The Vocabulary of Gratian's *Decretum*: Change Over Time[1]

Paul Evans

Abstract

The textual history of Gratian's *Decretum* is now understood well enough to enable researchers to comparatively read first-recension, second-recension, and vulgate versions of the text to identify evolutionary changes in its doctrine. Specific developments can be identified by close reading well-chosen selections from the text (e.g., Pennington, "The Law's Violence against Medieval and Early Modern Jews"). However, a systematic search for such doctrinal developments is best conducted with computational assistance, taking advantage of recent advances in the application of large language models to natural language processing tasks. Using the PIE lemmatizer in conjunction with the LASLA Latin models, it is possible to lemmatize samples reflecting different compositional stages of corresponding sections of the *Decretum* and to compare the results, isolating lemmas unique to specific stages in the development of the text. Distinctive vocabularies of different compositional stages surfaced by this method then provide a roadmap for close reading in context.

The work I will be presenting today was originally intended to be part of my dissertation. In the interest of completing the dissertation before the end of the year, I decided in March of 2021 that I had gotten as far as I was going to get, and that the chapter in which this work was to be discussed would have to be left out. "Hitherto shalt thou come, but no further" (Job 38:11).

---

[1] ICMCL XVII – Congress Programme
Tuesday (9 July): Cathedral Conference Suite
15.30-17.00: Sessions VI
Session 18: Texts and Manuscripts in the Age of Gratian and Beyond
Room: Claggett Auditorium Balcony.
Moderator: Anders Winroth, University of Oslo, Norway.
Speaker 18b: Paul Evans, University of San Diego, USA.
'The Vocabulary of Gratian's Decretum: Change Over Time'
The maximum duration of each presentation (not inclusive of questions) is:
Speakers in sessions lasting 90 minutes and comprising three presentations:
20 minutes.

The larger project used techniques from computational linguistics to analyze the authorship of the case statements and *dicta* in Gratian's *Decretum*. My conclusion was that the case statements were written by a single author who was not the author of the *dicta* either in the first or second recension or in *de Penitentia*. Results from stylometric analysis for authorship of the *dicta*, on the other hand, were not consistent with either the one-author theory championed by Ken Pennington or the two-author theory championed by Anders Winroth. Instead, the results suggested, but did not conclusively prove, that both the first and the second recension *dicta* were the work of multiple authors.

All of my work is in some sense part of an overarching project to enhance the effectiveness of close reading medieval texts, Gratian's *Decretum* in particular, by using computational assistance. The Digital Humanities world uses the term "distant reading" to describe this kind of computational assistance, emphasizing both its connection to and its contrast with traditional close reading. However each project that uses a distant or machine reading approach is different in the questions it seeks to answer, and therefore in the tools and techniques it uses to explore them. The dissertation project was concerned with the question of the authorship of the case statements and the first- and second recension *dicta*, including the *dicta* in *de Penitentia*. It employed stylometric authorship analysis that used a statistical technique – principal component analysis of the frequencies of commonly occuring function words – to arrive at its results. In contrast, this project is concerned with changes in the teaching – the doctrine – of the *Decretum* between the first and second recensions as indicated by the use of distinctive vocabulary in the first- and second recension *dicta*. It employs lemmatization – a linguistic technique – using the PIE lemmatizer and a large language model (LLM) based on the LASLA corpus.[2] What the two projects have in common, for now, is the

[2] Enrique Manjavacas, Ákos Kádár, and Mike Kestemont, "Improving Lemmatization of Non-Standard Languages with Joint Learning," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human*

same data set. The dissertation project relied on a data set of the texts of the case statements and the first- and second-recension *dicta* that I very carefully and laboriously prepared.[3] The work I am going to discuss today depends on that same data set. To recapitulate: the overarching goal is the same – to use computational distant reading techniques to enhance our close reading of the *Decretum* – but each individual project attempts to answer different questions using different tools and techniques (although sometimes the same data set).

It is worth noting that this is *not* my first attempt to make progress on the problem of computationally identifying topics added to the *dicta* between the first and second recensions.

At the stage in my PhD program when I was starting to think ahead to my dissertation proposal (this would have been around 2012), there was tremendous enthusiasm in the Digital Humanities world for a technique called unsupervised topic modeling and in particular for a topic-modeling tool called MALLET.[4] Inspired by Ken 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,[5] I hoped to use MALLET to systematically identify new topics added in the

---

Language Technologies, Volume 1 (Long and Short Papers) (Minneapolis, Minnesota: Association for Computational Linguistics, 2019), 1493–503.

[3] This research could usefully be expanded to include the rubrics and canons, and I have the workflow in place to do so. But the workload required to expand the corpus to include the canons would realistically require some level of grant funding.

[4] MAchine Learning for LanguagE Toolkit
Andrew Kachites McCallum, "MALLET: A Machine Learning for Language Toolkit" (2002).

[5] 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, https://muse.jhu.edu/article/602133.

second recension. The idea was to topic model all the *dicta* in the vulgate *Decretum* and then topic model just the *dicta* in the first recension, and then see what topics were left when the first recension topics were subtracted from the vulgate topics. Simple in concept, but prohibitively difficult in practice, for two reasons. First, the difficulty in determining the number of topics to look for (a necessary precondition for unsupervised topic modeling); and second, the fact that there was no obvious way to subtract topics.

So once it became clear that unsupervised topic modeling using MALLET was not going to be an effective way to identify topics added to Gratian's *Decretum* between the first and second recensions, the most promising alternative approach to the problem appeared to be using lemmatization to identify distinctive *vocabulary* (as a signpost pointing to new ideas) added between the first and second recensions.

Because when working in a highly inflected language like Latin, using words as the signposts pointing to corresponding ideas is not precise enough. To anticipate an example that we will look more closely at later in the presentation, the noun *calumnia* has 6 unique declined forms. A normal Latin verb has 120 conjugated forms, although not all of them are unique, and that does not include the participial forms. I did not count the number of unique forms that a first conjugation deponent verb like *calumnior*, *calumniari*, *calumniatus* has, but the number is large. So if we want to use distinctive vocabulary as a basis for determining whether or not an idea or topic is present in a Latin language text, we need to lemmatize every word form we encounter – that is, reduce it to its dictionary headword.

However the results of my initial experiments with the Classical Language Toolkit (CLTK), built on top of the Python Natural Language Toolkit (NLTK) and the best lemmatization tool available at the time, were not encouraging. The first- and second-recension *dicta* – 56,713 and 14,255 words respectively – might reasonably be expected to include a few hundred unique lemmas, but CLTK reported many thousands (over four thousand just for the first-recension *dicta*), the overwhelming majority of which were

false positives with no readily detectable pattern.[6] For my purposes at least, lemmatization was not ready for prime time, and that remained the case for many years, from around 2014 through around 2020.

That changed in early 2021, when Mike Kestemont made me aware of the PIE lemmatizer. Kestemont is a researcher at the Universiy of Antwerp specializing in medieval Latin and Middle Dutch literature. He is also a leading figure in the field of computational text analysis whose advice on such matters I very much make it my business to seek and follow. The early results of my experiments with using PIE to lemmatize the *dicta* data set were reasonably promising, so it was with some reluctance that I set aside the work in order to complete my dissertation.

I want to make it clear that PIE is not just a program that you run – you do not just type a command or click a button and get lemmatized text as output. PIE and PIE extended are libraries, packages, toolkits, that provide an extremely versatile set of software building blocks that can be called upon to perform a wide range of natural language processing functions, like part-of-speech tagging or lemmatization, from within a Python program.[7] They are based on large language models (LLMs) trained using machine learning techniques on annotated corpora of texts in the target language. In this case, we are using a model trained on the LASLA corpus of 1.7 million tokens or words of classical Latin text, each annotated with lemma, part of speech, and other morphological and syntactic information.

---

[6] This may be the place to get explicit about what I mean by *unique* lemmas. When comparing any two text samples (here, the first- and second-recension *dicta*), every lemma either appears in both, or is unique to one or the other.

[7] I would like to acknowledge Jake Bayon, an undergraduate Computer Science student at the University of San Diego, who set up the PIE lemmatization environment as independent study project with me during the Spring 2024 semester, and who learned something about Gratian in the process. PIE can only be installed with the 2019 Python 3.8 release – the current release is Python 3.12 – so setting up the lemmatization environment was not a trivial task.

If the term large language model or the acronym LLM sound familiar, they should. LLMs are the basis for the growing family of generative AI tools, ChatGPT and friends, that have been the subject of so much attention for the last two years.

Once the PIE lemmatization environment had been set up, I wrote a Python program that used PIE to create separate lists of every lemma found in the first- and second-recension *dicta*, and then compare the two lists to identify lemmas that appear only in the second-recension *dicta*. The program output a list of 725 unique lemmas present only in second-recension *dicta* and absent from the first-recension *dicta*.[8]

[Slide]

To define terms clearly, for the purpose of these experiments I am defining "first-recension *dicta*" as the text of the *dicta* as they are listed in the appendix of Winroth's *The Making of Gratian's Decretum*, and I am defining "second-recension *dicta*" as the words in the text of the *dicta* as they appear in the Friedberg edition when the words in the *dicta* listed by Winroth in the appendix have been taken away. D.54 d.p.c.23 is a good example. Winroth's appendix indicates that only the first sentence of the *dictum* appears in the first recension. Therefore, the first sentence of the *dictum* is assigned to the first recension text sample, and the remainder is assigned to the second recension text sample.

An understanding of whether an idea or topic is present in or absent from a selection of text can almost never be arrived at based on the presence or absence of a single lemma. Instead, human (as opposed to machine) readers look for the presence of families of related lemmas to signal the presence of an idea or topic in a selection of

[8] The 728 lines of program output included 3 numbers, which I discarded.

text. In reviewing the list of 725 unique lemmas one such family of lemmas in particular stood out to me: *calumpia*, *calumniator*, *calumpniatus*.[9]

We know that between 1140 and 1234, what we think of as the classical period in the history of medieval canon law, the concept of calumny took on a significance and a formal legal meaning that was derived from but was considerably more technically precise than its previous general use in Christian discourse. For example, during this period, oaths of non-calumniation at the onset of legal proceedings came to be required of all litigants in canonical courts.

We should expect to see at least three Latin lemmas associated with the concept of calumny:

- the deponent verb *calumnior*, *calumniari*, *calumniatus* including participial forms like *calumnians* and *calumniatus*,
- the feminine noun *calumnia* corresponding to the idea of calumny in the abstract,
- and the masculine noun *calumniator*.

One form related to the extended family of lemmas we are considering *does* appear in the first-recension *dicta*, *calumpniantibus*, lemmatized by PIE as the verb *calumpnio*. As a result, the lemma *calumpnio* does not appear on the list of lemmas unique to the second-recension *dicta*.

[Table Slide]

Conclusion (minor)

The long-term goal of this project has has been to find a way to use computationally-enabled distant reading – “reading machines” in the words of Stephen Ramsay – to efficiently direct the attention of trained researchers to specific sites in the text of

[9] *calumpia* is almost certainly a typo in the LASLA Latin language model for *calumpnia*.

Gratian's *Decretum* where new topics added between the first and second recensions are likely to be found for close reading.

And as a proof of concept, I think this effort was a success. It is certainly closer to an interesting result than I have gotten in the 12 years or so I have been thinking about and intermittently working on this problem.

Using the PIE lemmatizer in conjunction with the LASLA Latin large language model (LLM) to systematically lemmatize every word in the *dicta* and then listing all of the lemmas that do appear in the second recension *dicta* but do not appear in the first recension *dicta* allowed me to identify a family of lemmas (*calumnia*, *calumniator*, and *calumniatus*) that point towards the canonically significant concept of calumny.

And when we turn our attention to the substantive treatment of the topic of calumny in the *dicta*, there is variation in terms of the legal sophistication with which the concept is handled, moving generally in the direction of greater technical precision and sophistication. (I say "generally" because while we can assume the first recension *dicta* were written before the second recension *dicta*, we do not have enough information to speculate about the temporal relationship within the second recension *dicta*.

The concept of calumny makes its initial appearance in the form of a (slightly) misquoted scriptural platitude in the first-recension *dictum* C.24 q.3 d.p.c.9. The treatment of the concept in the second-recension *dictum* C.23 q.4 d.p.c.23 §3 is in a similar spirit, although in the latter case the scriptural allusions are mediated through a patristic source, Augustine's *de Dono perseverantiae*, a treatise on predestination. As noted previously, both of these *dicta* use the words associated with the concept of calumny in the same general, non-technical, sense they had in the first millennium of Christian discourse.

Not so with the second recension *dictum* C.23 q.3 d.p.c.8. Here we see a series of quotes from book 48 of the *Digest*, containing 7 words, 5 of them unique, mapping to all

4 of the expected lemmas related to the concept of calumny. It is of obvious interest that the quotations in this *dictum* are from Justinianic Roman law rather than from scriptural or patristic sources. And most interesting of all is the fact that the *dictum* contains two first-person sayings by Gratian 2, the author of the second-recension *dicta* (or at least of this *dictum*), in effect glossing the terms calumniator and calumnia. This shows an increased, but still limited, level of legal sophistication in the sense that the discussion draws on resources from Justinianic Roman law, but that Gratian's own intervention is relatively modest.

To summarize: the technique of searching among lemmas that are unique to the second recension using distant or machine reading was successful at least as a proof of concept insofar as it did surface one family of related lemmas pointing towards a canonically significant topic, calumny. On close reading, the sites in the text of the *dicta* identified by the results did indeed show an interesting change over time in the vocabulary of Gratian's *dicta* and to that extent in the teaching of the *Decretum* on this topic. Calumny was the most obvious topic (at least to me), and I was surprised that there were no other immediately obvious conceptually related families of lemmas in the results, but I encourage those of you who are interested to take a look at the list this week and let me know if you see something I did not. As I previously indicated, there is limited value in the results of machine reading by itself. The real value of the results of machine reading lies in the patterns that trained researchers *see* in them.

Conclusion (major)

Is it enough? No.

I have tried to emphasize that the project I have been presenting here today is very much a work in progress and that the results, although interesting, are limited to the *dicta*, and therefore should not be taken as anything more than a proof of concept.

Casual searching through the MGH e-text of the Friedberg edition that was created for the *Wortkonkordanz zum Decretum Gratiani* edited by Reuter and Silagi indicates that there are occurrences of forms of the words I have been focusing on – *calumnia*, *calumnior*, and *calumniator* – in the rubrics and canons.

A really thorough approach to the problem of systematically identifying new topics added to the *Decretum* between the first and second recensions is going to require a data set that includes the rubrics and canons as well as the *dicta* and case statements. (And might as well include the inscriptions while we are at it.) Ideally, such a data set would be in the form of a new e-text in TEI-P5 XML format incorporating texts from both the old Friedberg edition and the new Winroth edition-in-progress of the first recension. And this is where the scale of the undertaking starts to get really challenging. Even without the overhead of structuring the data set as a TEI-P5 document, I spent something like 12 person-weeks on corpus preparation for the *dicta* and the case statements as part of my dissertation project. Since the word count of the canons is roughly five times that of the *dicta*, one person-year is not an unreasonable initial estimate for corpus preparation for a comparable data set for the canons.

The work I have presented today is based on a highly customized version of a 20th century e-text of a 19th century print edition of the *Decretum*. The MGH e-text of the Friedberg edition is the indispensable free resource without which none of my work, and I suspect the work of many others, is possible. But like so many free things in life, someone paid a great deal of money to make it free (in this case, presumably the taxpayers of the Federal Republic of Germany in the 1980s and 1990s). But the MGH e-text is a resource that because of its archaic format is approaching the end of it useful life. If we want to continue to advance in our understanding of Gratian's *Decretum* with the help of electronic resources, we need to invest time, effort, and grant funding into a 21st century electronic text, or better still an electronic edition, of Gratian's *Decretum* that meets 21st century research needs.

Bibliography

Manjavacas, Enrique, Ákos Kádár, and Mike Kestemont. “Improving Lemmatization of Non-Standard Languages with Joint Learning.” In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 1493–503. Minneapolis, Minnesota: Association for Computational Linguistics, 2019.

McCallum, Andrew Kachites. “MALLET: A Machine Learning for Language Toolkit,” 2002.

Pennington, Kenneth. “Gratian and the Jews.” *Bulletin of Medieval Canon Law* 31, no. 1 (2014): 111–24. https://muse.jhu.edu/article/602133.

———. “The Law’s Violence Against Medieval and Early Modern Jews.” *Rivista Internazionale Di Diritto Comune* 23 (2013): 23–44.