur autem (1)



	Sg	Fr.
C.Prima	Queritur igitur	
C.1	Primum ergo queritur	Hic primum queritur
C.2	Primum igitur queritur	Hic primum queritur
C.3	Queritur igitur primum	Hic primum queritur
C.4	Queritur ergo primo	Hic primum queritur
C.5	Unde primo dubitatur	Primo queritur
C.6	Queritur ergo	Primo queritur
C.7	Queritur ergo	Hic primum queritur
C.8	Unde questio primum proponitur	Hic primo queritur
C.9	Unde queri potuit	Queritur
C.10	Primum ergo queritur	Modo primum queritur
C.11	Queritur ergo primum	Hic primum queritur
C.12	Queritur igitur primum	Modo primum queritur
C.13	Primum ergo queritur	Modo primum queritur
C.14	Unde primum queritur	Queritur

C.15	Unde primo queritur	Queritur autem
C.16	Primum ergo queritur	Hic primum queritur
C.17	Queritur ergo	Hic primum queritur
C.18	Unde queritur	Queritur
C.19	Unde queritur	Modo queritur
C.20	Queritur igitur	Nunc primum queritur
C.21	Unde primo queritur	Queritur
C.22	Queritur ergo	Primum queritur
C.23	Queritur igitur	Hic primum queritur
C.24		Hic primum queritur
C.25		Nunc primum queritur
C.26		Primum queritur
C.27	Unde prius queritur	Hic primum queritur
C.28		Hic primum queritur
C.29	Unde queritur	Hic primum queritur
C.30	Queritur	Queritur

C.31	Unde modo queritur	Queritur
C.32	Hic primum queritur	Hic primum queritur
C.33	Queritur	Queritur
C.34	Unde primo queritur	Nunc primum queritur
C.35	Unde queritur	Hic primum queritur
C.36	Hinc queritur	Queritur

# Appendix 1

## dicta.py

```
#!/usr/local/bin/python3
#
# Paul Evans (10evans@cua.edu)
# 25 Oct - 8 Nov 2015
# 17 Oct - 24 Oct 2013
#
# import re
# import sys
def main():
    f = open('./edF.txt', 'r')
    file = f.read()
    toc = open('./toc_all.txt', 'r')
    dictionary_Fr = {} # Friedberg
    dictionary_1r = {} # first recension
    dictionary_2r = {} # second recension
    # (?<=...) positive Lookbehind assertion.
    dicta = re.findall('(?:<\T [AP]\>|(?<=\T [AP]\>)(.*?))' # dictum starts with dictum ante or dictum
post tag.
    (? : # non-capturing group.
        '\<1 [CD][CP]?\>' # dictum ends with major division,
        '\<2 \d{1,3}\>' # or number of major division,
        '\<3 \d{1,2}\>' # or number of question,
        '\<4 \d{1,3}\>' # or number of canon,
        '\<P 1\>' # or Palea,
        '\<T [AIPRT]\>' # or inscription or text tag.
    ')', file, re.S) # re.S (re.DOTALL) makes '.', special character match any character
including newline.
print('expected 1277 dicta, found ' + str(len(dicta)) + ' dicta', file=sys.stderr)
```



```

for dictum in dicta:
    dictum = re.sub('\'<S \d{1,4}\>' ><L 1\> \-' \d{1,4}\'+', '', dictum) # remove page and line number tags.
    dictum = re.sub('\'<P 1\> \'-\\'[PALEA]\.\'+', '', dictum) # remove Palea tags.
    dictum = re.sub('\'-.*?\'+', '', dictum)
    dictum = re.sub(re.compile('\'-\'.*\?\'+'), re.S), '', dictum)
    dictum = re.sub('\'\s+', '', dictum)
    dictum = re.sub('\'^\s+', '', dictum) # remove Leading whitespace characters
    dictum = re.sub('\'\s+$', '', dictum) # remove trailing whitespace characters

key = toc.readline().rstrip()
if key in dictionary_Fr:
    # if there's already a dictionary entry with this key, merge the entries
    # print('duplicate key: ' + key, file=sys.stderr)
    dictum = dictionary_Fr[key] + ' ' + dictum
    dictionary_Fr[key] = dictum

keys = tuple(open('./toc_1r.txt', 'r'))
for key in keys:
    key = key.rstrip()
    dictionary_1r[key] = dictionary_Fr[key] # copy dictum from Friedberg dictionary into first-
recension dictionary

keys = tuple(open('./toc_2r.txt', 'r'))
for key in keys:
    key = key.rstrip()
    dictionary_2r[key] = dictionary_Fr[key] # copy dictum from Friedberg dictionary into second-
recension dictionary

keysandpatterns = [
    {'key': 'D.25 d.p.c.1', 'pattern': '(Ex hac epistola liquet, quid cuiusque officii sit\.)'},
    {'key': 'D.25 d.p.c.3', 'pattern': '(Nunc autem per.*?mentem eius remordeat\.)'},
    {'key': 'D.26 d.p.c.4', 'pattern': '(Iohannes etiam Baptista.*?alterm habuisse probantur\.)'},
    {'key': 'D.30 d.a.c.1', 'pattern': '(Illud autem, quod.*?que coniugium detestabatur,)'},
    {'key': 'D.31 d.p.c.11', 'pattern': '(Ut igitur ex.*?reddere non ualent. Sed obicitur illud

```



```

Tripartitae ystoriae:'), },
{'key': 'D.42 d.p.c.1', 'pattern': '(Hinc etiam Iohannes.*?de ecclesia eiciebat\.)' },
{'key': 'D.45 d.p.c.17', 'pattern': '(Hinc etiam alibi.*?uo patrem exhibeat.")' },
{'key': 'D.47 d.p.c.8', 'pattern': '(Necesse est etiam.*?sollicitam diligentiam exhibebit\?)' },
{'key': 'D.54 d.p.c.23', 'pattern': '(Ecce, quomodo servi.*?quomodo non admittantur\.)' },
{'key': 'D.63 d.p.c.28', 'pattern': '(Verum, quia imperatores.*?anathematis vinculo
innodaretur,)(.*?)(Postremo presentibus legatis.*?ecclesiae Dei conferentes\.)' },
{'key': 'D.63 d.p.c.34', 'pattern': '(Ex his constitutionibus.*?habita constitutum est\.)' },
{'key': 'D.68 d.p.c.2', 'pattern': '(Quod ergo consecratus.*?ad cautelam salutis\.)' },
{'key': 'C.1 q.1 d.p.c.51', 'pattern': '(Sed notandum est.*?Non sanat baptismus perfidorum,
etc.' )},
{'key': 'C.1 q.1 d.p.c.123', 'pattern': '(Quolibet ergo munere.*?falsa diiudicatur ordinatio\.)' },
{'key': 'C.1 q.4 d.p.c.9', 'pattern': '(Cum ergo de baptizatis.*?impedit nomen erroris\.)' },
{'key': 'C.1 q.4 d.p.c.12', 'pattern': '(Ignorabat autem Petrus.*?permittitur ignorare, alii
non\.)' },
# {'key': 'C.2 q.3 d.p.c.7', 'pattern': '(Notandum quoque est.*?quod obiecerat desistat\.)' }, # @.2
{'key': 'C.2 q.3 d.p.c.7', 'pattern': '(Notandum quoque est.*?in Libro Capitulorum:' ), # @.2
{'key': 'C.2 q.6 d.p.c.31', 'pattern': '(Forma uero appellationis.*?in scriptis fieri debent\.)' },
{'key': 'C.2 q.6 d.p.c.39', 'pattern': '(Cum autem in.*?suam agere oportet\.)' },
{'key': 'C.2 q.7 d.p.c.40', 'pattern': '(Cum ergo Petrus.*?suscipere reprehensionem
subditorum\.)' },
{'key': 'C.2 q.8 d.p.c.5', 'pattern': '(Sed Calixtus Papa.*?per epistolam accusare audeat\.)' },
{'key': 'C.3 q.1 d.p.c.6', 'pattern': '(Patet ergo, quod.*?quam uocentur ad causam\.)' },
{'key': 'C.3 q.11 d.p.c.3', 'pattern': '(Hoc autem intelligendum.*?auctoritatibus non
prohibetur\.)' },
# {'key': 'C.4 q.2 d.p.c.3', 'pattern': '(Sed obicitur illud.*?humanae actionis trahenda\.)' },
{'key': 'C.5 q.3 d.p.c.1', 'pattern': '(Ecce episcopus.*?se agere licet\.)' },
{'key': 'C.6 q.1 d.p.c.21', 'pattern': '(Verum hoc Augustini.*?accusatione ipse repellit\.)' },
{'key': 'C.11 q.1 d.p.c.34', 'pattern': '(Non ait propter.*?quam criminalem intelligens\.)' },
{'key': 'C.11 q.3 d.p.c.40', 'pattern': '(Premissis auctoritatibus, quibus.*?in se excepert\.)' },
{'key': 'C.13 q.1 d.p.c.1', 'pattern': '(In diocesi autem.*?qui secum erant\.)(.*?)(Quia ergo
nos.*?ad diocesianum transferre\.)' },
{'key': 'C.13 q.2 d.p.c.3', 'pattern': '(Item Ioseph, moriens.*?eo sepultus est\.)' },
igitur istorum.*?uoluntate tumulandi consistit\.)' },

```



```
{
  'key': 'C.13 q.2 d.p.c.8',
  'pattern': '(Hac nimirum auctoritate.*?quam prohibetur transscendere\.)'},
  {'key': 'C.14 q.1 d.p.c.1',
  'pattern': '(Quia ergo generaliter.*?prohibentur stare coram iudice\.)'},
  # 'stare coram iudice' occurs twice
  {'key': 'C.14 q.2 d.p.c.1',
  'pattern': '(Potest etiam intelligi.*?pauperum, testimonium dicant\.)'},
  {'key': 'C.14 q.5 d.p.c.14',
  'pattern': '(Sed hoc multipliciter.*?bonum possunt conuerti\.)'},
  {'key': 'C.15 q.1 d.p.c.3',
  'pattern': '(Ex eo autem.*?penam aut gloriam.")'},
  {'key': 'C.15 q.1 d.p.c.11',
  'pattern': '(Cum itaque qui.*?Obicitur autem)'},
  {'key': 'C.15 q.1 d.p.c.12',
  'pattern': '(Sunt quedam, que.*?muneris executionem impediunt\.)'},
  {'key': 'C.16 q.1 d.p.c.40',
  'pattern': '(.*?)(Ostendit ergo Ieronimus.*?ipsum imperfectis connumerans\.) (. *?)(Ecce sufficienter monstratum.*?assecurunt potestatis executionem\.).(*?$)'},
  # 'imperfectis' occurs twice
  {'key': 'C.16 q.1 d.p.c.47',
  'pattern': '(Quod autem dicitur.*?duos potest diuidi)'},
  {'key': 'C.16 q.1 d.p.c.53',
  'pattern': '(Sicut duo episcopatus.*?ad paucitatem redigeretur\.)'),
  {'key': 'C.16 q.3 d.p.c.15',
  'pattern': '(Potest etiam aliter.*?obici non potest\.)'},
  {'key': 'C.16 q.3 d.p.c.16',
  'pattern': '(Sed sola prescriptione.*?spatio prescribi possunt\.)'},
  {'key': 'C.17 q.2 d.p.c.2',
  'pattern': '(Ecce iste se.*?conceptit, et ore pronunciauit\.)'}, # 'et' occurs twice
  {'key': 'C.21 q.2 d.p.c.3',
  'pattern': '(Sed aliud est.*?omnibus modis prohibetur\.)'},
  {'key': 'C.22 q.1 d.p.c.16',
  'pattern': '(Sic etiam cum.*?creatorem iurat mendaciter\.)'},
  {'key': 'C.22 q.2 d.p.c.5',
  'pattern': '(Ille ergo falsum.*?esse quod iurat\.)'},
  {'key': 'C.23 q.4 d.p.c.26',
  'pattern': '(Potest in hac.*?personae quendam excommunicauerat,)'},
  {'key': 'C.23 q.4 d.p.c.27',
  'pattern': '(ostendens, quod peccata.*?potius dissimulanda sunt)'},
  {'key': 'C.23 q.4 d.p.c.30',
  'pattern': '(Quod autem peccatum.*?patiente tollerasse asseritur\.)'},
  {'key': 'C.23 q.8 d.p.c.25',
  'pattern': '(Hinc datur intelligi.*?Pontificis fieri debet\.)'},
  {'key': 'C.23 q.8 d.p.c.27',
  'pattern': '(Reprehenduntur ergo Gallicani.*?orationibus Deo commendent\.)'},
  {'key': 'C.29 q.1 d.a.c.1',
  'pattern': '(Quod autem coniugium.*?potest eam dimittere,)'},
  {'key': 'C.29 q.2 d.p.c.6',
  'pattern': '(Cum dicitur: "sciens.*?fraude decepta est;")'},
  {'key': 'C.30 q.4 d.p.c.5',
  'pattern': '(Notandum uero est.*?uiro suo cognoscitur\.)'},
  {'key': 'C.31 q.1 d.p.c.7',
  'pattern': '(Sed obicitur: Dauid.*?quam significacione futurorum\.)'},
  {'key': 'C.32 q.1 d.p.c.10',
  'pattern': '(Si ergo, ut.*?sed adulteri appellantur\.)'}
}
```



```

{'key': 'C.32 q.4 d.p.c.10', 'pattern': '(Ecce, quod nullo.*?nomine iudicantur indigni\.)'},
{'key': 'C.33 q.2 d.p.c.9', 'pattern': '(In premissis auctoritatibus.*?eis misericordia
inpendatur\.)'},
{'key': 'de Pen. D.1 d.a.c.1', 'pattern': '(Utrum sola cordis.*?promererি, iuxta illud)'}, #
d.a.c.1
{'key': 'de Pen. D.1 d.p.c.87', 'pattern': '(His auctoritatibus asseritur.*?iugiter confiteri
deberimus\.).*(.?) (Similiter et illud.*?de penitencia ait:)'},
{'key': 'de Pen. D.2 d.a.c.1', 'pattern': '(Alii dicunt penitenciam.*?tibi aliquid contingat." )'},
{'key': 'de Pen. D.2 d.p.c.24', 'pattern': '(Hec itaque karitas.*?redeunt et cetera." )'},
{'key': 'C.35 q.2 d.p.c.21', 'pattern': '(Hac auctoritate dum.*?ducat in uxorem." )'}, # ducat
]

for i in range (len(keysandpatterns)):
    key = keysandpatterns[i]['key']
    pattern = keysandpatterns[i]['pattern']
    result = re.search(pattern, dictionary_Fr[key])
    if result:
        if len(result.groups()) == 1:
            dictionary_1r[key] = fixString(result.group(1))
            dictionary_2r[key] = fixString(re.sub(pattern, '', dictionary_2r[key]))
        elif len(result.groups()) == 3:
            dictionary_1r[key] = fixString(result.group(1)) + ' ' + fixString(result.group(3))
            dictionary_2r[key] = fixString(result.group(2))
        elif len(result.groups()) == 5: # C.16 q.1 d.p.c.40
            dictionary_1r[key] = fixString(result.group(2)) + ' ' + fixString(result.group(4))
            dictionary_2r[key] = fixString(result.group(1)) + ' ' + fixString(result.group(3)) + ' ' +
fixString(result.group(5))

    else:
        print('no match: ' + key + '\n' + dictionary_Fr[key], file=sys.stderr)

# insert
key = 'C.3 q.1 d.p.c.2'
dictionary_1r[key] = ...Sed notandum est quod restitutio alia fit per presentiam iudicis, ueluti cum

```



dicitur a iudice: "Censeo te in integrum restituendum", qua restitutione animo tantum, non corpore possessio recipitur. Alia fit per executorem iudicis quando restitutus corporalem recipit possessionem. Queritur ergo que harum concedatur expoliatis, an illa tantum, que fit per sententiam iudicis, an illa etiam que fit per executorem sententiae, qua expoliatis presentialiter omnia reciduntur. Hec ultima expoliatis prestanta est. . .

```

# append
key = 'C.3 q.1 d.p.c.6'
dictionary_1r[key] = dictionary_1r[key] + ' His ita respondetur. Si uicum electionis ecclesie notum fuerit et ideo reprobati fuerint et si aliqua uiolentia in sedibus illis irrepsert electi restitutionem postulare non possunt. Si autem ecclesia eos per pacientiam tolerare voluerit et eis gradum honoris concesserit et si uiciosa fuerit eorum electio, tamen post electionem restituendi sunt, ante regularem ad synodi uocationem. . .'

# special fix
key = 'C.15 q.1 d.p.c.11'
dictionary_1r[key] = dictionary_1r[key][0:-1] + ':'

# interpolate
key = 'C.15 q.3 d.p.c.4'
pattern = '(Cum autem sacris.*?hoc non inferetur\.).*?)(Quamuis igitur sacris.*?credi non oportet\.).*?$)'
result = re.search(pattern, dictionary_1r[key])

if result:
    dictionary_1r[key] = fixString(result.group(1)) + ' Quecumque enim persone humanis legibus copulari prohibentur et diuinis, non omnium copula a sacris canonibus admittitur, quorum conuentio legibus imperatorum indulgetur. . . + fixString(result.group(3))
    dictionary_2r[key] = fixString(result.group(2)) + ' + fixString(result.group(4))

else:
    print('no match: ' + key + '\n' + dictionary_1r[key], file=sys.stderr)

# insert
key = 'C.21 q.3 d.a.c.1'
dictionary_1r[key] = 'Quod autem clerici secularium negotiorum procuratores esse non ualeant auctoritate Calcedonensis synodi probatur in qua sic statutum est legitur: . . .

# append
key = 'C.23 q.8 d.p.c.25'
dictionary_1r[key] = dictionary_1r[key] + ' Unde in quodam concilio statutum est ut episcopi non

```

```

proficiuntur ad comitatum nisi formatas ab apostolico acceperint.''
# append
key = 'de Pen. D.1 d.a.c.1'
dictionary_1r['de Pen. D.1 d.a.c.1'] = dictionary_1r[key].rstrip('.') + ' Leonis page:'''

all = open('./Gratian1.txt', 'w')
keys = tuple(open('./toc_1r.txt', 'r'))
for key in keys:
    key = key.rstrip()
    outfilename = './1r/' + key + '.txt'
    each = open(outfilename, 'w')
    each.write(dictionary_1r[key] + '\n')
    all.write(dictionary_1r[key] + '\n')
    each.close()
all.close()

all = open('./Gratian2.txt', 'w')
keys = tuple(open('./toc_2r.txt', 'r'))
for key in keys:
    key = key.rstrip()
    outfilename = './2r/' + key + '.txt'
    each = open(outfilename, 'w')
    each.write(dictionary_2r[key] + '\n')
    all.write(dictionary_2r[key] + '\n')
    each.close()
all.close()

def fixString(string):
    string = re.sub('\s+', ' ', string) # 2r
    string = re.sub('^\s+', ' ', string) # 2r
    string = re.sub('\s+$', ' ', string) # 2r
    if string[-1] == ' ' or string[-1] == ';':
        string = string[0:-1] + '.'

    string = string[0:-1] + '.'
```



```
if string[-1].isalpha():
    string = string + '_'
return string

if __name__ == '__main__':
    main()
```

## Appendix 2

### main.py

```
#!/usr/local/bin/python3
#
# Paul Evans (10evans@cua.edu)
# 8 February 2015 -
# 12 February 2015
#
# import re
# import sys
# import parse
def main():
    file = open('./edf.txt', 'r').read()
    decretum = parse.parse_all(preprocess(file))
    traverse(decretum[0])
    traverse(decretum[1])
    traverse(decretum[2])

def traverse(tree):
    for i in range(len(tree[1])):
        subtree = tree[1][i]
        if isinstance(subtree[1], list):
            tag = subtree[0]
            print(tag)
            traverse(subtree)
        elif isinstance(subtree[1], str):
            tag = subtree[0]
            text = subtree[1]
            print(tag, text)
    return
```



```

def preprocess(text):
    text = re.sub(re.compile('`-.*?\'+', re.S), '' , text) # remove comments
    text = re.sub(`\`S \d{1,4}\`', '' , text) # remove page number tags
    text = re.sub(`\`L \d{1,2}\`', '' , text) # remove Line number tags
    text = re.sub(`\`P 1\`|\`P 0\`', '' , text) # remove Palea tags
    text = re.sub(`\`$+', '' , text) # remove multiple whitespace characters
    text = re.sub(`\$+', '' , text) # remove trailing whitespace characters
    return(text)

if __name__ == '__main__':
    main()

parse.py
#!/usr/local/bin/python3
#
# Paul Evans (10evans@cua.edu)
# 23 January 2015 -
# 12 February 2015
#
import re
import sys
def parse_all(text):
part_list = []
m = re.search(`(\`1 D\`.*?)(\`1 C\`.*?)(\`1 DC\`.*?$', text, re.S)
part_list.append(`\`1 D\`', parse_part_1(m.group(1)))
part_list.append(`\`1 C\`', parse_part_2(m.group(2)))
part_list.append(`\`1 DC\`', parse_part_3(m.group(3)))
return(part_list)

# D.1-101
def parse_part_1(text):
    distinction_list = []

```



```

distinctions = re.findall('(?:<1 D\\>)(.*?)(?=\\<1 D\\>|$)', text)
for distinction in distinctions:
    distinction = distinction.strip(' ')
    m = re.match('(?:<2 \\d{1,3}\\>) (\\<T A\\>) (.*) (\\<4 1\\>.*?)$', distinction)
    tag = m.group(1)
    node = (m.group(2), m.group(3)) # d.a.c.1 tag-text tuple
    canon_list = parse_canons(m.group(4))
    canon_list.insert(0, node)
    distinction_list.append((tag, canon_list))

# C.1-36
def parse_part_2(text):
    case_list = []
    cases = re.findall('(?:<1 C\\>)(.*?)(?=\\<1 C\\>|$)', text)
    for case in cases:
        case = case.strip(' ')
        m = re.match('(?:<2 \\d{1,2}\\>) (\\<T Q\\>) (.*) (\\<3 1\\>.*?)$', case)
        tag = m.group(1)
        node = (m.group(2), m.group(3)) # d.init. tag-text tuple
        question_list = parse_questions(m.group(4))
        question_list.insert(0, node)
        case_list.append((tag, question_list))
    return(case_list)

# de Consecratione
def parse_part_3(text):
    distinction_list = []
    distinctions = re.findall('(?:<1 DC\\>)(.*?)(?=\\<1 DC\\>|$)', text)
    for distinction in distinctions:
        distinction = distinction.strip(' ')
        m = re.match('(?:<2 \\d\\>) (\\<4 1\\>.*?)$', distinction)
        tag = m.group(1)

```

```

canon_list = parse_canons(m.group(2))
distinction_list.append((tag, canon_list))
return(distinction_list)

def parse_questions(text):
    question_list = []
    questions = re.findall('(\<3 \d{1,2}\>.*?)(?=\<3 \d{1,2}\>|\$)', text)
    for question in questions:
        question = question.strip(' ')
        m0 = re.match('(\<3 \d{1,2}\>) (\<T A\>) (.*)? (\<1 DP\>.*?)$', question) # C.33 q.3 (de Pen.)
        m1 = re.match('(\<3 \d{1,2}\>) (\<T A\>) (.*)? (\<4 1\>.*?)$', question)
        m2 = re.match('(\<3 \d{1,2}\>) (\<T A\>) (.*)?$', question) # C.11 q.2, C.17 q.3, C.22 q.3, C.29
        q.1
        if m0:
            tag = m0.group(1)
            node = (m0.group(2), m0.group(3)) # d.a.c.1 tag-text tuple
            distinction_list = parse_de_pen(m0.group(4))
            question_list.append((tag, [node, ('<1 DP>', distinction_list)]))
        elif m1:
            tag = m1.group(1)
            node = (m1.group(2), m1.group(3)) # d.a.c.1 tag-text tuple
            canon_list = parse_canons(m1.group(4))
            canon_list.insert(0, node)
            question_list.append((tag, canon_list))
        elif m2:
            tag = m2.group(1)
            node = (m2.group(2), m2.group(3)) # d.a.c.1 tag-text tuple
            question_list.append((tag, [node]))
    return(question_list)

# de Penitentia
def parse_de_pen(text):
    distinction_list = []

```



```

distinctions = re.findall('(?:<1 DP\\>)(.*?)(?=\\<1 DP\\>|$)', text)
for distinction in distinctions:
    distinction = distinction.strip(' ')
    m = re.match('(<2 \\d\\>) (<T A\\>) (.*) (\<4 1\\>.*?)$', distinction)
    tag = m.group(1)
    node = (m.group(2), m.group(3)) # d.a.c.1 tag-text tuple
    canon_list = parse_canons(m.group(4))
    canon_list.insert(0, node)
    distinction_list.append((tag, canon_list))
return(distinction_list)

# return List of canons
def parse_canons(text):
    canon_list = []
    canons = re.findall('(<4 \\d{1,3}\\>.*?)(?=\\<4 \\d{1,3}\\>|$)', text)
    for canon in canons:
        canon = canon.strip(' ')
        m = re.match('(<4 \\d{1,3}\\>) (.*)$', canon)
        if m:
            nodes = parse_nodes(m.group(2))
            else: # C.1 q.4 c.6
                m = re.match('(<4 \\d{1,3}\\>)$', canon)
                nodes = []
                canon_list.append((m.group(1), nodes))
    return(canon_list)

# return List of terminal nodes (tag-text tuples)
def parse_nodes(text):
    node_list = []
    nodes = re.findall('(<T [AIPRT]\\>.*?)(?=\\<T [AIPRT]\\>|$)', text)
    for node in nodes:
        node = node.strip(' ')
        m = re.match('(<T [AIPRT]\\>) (.*)$', node)

```



```
    node_list.append((m.group(1), m.group(2)))
return(node_list)

if __name__ == '__main__':
    main()
```

## Appendix 3

Citation	Column	Reported by	Date	Error	Correction
D.6 d.p.c.3	11	Anders Winroth	6-Oct-19	quantam ad moralem intelligentiam	quantum ad moralem intelligentiam
D.23 c.2	79	Anders Winroth	23-Aug-19	carnis passione, mortuus uera corporis sui morte, resurrexit uera	
D.54 d.p.c.22	214	Paul Evans	6-Apr-13	Quid autem serui ecclesiarum contra sententiam restitutionem inpetrauerit	Quod autem serui ecclesiarum contra sententiam restitutionem inpetrauerit
C.2 q.6 c.41	483	Paul Evans	13-Jan-19	a communione ait remouendus in episcoporum iudicio	a communione sit remouendus in episcoporum iudicio
C.4 d.init.	536	Paul Evans	1-Jul-19	Quidam longa inualetudine grauiatus episcopus	Quidam longa inualetudine grauiatus episcopus
C.4 d.init.	536	Paul Evans	1-Jul-19		
C.7 d.init.	566	Paul Evans	8-Jul-19		

## Bibliography

- Adair, Douglass. "The Authorship of the Disputed Federalist Papers." *The William and Mary Quarterly* 1, no. 2 (1944): 98–122.
- — —. "The Authorship of the Disputed Federalist Papers: Part II." *The William and Mary Quarterly* 1, no. 3 (1944): 235–64.
- Appelt, Heinrich, ed. *Die Urkunden Der Deutschen Könige Und Kaiser*. Vols. X, 2. Monumenta Germaniae Historica. Diplomata Regum et Imperatorum Germaniae. Berlin: Weidmannsche Verlagsbuchhandlung, 1979.
- Bellomo, Manlio. *The Common Legal Past of Europe: 1000-1800. Studies in Medieval and Early Modern Canon Law*, v. 4. Washington, D.C: Catholic University of America Press, 1995.
- Beyer, Alfred. *Lokale Abbreviationen Des Decretum Gratiani: Analyse Und Vergleich Der Dekretabbreviationen "Omnes Leges Aut Divine" (Bamberg), "Humanum Genus Duobus Regitur" (Pommersfelden) Und "de His Qui Intra Clastra Monasterii Consistunt"* (Lichtenthal, Baden-Baden). Bamberger Theologische Studien ; Bd. 6. Frankfurt am Main ; PLang, 1998.
- Brasington, Bruce Clark, ed. *Order in the Court: Medieval Procedural Treatises in Translation*. Medieval Law and Its Practice 21. Leiden: Brill, 2016.
- — —, ed. *Ways of Mercy: The Prologue of Ivo of Chartres ; Edition and Analysis*. Vita Regularis, Bd. 2. Münster : Piscataway, N.J: LIT ; Distributed in North America by Transaction Publishers, 2004.
- Brundage, James A. "The Teaching and Study of Canon Law in the Law Schools." In *The History of Medieval Canon Law in the Classical Period, 1140-1234: From Gratian to the Decretals of Pope Gregory IX*, edited by Wilfried Hartmann and Kenneth Pennington, 98–120. History of Medieval Canon Law. Washington, D.C: Catholic University of America Press, 2008.

- Burrows, John. "'Delta': A Measure of Stylistic Difference and a Guide to Likely Authorship." *Literary and Linguistic Computing* 17, no. 3 (September 2002): 267–87.
- — —. "Questions of Authorship: Attribution and Beyond: A Lecture Delivered on the Occasion of the Roberto Busa Award ACH-ALLC 2001, New York." *Computers and the Humanities* 37, no. 1 (February 2003): 5–32.
- Campbell, Gordon, Thomas N. Corns, John K. Hale, David I. Holmes, and Fiona J. Tweedie. "The Provenance of "de Doctrina Christiana" ." *Milton Quarterly* 31, no. 3 (1997): 67–121.
- Campbell, Gordon, Thomas N. Corns, John K. Hale, and Fiona J. Tweedie. *Milton and the Manuscript of de Doctrina Christiana*. Oxford ; New York: Oxford University Press, 2007.
- Chodorow, Stanley. *Christian Political Theory and Church Politics in the Mid-Twelfth Century; the Ecclesiology of Gratian's Decretum*. Publications of the Center for Medieval and Renaissance Studies, U.C.L.A., 5. Berkeley: University of California Press, 1972.
- — —. "Review of the Making of Gratian's Decretum by Anders Winroth." *The English Historical Review* 118, no. 475 (February 2003): 174–76.
- Christensen, Katherine. "Introduction." In *The Treatise on Laws: (Decretum DD. 1-20)*, ix–xxvii. Studies in Medieval and Early Modern Canon Law, v. 2. Washington, D.C.: Catholic University of America Press, 1993.
- Corbino, Alessandro, and Bernardo Santalucia, eds. *Justiniani Augusti Pandectarum Codex Florentinus*. Firenze: Olschki, 1988.
- Craig, Hugh. "Stylistic Analysis and Authorship Studies." In *A Companion to Digital Humanities*, edited by Susan Schreibman, Raymond George Siemens, and John Unsworth, 273–88. Blackwell Companions to Literature and Culture 26. Malden, MA: Blackwell Pub, 2004.



De Jong, Mayke. *The Penitential State: Authority and Atonement in the Age of Louis the Pious*, 814–840. Cambridge, UK ; New York: Cambridge University Press, 2009.

Donahue, Charles. "On Translating the "Digest"." *Stanford Law Review* 39, no. 4 (1987): 1057–77.

Dufour, Jean, Gérard Giordanegno, and André Gouron. "L'attrait Des 'Leges': Note Sur La Lettre d'un Moine Victorin (Vers 1124/1127)." *Studia et Documenta Historiae et Iuris* 45 (1979): 504–29.

Duvernoy, Jean, ed. *Le Registre d'inquisition de Jacques Fournier, évêque de Pamiers* (1318–1325). Bibliothèque Méridionale. 2. Ser, t. 41. Toulouse: É. Privat, 1965.

Eder, Maciej. "Does Size Matter? Authorship Attribution, Small Samples, Big Problem." *Literary and Linguistic Computing* 30, no. 2 (June 2015): 167–82.

Eder, Maciej, Mike Kestemont, and Jan Rybicki. "Stylometry with R: A Suite of Tools." In *Digital Humanities 2013: Conference Abstracts*, 487–89. Lincoln, NE: University of Nebraska–Lincoln, 2013. <http://dh2013.unl.edu/abstracts/ab-136.html>.

—. "Stylometry with R: A Suite of Tools." In *Digital Humanities 2013: Conference Abstracts*, 487–89. Lincoln, NE: University of Nebraska–Lincoln, 2013. <http://dh2013.unl.edu/abstracts/>.

Eichbauer, Melodie H. "Gratian's Decretum and the Changing Historiographical Landscape." *History Compass* 11, no. 12 (December 2013): 1111–25.

Eichbauer, Melodie Harris. "St. Gall Stiftsbibliothek 673 and the Early Redactions of Gratian's Decretum." *Bulletin of Medieval Canon Law* 27 (2007): 105–39.

Evert, Stefan, Thomas Proisl, Fotis Jannidis, Isabella Reger, Steffen Pielström, Christof Schöch, and Thorsten Vitt. "Understanding and Explaining Delta Measures for Authorship Attribution." *Digital Scholarship in the Humanities* 32, no. suppl\_2 (December 2017): ii4–ii16.



Fowler-Magerl, Linda. *Clavis Canonum: Selected Canon Law Collections Before 1140: Access with Data Processing*. Monumenta Germaniae Historica. Hilfsmittel 21. Hannover: Hahnsche, 2005.

Fuhrmann, Horst. *Einfluß Und Verbreitung Der Pseudoisidorischen Fälschungen : Von Ihrem Auftauchen Bis in d. Neuere Zeit*. Schriften Der Monumenta Germaniae Historica, Deutsches Institut Für Erforschung Des Mittelalters ; Bd. 24. Stuttgart: Hiersemann, 1972.

— — —. "The Pseudo-Isidorian Forgeries." In *Papal Letters in the Early Middle Ages*, edited by Detlev Jasper and Horst Fuhrmann, 137–95. History of Medieval Canon Law. Washington, D.C.: Catholic University of America Press, 2001.

Gabriel, Astrik L. "Paris, University of." In *Dictionary of the Middle Ages*, edited by Joseph R. Strayer, 408–10. New York: Scribner, 1982.

Gilchrist, J. T., ed. *The Collection in Seventy-Four Titles: A Canon Law Manual of the Gregorian Reform*. Mediaeval Sources in Translation 22. Toronto: Pontifical Institute of Mediaeval Studies, 1980.

Grebner, Gundula. "Lay Patronate in Bologna in the First Half of the 12th Century: Regular Canons, Notaries, and the Decretum." In *Europa Und Seine Regionen: 2000 Jahre Rechtsgeschichte*, edited by Andreas Bauer and Karl H. L. Welker, 107–22. Vienna, 2007.

Hartmann, Wilfried, and Kenneth Pennington, eds. *The History of Medieval Canon Law in the Classical Period, 1140-1234: From Gratian to the Decretals of Pope Gregory IX*. History of Medieval Canon Law. Washington, D.C.: Catholic University of America Press, 2008.

Haskins, Charles Homer. *The Renaissance of the Twelfth Century*. Cambridge: Harvard University Press, 1927.

Herrmann, J. Berenike, Karina van Dalen-Oskam, and Christof Schöch. "Revisiting Style, a Key Concept in Literary Studies." *Journal of Literary Theory* 9, no. 1 (2015): 25–52.

Hinschius, Paul, ed. *Decretales Pseudo-Isidorianaes, et, Capitula Angilramni: Ad Fidem Librorum Manuscriptorum Recensuit, Fontes Indicavit, Commentationem de Collectione Pseudo-Isidori Praemisit*. Aalen: Scientia Verlag, 1963.

Hockey, Susan. "The History of Humanities Computing." In *A Companion to Digital Humanities*, edited by Susan Schreibman, Raymond George Siemens, and John Unsworth, 3–19. Blackwell Companions to Literature and Culture 26. Malden, MA: Blackwell Pub, 2004.

Hoeflich, Michael H., and Jasonne M. Grabher. "The Establishment of Normative Legal Texts: The Beginnings of the Ius Commune." In *The History of Medieval Canon Law in the Classical Period, 1140-1234: From Gratian to the Decretals of Pope Gregory IX*, edited by Wilfried Hartmann and Kenneth Pennington. History of Medieval Canon Law. Washington, D.C: Catholic University of America Press, 2008.

Hove, Alphonse van. *Prolegomena Ad Codicem Iuris Canonici*. Editio artera auctior et emendatior. Commentarium Lovaniense in Codicem Iuris Canonici, v. 1, t. 1. Mechliniae: H. Dessain, 1945.

Hunt, Richard William. "The Introductions to the 'Artes' in the Twelfth Century." In *The History of Grammar in the Middle Ages: Collected Papers*, edited by G. L. Bursill-Hall, 117–44. Amsterdam Studies in the Theory and History of Linguistic Science. Series 3: Studies in the History of Linguistics v.5. Amsterdam: J. Benjamins, 1980.

Hyde, J. K. "Bologna, University of." In *Dictionary of the Middle Ages*, edited by Joseph R. Strayer, 2:311–13. New York: Scribner, 1982.

Jaffé, Philipp, ed. *Regesta Pontificum Romanorum*. 2nd ed. Vol. 1. Graz: Akademische Druck, 1956.

Jansen, Katherine Ludwig, Joanna H Drell, and Frances Andrews, eds. "The Laws of King Roger II (ca. 1140s)." In *Medieval Italy: Texts in Translation*, 175–86. The Middle Ages Series. Philadelphia: University of Pennsylvania Press, 2009.

Jansen, Katherine Ludwig, Joanna H. Drell, and Frances Andrews, eds. *Medieval Italy: Texts in Translation*. The Middle Ages Series. Philadelphia: University of Pennsylvania Press, 2009.

— — —, eds. "Roman Law and Legal Studies: Three Texts (ca. 1124-66)." In *Medieval Italy: Texts in Translation*, 167–72. The Middle Ages Series. Philadelphia: University of Pennsylvania Press, 2009.

Jockers, Matthew Lee. *Macroanalysis: Digital Methods and Literary History*. Topics in the Digital Humanities. Urbana: University of Illinois Press, 2013.

Kalb, Herbert. *Studien Zur Summa Stephans von Tournai: Ein Beitrag Zur Kanonistischen Wissenschaftsgeschichte Des Späten 12. Jahrhunderts*. Forschungen Zur Rechts- Und Kulturgeschichte, Bd. 12. Innsbruck: Universitätsverlag Wagner, 1983.

Kantorowicz, Hermann. *Albertus Gandinus Und Das Strafrecht Der Scholastik*. Vol. 1: Die Praxis. Berlin: J. Guttentag, 1907.

Kestemont, Mike. "Documentary: 'Hildegard of Bingen: Authorship and Stylometry' [HD]," July 18, 2013. <https://vimeo.com/70881172>.

— — —. "Function Words in Authorship Attribution from Black Magic to Theory?" In *Proceedings of the 3rd Workshop on Computational Linguistics for Literature (CLfL)*, 59–66. Gothenburg, Sweden: Association for Computational Linguistics, 2014.

Kestemont, Mike, Sara Moens, and Jeroen Deploige. "Collaborative Authorship in the Twelfth Century: A Stylometric Study of Hildegard of Bingen and Guibert of Gembloux." *Literary and Linguistic Computing* 30, no. 2 (June 2015): 199–224.

— — . "Stylometry and the Complex Authorship in Hildegard of Bingen's Oeuvre." In *Digital Humanities 2013: Conference Abstracts*, 255–58. Lincoln, NE: University of Nebraska–Lincoln, 2013. <http://dh2013.unl.edu/abstracts/ab-126.html>.

Kirschenbaum, Matthew. "What Is Digital Humanities and What's It Doing in English Departments?" In *Debates in the Digital Humanities*, edited by Matthew K. Gold, 3–11. Minneapolis: University of Minnesota Press, 2012.

Kretzschmar, Robert. *Alger von Lüttichs Traktat "de Misericordia et Iustitia": Ein Kanonistischer Konkordanzversuch Aus Der Zeit Des Investiturstreits: Untersuchungen Und Edition. Quellen Und Forschungen Zum Recht Im Mittalter*, Bd. 2. Sigmaringen: J. Thorbecke, 1985.

Krüger, Paul, and Theodor Mommsen, eds. *Corpus Iuris Civilis*. Berolini: apud Weidmannos, 1928.

Kuttner, Stephan. "De Gratiani Opere Noviter Edendo." *Apollinaris* 21 (1948): 118–28.

— — . "Research on Gratian: Acta and Agenda." In *Studies in the History of Medieval Canon Law*. Collected Studies CS325. Aldershot, Hampshire, Great Britain : Brookfield, Vt., USA: Variorum ; Gower, 1990.

— — . "The Father of the Science of Canon Law." *Jurist* 1 (1941): 2–19.

Landau, Peter. "Gratian and the Decretum Gratiani." In *The History of Medieval Canon Law in the Classical Period, 1140-1234: From Gratian to the Decretals of Pope Gregory IX*, edited by Wilfried Hartmann and Kenneth Pennington, 22–54. History of Medieval Canon Law. Washington, D.C: Catholic University of America Press, 2008.

— — . "Gratians Arbeitsplan." In *Iuri Canonico Promovendo: Festschrift Für Heribert Schmitz Zum 65. Geburtstag*, edited by Winfried Aymans and Karl-Theodor Geringer, 691–707. Regensburg: F. Pustet, 1994.

- Larrainzar, Carlos. "El Borrador de La 'Concordia' de Graciano: Sankt Gallen, Stiftsbibliothek MS 673 (=Sg)." *Ius Ecclesiae: Rivista Internazionale Di Diritto Canonico* 11, no. 3 (1999): 593–666.
- Larson, Atria A. "An Abbreviatio of the First Recension of Gratian's Decretum in Munich?" *Bulletin of Medieval Canon Law* 29 (2011): 51–118.
- — —. *Master of Penance: Gratian and the Development of Penitential Thought and Law in the Twelfth Century*. Studies in Medieval and Early Modern Canon Law, volume 11. Washington, D.C: The Catholic University of America Press, 2014.
- Leader, Damian Riehl. "Oxford University." In *Dictionary of the Middle Ages*, edited by Joseph R. Strayer, 9:320–22. New York: Scribner, 1982.
- Luscombe, D. E. "Abelard and the Decretum of Gratian." In *The School of Peter Abelard: The Influence of Abelard's Thought in the Early Scholastic Period*, 214–23. Cambridge Studies in Medieval Life and Thought, new series, v. 14. London: Cambridge U.P, 1969.
- Manaresi, Cesare, ed. *I Placiti Del "Regnum Italiae."* . Fonti Per La Storia d'Italia, Pubblicate Dall' Istituto Storico Italiano Per Il Medio Evo 97. Roma: Tip. del Senato, 1955.
- Markus, R. A. *Gregory the Great and His World*. Cambridge ; New York: Cambridge University Press, 1997.
- Martyn, John R. C., trans. *The Letters of Gregory the Great*. Mediaeval Sources in Translation 40. Toronto: Pontifical Institute of Mediaeval Studies, 2004.
- McLaughlin, Terence P., ed. *The Summa Parisiensis on the Decretum Gratiani*. Toronto: Pontifical Institute of Mediaeval Studies, 1952.
- Minnis, A. J. *Medieval Theory of Authorship: Scholastic Literary Attitudes in the Later Middle Ages*. 2nd ed. Middle Ages Series. Philadelphia: University of Pennsylvania Press, 1988.



Moore, R. I. *The Formation of a Persecuting Society: Authority and Deviance in Western Europe, 950-1250*. 2nd ed. Malden, MA: Blackwell Publishing, 2007.

Moretti, Franco. *Distant Reading*. London: Verso, 2013.

Mosteller, Frederick, and David L. Wallace. *Inference and Disputed Authorship: The Federalist*. Addison-Wesley Series in Behavioral Science: Quantitative Methods. Reading, Mass: Addison-Wesley Pub. Co, 1964.

Müller, Wolfgang P. "The Recovery of Justinian's Digest in the Middle Ages." *Bulletin of Medieval Canon Law* 20 (1990): 1–29.

Nardi, Paolo. "Fonti Canoniche in Una Sentenza Senese Del 1150." In *Life, Law and Letters: Historical Studies in Honor of Antonio García Y García*, edited by Peter Linehan, 661–70. *Studia Gratiana* 29. Rome, 1998.

Noonan, John T. "Catholic Law School - A.D. 1150." *Catholic University Law Review* 47 (1997): 1189–1205.

— — —. "Gratian Slept Here: The Changing Identity of the Father of the Systematic Study of Canon Law." *Traditio* 35 (January 1979): 145–72.

Norberg, Dag Ludvig, ed. *S. Gregorii Magni Registrum Epistularum, Libri VIII-XIV, Appendix*. Corpus Christianorum. Series Latina, 140 A. Turnholti: Brepols, 1982.

Pennington, Kenneth. "A Short History of Canon Law from Apostolic Times to 1917," 2002.

<http://legalhistorysources.com/Canon%20Law/PenningtonShortHistoryCanonLaw.pdf>.

— — —. "Gratian and the Jews." *Bulletin of Medieval Canon Law* 31, no. 1 (2014): 111–24.

— — —. "Law, Criminal Procedure." In *Dictionary of the Middle Ages. Supplement 1*, edited by William C. Jordan, 309–20. New York: Scribner, 2004.

— — . “The Beginnings of Law Schools in the Twelfth Century.” In *Les écoles Du XIIe Siècle*, edited by Cédric Giraud. Leiden: Brill, 2018.

— — . “The Biography of Gratian, the Father of Canon Law.” *Villanova Law Review* 59 (2014): 679–706.

— — . “The Law’s Violence Against Medieval and Early Modern Jews.” *Rivista Internazionale Di Diritto Comune* 23 (2013): 23–44.

Radding, Charles. *The Origins of Medieval Jurisprudence: Pavia and Bologna, 850- 1150*. New Haven: Yale University Press, 1988.

Radding, Charles, and Antonio Ciaralli. *The Corpus Iuris Civilis in the Middle Ages: Manuscripts and Transmission from the Sixth Century to the Juristic Revival*. Brill’s Studies in Intellectual History, v. 147. Leiden ; Boston: Brill, 2007.

Rashdall, Hastings. *The Universities of Europe in the Middle Ages*. A new edition, ed. by F. M. Powicke and A. B. Emden. London: Oxford Univ. Press, 1936.

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2015. <https://www.R-project.org/>.

Reuter, Timothy, and Gabriel Silagi, eds. *Wortkonkordanz Zum Decretum Gratiani*. Monumenta Germaniae Historica. Hilfsmittel 10. München: Monumenta Germaniae Historica, 1990.

Ridder-Symoens, Hilde de, ed. *Universities in the Middle Ages*. History of the University in Europe, v. 1. Cambridge [England] ; New York: Cambridge University Press, 1992.

Rufinus. *Summa Decretorum*. Edited by Heinrich Singer. Aalen : Paderborn: Scientia Verlag ; F. Schöningh, 1963.

Rumrich, John. “Stylometry and the Provenance of de Doctrina Christiana.” In *Milton and the Terms of Liberty*, edited by Graham Parry and Joad Raymond, 125–36.



Studies in Renaissance Literature, v. 7. Cambridge, UK ; Rochester, NY: D.S. Brewer, 2002.

Samuels, Lisa, and Jerome McGann. "Deformance and Interpretation." *New Literary History* 30, no. 1 (1999): 25–56.

Schulte, Johann Friedrich von, ed. *Die Summa Des Paucapalea über Das Decretum Gratiani*. Giessen: E. Roth, 1890.

Schulte, Johann Frierich, ed. *Die Summa über Das Decretum Gratiani*. Aalen: Scientia Verlag, 1965.

Simon of Bisignano. *Summa in Decretum Simonis Bisinianensis*. Edited by Pier Virginio Aimone-Braida. Monumenta Iuris Canonici. Series A, Corpus Glossatorum ; Vol. 8. Città del Vaticano: Biblioteca Apostolica Vaticana, 2014.

Singer, Heinrich. *Summa Decretorum*. Aalen : Paderborn: Scientia Verlag ; F. Schöningh, 1963.

Somerville, Robert, and Bruce Clark Brasington, eds. *Prefaces to Canon Law Books in Latin Christianity: Selected Translations, 500-1245*. New Haven, Conn: Yale University Press, 1998.

Sommar, Mary E. *The Correctores Romani: Gratian's Decretum and the Counter-Reformation Humanists*. Pluralisierung & Autorität, Bd. 19. Berlin : New Brunswick: Lit ; distributed by Transaction Publishers, 2009.

Southern, R. W. *Scholastic Humanism and the Unification of Europe*. Vol. 1. Oxford, UK ; Cambridge, Mass., USA: Blackwell, 1995.

Thompson, Augustine, and James Gordley, trans. *The Treatise on Laws: (Decretum DD. 1-20)*. Studies in Medieval and Early Modern Canon Law, v. 2. Washington, D.C: Catholic University of America Press, 1993.

- Verger, Jacques. "Patterns." In *Universities in the Middle Ages*, edited by Hilde de Ridder-Symoens, 35–74. History of the University in Europe, v. 1. Cambridge [England] ; New York: Cambridge University Press, 1992.
- Watson, Alan, ed. *The Digest of Justinian*. Philadelphia, Pa: University of Pennsylvania Press, 1985.
- Wei, John C. *Gratian the Theologian*. Studies in Medieval and Early Modern Canon Law, Volume 13. Washington, D.C: Catholic University of America Press, 2016.
- Weigand, Rudolf. *Die Glossen Zum "Dekret" Gratians: Studien Zu Den Frühen Glossen Und Glosskompositionen*. *Studia Gratiana* 25-26. Rome, 1991.
- — —. "Frühe Kanonisten Und Ihre Karriere in Der Kirche." *Zeitschrift Der Savigny-Stiftung Für Rechtsgeschichte. Kanonistische Abteilung* 76 (1990): 135–55.
- — —. "The Development of the *Glossa Ordinaria* to Gratian's *Decretum*." In *The History of Medieval Canon Law in the Classical Period, 1140-1234: From Gratian to the Decretals of Pope Gregory IX*, edited by Wilfried Hartmann and Kenneth Pennington, 55–97. History of Medieval Canon Law. Washington, D.C: Catholic University of America Press, 2008.
- Winroth, Anders. "Law Schools in the Twelfth Century." In *Mélanges En L'honneur d'Anne Lefebvre-Teillard*, edited by Bernard d' Alteroche, Florence Demoulin-Auzary, Olivier Deschamps, and Franck Roumy, 1057–64. Paris: Éd. Panthéon-Assas, 2009.
- — —. *The Making of Gratian's Decretum*. Cambridge: Cambridge University Press, 2000.
- — —. "The Teaching of Law in the Twelfth Century." In *Law and Learning in the Middle Ages*, edited by Mia Münster-Swendsen and Helle Vogt, 41–62. Copenhagen: DJØF, 2006.
- — —. "Uncovering Gratian's Original *Decretum* with the Help of Electronic Resources." *Columbia Library Columns* 46, no. 1 (1997): 26–31.



- — —. "Where Gratian Slept: The Life and Death of the Father of Canon Law." *Zeitschrift Der Savigny-Stiftung Für Rechtsgeschichte: Kanonistische Abteilung* 99, no. 1 (August 2013): 105–28.
- Zechiel-Eckes, Klaus. "Ein Blick in Pseudoisidors Werkstatt. Studien Zum Entstehungsprozeß Der Falschen Dekretalen. Mit Einem Exemplarischen Editorischen Anhang (Pseudo-Julius an Die Orientalischen Bischöfe, JK +196)." *Francia* 28, no. 1 (2001): 37–90.
- Zipf, George Kingsley. *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. Cambridge, Mass: Addison-Wesley Press, 1949.
- — —. *The Psycho-Biology of Language: An Introduction to Dynamic Philology*. Boston: Houghton Mifflin Company, 1935.