

## Naming Latin Texts in the *Thesaurus linguae Latinae* Index of Sources

Adam Gitner

**Abstract:** This paper discusses a specific category of names in the *Thesaurus linguae Latinae*: names of ancient authors and works used as source citations. These names cover nearly all known ancient Latin texts until about 600 CE, including ca. 1,035 authors or text groups and ca. 2,880 identifiable works. They were most recently printed as the *Index librorum scriptorum inscriptionum ex quibus exempla afferuntur* (1990) and, in partly updated form, appear [online](#). Currently efforts are underway to transform this into an open-access database as part of a larger digital transformation of the project. The *Index* includes information about the dates of authors and works, authority control data, modern editions, the existence of ancient translations into or out of Greek, sample citations, and other meta-data. The paper discusses theoretical and practical issues involving these names and their future data structure.\*

### 1. Introduction

If you open any page of the *Thesaurus linguae Latinae* (*TLL*),<sup>1</sup> the most comprehensive dictionary of ancient Latin, one of the first things you will notice is a densely printed sequence of mostly abbreviated source citations. For instance, the following passage comes from the article *omen*, written by Frits Oomes and published in 1974, illustrating omens involving names (*nomina*):<sup>2</sup>

---

\* I would especially like to thank Josine Schrickx, Nora Götze, and the rest of my colleagues on the *TLL* Digital Workgroup: Massimo Cè, Oliver von Criegern, Roberta Marchionni, and Martin Shedd. The views expressed are my own, as are the errors and omissions. I would also like to thank Manfred Flieger and Cornelis van Leijenhorst for advice and details about the history and use of the *Index* and Yannick Anné and Tim Denecker at Brepols for assistance with word counts of Latin literature.

1 The *TLL* is available online in open access excluding the three most recent years of publication: <http://tll-open.badw.de/> (last access 04.04.2025). The subscriber-only De Gruyter platform offers the latest publications and the ability to do a full-text search: <https://tll.degruyter.com/> (last access 04.04.2025). On using the *TLL*, see n. 4.; on its history, see e.g. Bögel et al. (1996), Corbeill (2007), Flow (2015), and the literature listed at <https://thesaurus.badw.de/ueber-den-tll/literaturhinweise.html> (last access 04.04.2025).

2 It is impossible to consult the article *nomen* since the *N* volume is not yet complete. The first *N* fascicle appeared in 2011 (volume XI, 1 fasc. 1: *n – navalis*) and work has progressed to *netura*. For the publication history of the *TLL*, see <https://thesaurus.badw.de/ueber-den-tll/publikationen.html> (last access 04.04.2025).

<p><i>γ de nominibus, appellatione (nom. propr. praeter p. 576, 4; cf. etiam supra l. 39. 45):</i> <u>PLAVT. Persa 625 nomen atque -n</u> (sc. 'Lucris', cf. <i>de lacu Lucrino</i>: <u>PAVL. FEST. p. 121</u> in vectigalibus . . . primus locatur . . . -is boni gratia ut in dilectu . . . primi nominantur Valerius, Salvius, Statorius). <u>CIC. Verr. II 2, 18</u> -n . . . famae . . . cum ex nomine <i>Verris</i> augurabantur (<i>item 2, 19</i> paratur <i>Verres</i> ex illo -e urbano ad everrendam provinciam). <u>Scaur. 30</u> -n <u>nominis Valerii</u> (<u>Phil. 7, 11</u> Brutum -e [hom- var. l.] quodam illius generis et n. natum). <u>VARRO ling. 5, 159</u> a bono -e . . . appellarunt (sc. <i>Vicum Ciprium</i>); nam <i>ciprum</i> Sabine bonum. <u>PROP. 4, 1, 25</u> Alba . . . albae suis -e nata. <u>4, 10, 46</u> Feretri, -e (crimine <i>recc.</i>) quod certo dux ferit ense ducem. <u>Ov. fast. 1, 616</u> auspiciis . . . deis tanti cognominis heres -e suscipiat, quo pater, orbis onus sc. <i>Augustus</i> (ab <i>augendo</i> vel <i>augurio</i>). <u>Liv. 5, 34, 9</u> <i>Insubres</i> -n sequentes loci (sc. <i>agri Insubrii</i>). <u>7, 25, 11</u> ob -n faustum . . . cognominis. <u>29, 27, 12</u> cum Pulchri</p> <p>promunturium id vocari audisset, 'placet -n' inquit <i>egs.</i> <u>PAVL. FEST. p. 34</u> Beneventum . . . appellari . . . melioris -is causa. <u>MELA 2, 56</u> nomen <i>Epidamni</i> mutavere, quia velut in <i>damnum</i> <i>ituris</i> -n id visum est. <i>al. similiter explicatur appellativum</i>: <u>VARRO ling. 5, 116</u> ab -e (hom- trad.) pilum, qui hostem feriret, ut <i>perilum</i>.</p>	<p>70</p> <p>75</p> <p>80</p>
---	-------------------------------

Fig. 1: TLL, s.v. *omen*, IX 2, 575, 70–576, 5 (Oomes [1974]).

The source citations (*Zitierweisen*) have all been underlined in red, such as “PLAVT. Persa 625”, “PAVL. FEST. p. 121”, “PROP. 4, 1, 25”.<sup>3</sup> Since *TLL* articles are arranged in strict chronological order, the first citation in this section, from Plautus’s *Persa*, is also the oldest surviving attestation of the word *omen* in the relevant sense. Strikingly it already illustrates the jingle *nomen atque omen* ‘name and omen’, which is part of the title of this volume. Though the conventions of the *TLL* may seem intimidating at first glance, with patience one can extract much more information about the meaning, range, and distribution of the word *omen*.<sup>4</sup>

As an onomastic convention, the use of standardized source references is essential for modern lexicography. Besides saving space, it provides the necessary scholarly precision so that the reader knows exactly which part of which edition is under discussion. This is especially important for ancient texts, which have been subject to proliferating and incompatible systems of pagination, capitulation, and numeration since the advent of printing. For instance, for the poet Propertius (“PROP.”), modern editors disagree not just about how to number individual poems, but how many books have been transmitted, and where the divisions lie. Similarly, when a text is cited by page number, as in “PAVL. FEST. p. 121”, how does the reader know which edition is referenced?

Answering these questions is the job of a meta-lexicographic reference work that has developed alongside the *TLL*: The *Index librorum scriptorum inscriptionum ex quibus exempla afferuntur* or here *Index* for short. In its latest incarnation, the *Index* exists as an open access HTML website in tabular format.<sup>5</sup> The website incorporates hundreds of *addenda* and *corrigenda* that have been adopted since the *Index* was last printed in 1990.<sup>6</sup> Since the *TLL* aims to cover *all* Latin produced before 600 in any medium, the *Index* provides, in effect, a database of nearly all known Latin texts within this time period. Currently it has nearly 5,000 entries, five times as many as the list of abbreviations in the *Oxford Latin*

3 When a following citation shares elements with the preceding citation, the identical elements are usually not repeated. For instance, “Scaur. 30” appears instead of “CIC. Scaur. 30” because the previous citation is also from Cicero.

4 A good place to start is the specimen article provided on the *TLL* website: <https://thesaurus.badw.de/en/hilfsmittel-fuer-benutzer/specimen-article.html> (last access 04.04.2025). The canonical account is the *Praemonenda* (Keudel [1990]).

5 <https://thesaurus.badw.de/tll-digital/index/a.html> (last access 22.04.2025).

6 Krömer / van Leijenhorst (1990); for the *addenda* see n. 32.

*Dictionary*, whose coverage only extends to around 200. But the *Index* is much more than just a list of abbreviations: it provides datings considered authoritative by the *TLL*, prosopographic information, concordances of various kinds, bibliographies of scholarly editions, and even information about translation from and into Greek. These features of the *Index* are often overlooked, and have scholarly value far beyond the lexicographic use for which it is primarily designed.

Currently the *Index* is undergoing a digital make-over as part of a more profound digital transformation of the *TLL*. Perhaps the most significant goal is to release all new lexicographic articles in open-access XML starting in 2026, which requires a more thorough digitalization of many components of its lexicographic workflow. As part of this, the *Index* data is being migrated to a new database that can facilitate the production of XML and handle complex queries.

Because of this digital transformation it is an opportune moment to take stock of the current *Index* and related onomastic conventions. First, I will discuss the underlying syntax and semantics of *TLL* source citations. These conventions deserve to be more widely known and adopted by classicists because of their consistency and breadth of coverage. Yet, as those of us working on transforming the *Index* have discovered, the conventions are more complex than they seem, raising difficult questions about the nature of authorship and ancient titulature. Explaining how these work in detail will be helpful to users of the *TLL* in general, and it will also facilitate the greater formalization of these names in future digital formats, such as Canonical Text Services. Second, I will provide an overview of the *Index* as a reference work for the names of ancient Latin sources, emphasizing the scope and variety of information available there. By way of conclusion, I will sketch some possible digital futures for the *Index* and its onomastic conventions.

## 2. The Syntax of *TLL* Source Citations

### 2.1 General Remarks

The practice of abbreviating Latin lexicographic sources can be traced back to antiquity,<sup>7</sup> and was elaborated by medieval lexicographers, such as Papias in the eleventh century.<sup>8</sup> But there is a crucial difference between the sporadic and inconsistent use of abbreviations one finds in ancient glossography – for instance variation among *Virg.*, *Virgl.*, and *Virgilius* depending on how much space is at the end of a line – and the imposition of a *coherent system* of source abbreviation that is unambiguous and internally consistent. The increasing formalization and precision in citation practice since Robert Estienne’s original *Thesaurus linguae Latinae*, first printed in 1531, reflects the growing professionalization of classical scholarship as well as the greater complexity of available Latin source material.

The modern *TLL* has developed these conventions over time into a comprehensive ontology for citing, as far as possible, all known ancient Latin texts produced before 600. The conventions strive to be as intuitive and self-explanatory as possible, so that a reader does not have to spend too much time decoding an unfamiliar abbreviation. Instances such as “PLAVT. Persa 625” (Plautus, *Persa*, line 625) and “AVG. in euang. Ioh. 10, 10” (Augustine, in *Iohannis euangelium tractatus*, tractate 10, chapter 10) are probably decipherable by most classicists even if they have never read these texts before. However, the pattern implicit in these straightforward cases has been extended over time – more by

---

7 For instance, in the medieval manuscripts of Arusianus Messius’s *Exempla elocutionum*, there is inconsistent abbreviation of the names of Cicero, Sallust, Terence, Virgil, and their works; the most recent editor (Di Stefano 2011, p. XCVII) has preserved these on the grounds that it goes back to fourth-century practice. On the possibly related practice of *sigla personarum* in Latin dramatic texts, see e.g. Andrieu 1940.

8 On source marks in the eighth-century *Liber glossarum* see Grondeux (2015), in Papias, see Grondeux (2023), and in medieval usage more broadly, Grabowski (2024), 207–220.

improvisation than by any abstract rules – to cover some extremely complex sources, such as different witnesses of the *Vetus Latina*, e.g. “VET. LAT. I Cor. 4, 13 (cod. 75. *sim. al. et VVLG.*)”,<sup>9</sup> and different recensions of the Latin Oribasius, e.g. “ORIBAS. eup. 2, 1 A 24 La p. 485”.<sup>10</sup> Moreover, non- and para-literary sources, which have become available in greater abundance and diversity since the nineteenth century, require different paradigms of citation that do not center individual authorship. In the *TLL*, these texts may be cited by their place of publication (e.g. “CIL I<sup>2</sup> 3191”) or various combinations of genre and publication (e. g. “TAB. devot. Audollent 233, 32” for a curse tablet edited by Audollent 1904). To give one amusing example, there is even an abbreviation for inscriptions preserved on ancient gaming pieces, so called *tabulae lusoriae* (“INSCR. tess. lusor.”).<sup>11</sup>

The brief, canonical account in the *TLL Praemonenda* does not offer terminology for describing all the elements of complex citations, nor does it say anything about how abbreviations are assigned.<sup>12</sup> To some extent the lack of strict codification is actually an advantage because it creates room for the improvisation that is sometimes necessary when citing complex sources and allows tolerance for the minor internal inconsistencies that have inevitably developed over time. What I offer here are some personal observations based on my experience as a lexicographer working on the dictionary and based on collective discussions in the *TLL* Digital Workgroup, but they are not meant to be an official or exhaustive account.

While the *TLL* system tries to be as comprehensive as possible, no single system is going to be adequate or appropriate for every scholarly task. Very often the form of a source citation is a compromise choice among conflicting criteria, such as brevity versus clarity or the latest scholarship versus backwards compatibility. Specialists in particular fields will sometimes need greater precision. For instance, a papyrologist may want Mertens-Pack abbreviations or *TM* identification numbers,<sup>13</sup> while a scholar of ancient grammar may need to distinguish the titles and authors of works that the *TLL* lumps together.<sup>14</sup> Ideally the *TLL* adapts to changing practices in these specialty fields, but every proposed change must be carefully weighed up against internal considerations, including the needs of non-specialist readers.

---

9 Translated into discursive prose this means: the *Old Latin* Bible text of the First Epistle to the Corinthians, chapter 4, verse 13, as found in codex number 75 (numbered according to Gryson 1999–2004), which is similar to the text in some other *Old Latin* witnesses and Jerome’s *Vulgata*.

10 Translated into discursive prose: the Latin translation of Oribasius’s *Euporista* book 2, chapter 1, under the Greek letter *lambda* entry number 24, in the La recension, which is found on page 485 of Molinier’s edition (1873–1876).

11 These *tesserae lusoriae* sometimes preserve lexicographically relevant occurrences of relatively rare words, such as *nugator* ‘trifler’ at INSCR. tess. lusor. (cil x 8070) 13; see Baratta (2019).

12 Keudel (1990), 12 (under the subtitle *ad exempla afferenda et explicanda*). The account there is effectively bipartite, distinguishing between *nomina testium* (e.g. “CIC.”) and *opera locorumque nomina* (e.g. “Verr. II 2, 156”).

13 For some proposals about citing Latin literary papyri see Scappaticcio (2019). The Mertens-Pack<sup>3</sup> conventions are managed by CEDOPAL (<http://www.cedopalmp3.uliege.be/> [last access 21.04.2025]); for *TM* stable identifiers, see [https://www.trismegistos.org/about\\_how\\_to\\_cite.php](https://www.trismegistos.org/about_how_to_cite.php) (last access 21.04.2025).

14 For Roman grammarians, the most up-to-date list of sources, abbreviations, and editions is the *Index Grammaticorum Latinorum*, managed by Michela Rosellini and Elena Spangenberg Yanes as part of the Latin Grammarians Collection: <https://latingrammarianscollection.uniroma1.it/site/index-grammaticorum-latinorum-intro> (last access 21.04.2025).

## 2.2 *Tria nomina*?

Like Roman names, *TLL* citations have two obligatory elements, but can expand up to three or even four elements. The most general identifier is placed on the left and as one moves rightward one encounters increasingly specific hierarchies. The structure is best illustrated with some labeled examples:

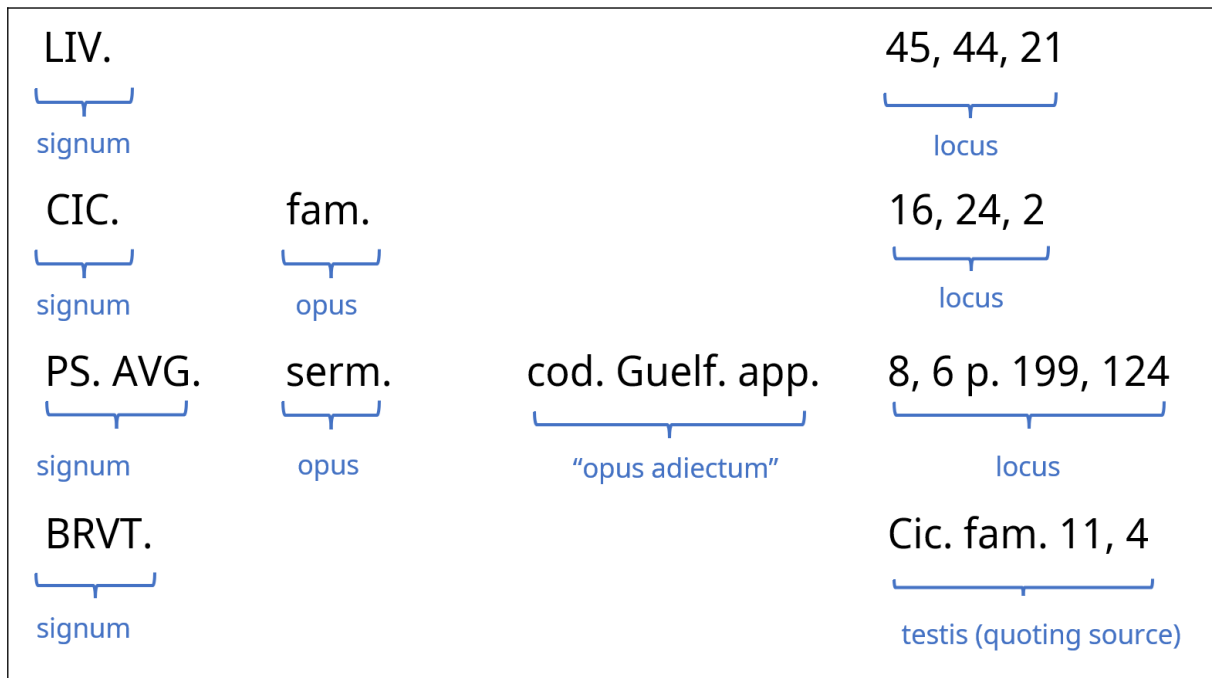


Fig. 2: Sample *TLL* Citations.

## 2.3 *Signum*

The first obligatory element is a word written in all capital letters here (or small caps in print). In the absence of an official title, I call this the *signum* in the sense ‘nickname’, as used in late antique onomastic formulae for various unofficial names.<sup>15</sup> In a *TLL* article, capitalization and Roman type are only used when the cited passage contains the lemma word; all other citations are written in sentence case and italicized.<sup>16</sup>

At first glance it is tempting to consider this element the ‘author’, and indeed in most of these examples this is so: “LIV.” seems to refer to the historian T. Livius and “CIC.” to Cicero. But the situation is more complex in the case of “PS. AVG.” for ‘Pseudo-Augustine’, which, as used in the *Index*, is a cover term for several anonymous authors whose works have been intentionally or unintentionally confused with those of Augustine. In fact, if we look at “LIV.” in the *Index*, we will see it is also used to identify later historical works *derived from* Livy, but not written by him: namely, the so-called Oxyrhynchus Epitome (“LIV. epit. Oxyrh.”). Moreover, even where authorship can be confidently assigned, it is not always given prominence in first position: for instance, Augustus’s *res gestae* are labeled “R. GEST. div. Aug.” rather than “AVG. r. gest.”.

Consequently, it is safest to regard the *signum* in the broadest sense as designating a group of works (or, when there is no further subdivision, a single work). In some cases these are defined by common authorship, but often the relationship is extremely slippery. In addition to the simple case of an au-

15 For discussion of the ‘detached *signum*’ (e.g. INSCR. Ann. Épigr. 1914 n. 63 *Aelius Saturninus signo Calvus*) see Kajanto (1966), 57–75.

16 Keudel (1990), 12. Furthermore, when writing Latin titles in sentence case, the *TLL* usually only capitalizes proper names: e. g. the works cited as “RHET. Her.” and “VIR. ill.”. The capitalization conventions can be inferred from the *notarum explicatio* column in the *Index*.

thor's name, there are at least five additional typologies among the ca. 1,374 unique *signa* in the *Index*. These include:

1. a fictive or pseudonymous author (e.g. "CANDID.", "PS. ALEX.")
2. an original Greek author of a work translated into Latin (e.g. "GALEN.")
3. a textual genre (e.g. "DECRET." for official decrees, "EPIST." for letters)
4. a modern edition or corpus (e.g. "GLOSS." for Goetz's six-volume *Corpus glossariorum Latinorum*)
5. a writing material or medium (e.g. "TAB. cer." for wax tablets, "PAP." for papyri)

While the *signum* is ordinarily capitalized, and can consist of several space-separated words (e.g. "CAEL. AVR." for Caelius Aurelianus), there are some exceptions where an additional word that belongs to the *signum* is not actually capitalized (e.g. "SCIP. mai." and "SCIP. min." for Scipio Africanus and Scipio Aemilianus).<sup>17</sup> Another peculiarity is that the current *signa* include several instances of homonymy. For instance, there are currently three different fragmentarily preserved authors named Brutus, who are all historically distinct (and distinguished by the *Index*) but notated by the same *signum* "BRVT."<sup>18</sup> For human readers this is not a big difficulty, since the rest of the citation and *Index* provide disambiguation, but for machine processing this could be tricky.

### 2.4 Locus

The second obligatory element is the *locus*, indicating a precise location within a text. The notation depends on the way the work is structured in the standard edition. Ideally the text has been divided into logical units, such as books and chapters, that are independent of the pagination in a specific edition, but this is not always the case. When the sections are inconveniently long, sometimes the *TLL* also adds a page number (or page and line number) for convenience. Inscriptions and papyri may have columns, line numbers, and other complications. There may also be named elements in the textual hierarchy, such as a prologue ("prol.") or preface ("praef.").

Sometimes the *locus* is replaced by a cross-reference to a quoting source, labelled in fig. 2 as *testis* (e.g. "BRVT. Cic. fam. 11, 4"). The cross-reference takes the form of a complete citation, with its own *signum*, *opus* (if necessary), and *locus*, but all written in sentence case. This is a form of simple recursion. In effect it allows the *TLL* to attach a different *signum* to parts of a text or corpus that belong to different authors or textual typologies, such as embedded quotations, reported *senatus consulta*, and the like. The new *signum* usually emphasizes a feature of the embedded text that is important for its chronological placement or linguistic register.<sup>19</sup> The citation of fragments by their embedded location, rather than by a fragment number in a standard collection (when available) also helps to future-proof the *TLL*.<sup>20</sup> Fragment numbers have a tendency to change over time and not all collections are equally accessible.

---

17 By contrast, no distinction is made between the Elder and the Younger Seneca (both "SEN.") nor between the Elder and the Younger Pliny (both "PLIN."). Other cases where the *signum* arguably includes uncapitalized text: "BREV. expos.", "CVRIO avus" and "CVRIO pater", "OP. imperf. in Matth.", "ORAT. imp.", "PERVIG. Ven.", "RHET. Her.", "REG. urb.", "SERV. auct.", "SYMM. pater", "VIR. ill.", "VIRG. gramm."

18 M. Iunius Brutus (Caesar's assassin), D. Iunius Brutus Albinus, and the juriconsult M. Iunius Brutus.

19 Nevertheless, the *TLL* does not provide a distinct *signum* for every possible instance where an embedded text potentially has a different author or textual typology from its surrounding context. Much discretion is left to individual authors and editors.

20 On the complexity of citing fragmentary texts see Berti (2024).

## 2.5 *Opus*

In between the *signum* and the *locus* there is often an obligatory middle element, labelled here *opus*. It provides an important specification to the *signum*, typically identifying a single work, such as Virgil's *Aeneid* ("VERG. Aen."), as opposed to others by the same author or in the same work-group (e.g. "VERG. georg."). Where an original ancient title is unknown or impractical to use, the *opus* may be an abbreviation describing the genre (e.g. "serm." for *sermone*) or the published corpus where the relevant text is found (e.g. "gramm." for Keil's *Grammatici latini*).<sup>21</sup> Thus the *opus* element is more than just a 'work' in the strict sense. When the *signum* on its own refers unambiguously to a single work (e.g. "LIV." for Livy's *Histories* or "AETNA" for the Pseudo-Virgilian *Aetna*), there is no need to specify an *opus*.

The *opus* element is not typographically distinguished from the *locus*, nor does the current *Index* consistently identify where it starts and ends (§3.2). This sometimes creates ambiguity, which is not a problem at present, but will need to be resolved on a case-by-case basis if the citation syntax is implemented more formally. For instance, should an *appendix* ("app.") or *supplementum* ("suppl.") be regarded as a sub-section of an *opus*, and hence part of the *locus*, or a separate work entirely?<sup>22</sup> This may hinge on historical factors, such as whether the *appendix* was added by the author themselves or is an entirely separate work that became attached in transmission.

## 2.6 'Opus adiectum'

Because of these and other textual complications, the *TLL* currently acknowledges a further specification to the *opus*, here labelled *opus adiectum*. This element is signalled in the current *Index* by a second level of indentation, as discussed below (§3.2). In the figure above, "PS. AVG. serm. cod. Guelf. app." refers to a corpus of Pseudo-Augustinian sermons transmitted in a ninth-century manuscript (Wolfenbüttel, Herzog August Bibliothek, Cod. Guelf. 12 Weiss.).<sup>23</sup> While it would be possible to consider "cod. Guelf. app." part of the *locus*, this would efface the common features that set these sermons apart from other Pseudo-Augustinian sermons, such as their shared transmission and publication history. The label "*opus adiectum*" tentatively offered here is an awkward solution. It would be more accurate to say that "PS. AVG." and "serm." designate nested work-groups, and "cod. Guelf. app." functions as the specific *opus*.

## 2.7 Parenthetical *Addenda*

A final peculiarity to mention is the presence of obligatory parentheses or other *addenda* within a source citation. Typically these provide an alternative form of citation that may be superfluous, but nevertheless helps orient the reader. For instance, in citing the *Acta fratrum Arvalium*, the CIL number is always provided (e.g. "ACT. Arv. a. 105 (CIL VI 2075) II 7") and in citing the poems of Ausonius, the older Souchay numeration is given for reference (e.g. "AVSON. 27, 24 (417 S.), 124"). These obligatory parentheses, which interrupt the source citation, are different from an optional, explanatory parenthesis that sometimes follows the citation.

---

21 E.g. the treatise referred to by "CONSENT. gramm. V 385, 19" actually has the ancient title *De nomine et verbo*; in "LIV. ANDR. trag. 5" the *trag.* refers to Ribbeck's (1897) collection of fragments, instead of using the name of the tragedy, *Aegisthus*.

22 Some appendixes attached to literary corpora that are (or contain) arguably separate, anonymous compositions: e.g. "AVSON. app.", "CLAVD. carm. min. app.", "OPTAT. app.", "PHAEDR. app.". The *appendix* to Augustus's *res gestae* cannot have been written by Augustus since it was composed after his death ("R. GEST. div. Aug. app.").

23 Online facsimile and description available here: <https://diglib.hab.de/?db=mss&list=ms&id=12-weiss&catalog=Butzmann> (last access 04.04.2025). More specifically, the *opus adiectum* refers to a subset of sermons in this collection published by Morin in an appendix (1917).

### 3. The *TLL Index* of Sources as a Reference Work

To make sense of an unfamiliar source citation a reader should turn first to the *Index* online.<sup>24</sup> This is a master list of nearly all source citations currently in use, mapping abbreviations to full names and titles, as well as dates, editions, and other metadata. I say “nearly all” because in some cases, mainly for non-literary sources, only some representative citations are given rather than a complete listing of all surviving texts. Because it continues to be written in Latin, like the *TLL* itself,<sup>25</sup> explaining how it works in English should make it more accessible.

#### 3.1 History

The *Index* is a continually evolving document, illustrating changes to *TLL* practice over its 130-year history, as well as advances in Latin scholarship. Its history has determined some its most important characteristics as a repertoire, as well as some constraints. Because of the need to maintain backwards compatibility, every change to existing citations must be weighed carefully against the harm of long-term inconsistency, and a high evidentiary bar must be met before adopting new scholarly findings.

The earliest list of source abbreviations was a short document reproduced by hectograph for internal use.<sup>26</sup> Meanwhile, readers of the earliest fascicles had to rely on a single page of *notanda* for guidance.<sup>27</sup> The need for a more comprehensive list of sources was finally met in 1904 with the publication of the first edition of the *Index* as a supplementary volume to the *TLL* under Friedrich Vollmer as General Editor.<sup>28</sup> This publication established the columnar layout and the main categories of data that remain today. This first edition of the *Index* runs to 109 pages and contains approximately 3,815 entries. Three short supplements appeared over the years, which were eventually combined together in 1958 into a separate, thirteen-page *Supplementum*, incorporating additional changes.<sup>29</sup>

The second edition from 1990 involved a systematic revision of every entry and a thorough search through the literature to discover missing items.<sup>30</sup> Dietfried Krömer was responsible for the general conception and the sections A to O, while Cornelis van Leijenhorst was enlisted as co-author to complete the sections from P to Z, as well as inscriptions and papyri. Colleagues and subject experts were frequently consulted, especially H. J. Frede on Christian material. Significant changes and scholarly progress are especially apparent in the presentation of late antique texts, some of which appear for the first time or with more accurate information, while other works long accepted as ancient have been rejected.<sup>31</sup>

The current HTML presentation of the *Index* incorporates numerous *addenda quaedam indicii*, most recently updated in 2022 and downloadable separately.<sup>32</sup> Both the website, which went online in 2017

24 See n. 5.

25 For a defense of Latin in the *TLL* see Marchionni (2015).

26 Vollmer (1904); see Bögel et al. (1996), 54–55, 146.

27 The *notanda* appeared on the last page (xiv) of the *praefatio* that was enclosed with vol. I, fascicle 1 (1900), and reprinted at the beginning of vol. III (1907–1912). Though the *notanda* expand only a few abbreviations used in citations, they nevertheless formulate an important principle that source abbreviations should be formulated generously.

28 Bögel et al. (1996), 145–153. For internal and external contributors, including H. Dessau and L. Traube, see Vollmer (1904), first unnumbered page.

29 The three supplements appeared in vol. III, iv–v (1906), vol. V.1, iv–v (1910) and in a double-leaved printing in November 1936 (presumably included with fascicle VII.1.3); see Bögel et al. (1996), 154 n. 2.

30 Krömer / van Leijenhorst (1990), v–vii.

31 On the changes to late antique material see especially Krömer (2003) and the review by Doignon (1993).

32 The *addenda*, chosen by the *TLL* Editors, are available as a separate PDF file: [https://thesaurus.badw.de/fileadmin/user\\_upload/Files/TLL/addenda.pdf](https://thesaurus.badw.de/fileadmin/user_upload/Files/TLL/addenda.pdf) (last access 04.04.2025). These include newly

with support from the IT Department of the Bavarian Academy, and the *addenda* are maintained by Josine Schrickx, current General Editor, and Marijke Ottink. The electronic version also includes its own directions for use.<sup>33</sup> Meanwhile the on-going work transforming the *Index* into a database has been overseen by the *TLL* Digital Workgroup, coordinated by Nora Götze.

## 3.2 Using the Online *Index*

To understand the contents of the *Index*, we can turn to a section from the beginning of the HTML table:

Zurück zur Projektseite		Hinweise zur Benutzung		Sprung nach: A B C D E F G H I L M N O P Q R S T V Z	
A	aetas	notae	notarum explicatio	editiones	
	cos. 331. † 338 post 326	ABLAB. epigr. 2	Abliabius Constantini Magni familiaris epigramma a Sidon. epist. 5, 8, 2 traditum (vix genuinum), vers. 2 epistula ad Orcistenos notatur Epist. praef. praet. (CIL III 7000)	FPL Morel (1927) p. 159; cf. FPL Buechner (1982) p. 190	
171	versa non post 484?	ACAC. epist. Ver. 4 p. 5, 22	Acacii patriarchae Constantinopolitani ad Simplicium papam epistula (scripta 477/478), cuius versio latina (graeca perit) collectione Veronensi (epist. 4) servatur, p. 5 lin. 22 B: SIM 8	Schwartz, Abh. Münch. Ak. NF 10 (1934) p. 4 sq.	
	* 170 a. Chr., † saec. II <sup>n</sup> .	ACC. carm. frg. 24, 2	L. Accius ex Vmbria Pisaurensis carminum (praeter scaenica) fragmenta, fragm. 24 Morel (olim 26 Baehrens) vers. 2 nonnulla fragmenta quidam oratione soluta conscripta esse iudicant	FPR Baehrens (1886) p. 266–271; FPL Morel (1927) p. 34–39; cf. FPL Buechner (1982) p. 46–51	
1		ACC. carm. frg. 24, 2	praetext. 41	TRF Ribbeck ( <sup>3</sup> 1897) p. 326–331	
2			trag. 701	TRF Ribbeck ( <sup>3</sup> 1897) p. 157–262; cf. D'Antò (1980)	
			--	Ps. Acro v. Schol. Hor.	
J			Act.	acta p a g a n a inscriptionibus tradita	
	21 a. Chr.–241 p. Chr.	Arv. a 105 (CIL VI 2075) II 7 a. 27 (CIL VI 2024) f 8 Dom. (CIL VI 2071) l 5	[Arv. Dom. D l 5]	CIL VI 2023–2119. 32338–32398. 37164 sq.; cf. Pasoli (1950)	
		ACT. Arv. a. 105 (CIL VI 2075) II 7		ad ann. 105, col. II lin. 7 ad ann. 27, fragm. f lin. 8 aetatis imp. Domitiani, (fragm. D) col. I lin. 5	
				titulos recentius repertos similiter afferimus, sc. indicatis aetate ac editione, e. g. Act. Arv. a. 53 (Année Epigr. 1977 n. 18)	
				Pighi, De lud. saec. (1941; c. addend. 1965)	
			lud. saec.	ludorum saecularium	

Fig. 3: *Index* online, from ABLAB. to ACT. lud.saec.

For most users, the most important column is the one labeled *notae*, providing sample abbreviations, some of which are highlighted in yellow, such as “ABLAB. epigr. 2”, “ACAC. epist. Ver. 4 p. 5, 22”, “ACC. carm. frg. 24, 2”, “ACT. Arv. a. 105 (CIL VI 2075) II 7”. As the illustration shows, a complete citation is formed from combining the text of the capitalized *signum* with all the text in the relevant indented lines below it until one reaches a *locus*.

It is important to note the presence of two levels of indentation within the *notae* column distinguishing between the *opus* and *opus adiectum* elements. Where there is a single indentation, this indicates the beginning of an *opus* element (e.g. before “epigr.”, “carm. frg.” and “Arv.”). Where there is a double indentation, this either marks the beginning of an *opus adiectum*, such as before “a. 105 (CIL VI 2075) II 7”, or emphasizes different types of *loci* belonging to the same *opus*. Nevertheless, some citations are compressed on a single line, such as “ACAC. epist. Ver. 4 p. 5, 22”, and in these cases the division among the elements is ambiguous.

The sample citation *always* includes a complete *locus* so that it is clear which numeration or system of reference the *TLL* follows. This is extremely important in order to avoid confusion among the many competing systems of capitulation sometimes present within a single edition. The sample *locus* is usually the *last citable section of the work* since it is at the end of a work, rather than the beginning, where differences in capitulation are most evident.

published ancient Latin works and changes to the citation format of previously known works; new editions of known texts are not added unless they prompt the *TLL* to change its citation format.

33 <https://thesaurus.badw.de/en/tll-digital/index/directions-for-use.html> (last access 04.04.2025).

The full expansion of each element in the abbreviation is found in the adjacent column, labeled *notarum explicatio*. This column also contains additional information that may be relevant to understanding and finding the text in question. For instance, the citation “ACAC. epist. Ver. 4 p. 5, 22” is expanded as follows:

*Acacii patriarchae Constantinopolitani ad Simplicium papam epistula (scripta 477/478), cuius versio latina (graeca periit) collectione Veronensi (epist. 4) servatur; p. 5 lin. 22.*

Letter by Acacius, Patriarch of Constantinople to Pope Simplicius (written in 477–478), of which a Latin translation (the Greek has been lost) is preserved in the *collectio Veronensis* (letter 4), page 5, line 22.

Each component of the abbreviation (“epist. Ver.”) is expanded, and the numbers in the *locus* (“4 p. 5, 22”) are glossed. Furthermore, important contextual details in this case are provided about the text’s addressee, its status as a translation from a lost Greek original, and its original date of composition (which differs from the date of its translation, indicated in the *aetas* column).

Incidentally, the use of the genitive case for the proper name, *Acacii*, in the *notarum explicatio* is significant. Whereas the nominative in apposition is regularly used to indicate genuine authorship (e.g. “Ablabius”, “L. Accius”), the genitive usually indicates a more complex relationship, such as authorship of the original Greek text or pseudonymity.

### 3.3. Metatextual Information in the Index

In the figure below I have identified six kinds of metatextual data that are consistently or occasionally present:

A	aetas	notae	notarum explicatio	editiones
	cos. 331. † 338	ABLAB.	Ablabius Constantini Magni familiaris	
1	post 326	epigr. 2	epigramma a Sidonius	FPL Morel (1927) p. 159; cf. FPL Buechner (1982) p. 190
		–	epistula ad Orcis	
171	versa non post 484?	ACAC. epist. Ver. 4 p. 5, 22	Acacii patriarchae 477/478), cuius servatur, p. 5 lin. 22	Schwartz, Abh. Munch. Ak. NF 10 (1934) p. 4 sq.
	* 170 a. Chr., † saec. I <sup>n</sup> .	Acc.	L. Accius ex Vmbria Pisauraensis	
1		carmin. fig. 24, 2	[carmin. fig. 26, 2] carminum (praeter scaenica) fragmenta, fragm. 24 Morel (olim 26 Baehrens) vers. 2	FPR Baehrens (1896) p. 266–271; FPL Morel (1927) p. 34–39; cf. Buechner (1982) p. 46–51
2		41	fragmenta quidam oratione soluta conscripta esse iudicant	
2			rum fragmenta, vers. 41	TRF Ribbeck ( <sup>3</sup> 1897) p. 326–331
			rum fragmenta, vers. 701 Ribbeck	TRF Ribbeck ( <sup>3</sup> 1897) p. 157–262; cf. D’Antò (1980)
		–	Ps. Acro v. SCHOL. Hor.	
J		Act.	acta p a g a n a inscriptionibus tradita	
	21 a. Chr.–241 p. Chr.	Arv.		CIL VI 2023–2119. 32338–32398. 37164 sq.; cf. Pasoli (1950)
		a. 105 (CIL VI 2075) II 7		
			ad ann. 27, fragm. f lin. 8	
		[Arv. Dom. D I 5]	aetatis imp. Domitiani, (fragm. D) col. I lin. 5	

Fig. 4: Metatextual data in the Index.

Starting on the left, the first category is the *Ordnungsnummer*, found in the first column (labeled “A”). This number, which is important for ordering passages in a *TLL* article, provides an approximate relative chronological position for each author or text, starting with 1 for fragments of early Roman poetry and now ending at 226 for some of the latest, sixth- or seventh-century cited texts (e.g. “CHILP. hymn. Medard.”). The numbers also correspond to the shelfmark in the *TLL* Library.

Second, the date of the author or work is found in the second column (*aetas*) where available. As even this short sample shows, the kind of dating and its precision vary considerably based on the available

information.<sup>34</sup> For instance, the author Ablabius (“ABLAB.”) is identified only by the date of his consulship (331) and death (338), but the epigram Sidonius attributes to him is dated more precisely “after 326”.<sup>35</sup>

The third category are out-of-date source abbreviations and numeration, used in older volumes of the *TLL* but now superseded. An out-of-date abbreviation always appears between square brackets (“[carm. frg. 26, 2]”). The third column (the lefthand side of *notae*) contains *all* abbreviations, whether current or not. When an out-of-date abbreviation appears in the third column, the current abbreviation appears in the fourth column (the righthand side of *notae*); when a current abbreviation appears in the third column, an out-of-date equivalent may be listed in the fourth column.<sup>36</sup> As the *notarum explicatio* makes clear, the fragments of Accius (“ACC.”) were originally numbered according to Baehrens’s edition, but now according to Morel.

Cross-references of various kinds constitute the fourth kind of metadata. For instance, a reader might expect to find the Pseudo-Acronian scholia to Horace listed under ‘PS. ACRO’. Accordingly the name ‘PS. ACRO’ is inserted alphabetically in the *notarum explicatio* column between ‘ACC.’ and ‘ACT.’, but the reader is advised to find the text under ‘SCHOL. Hor.’. The presence of a cross-reference is indicated in the *nota* column by double dashes (‘--’), when it refers to the *signum* level, and by a single dash (‘-’), when it refers to an *opus*. There are about 820 cross-references in the current *Index*.

The fifth kind is authority control information (*Normdaten*). This has become extremely important in information science as a way to guarantee that names, especially used across platforms in linked open data, refer persistently to the same person or object.<sup>37</sup> The second edition of the *Index* already provided a rudimentary form of this: whenever an italicized letter “*B*” appears in the *notarum explicatio* column, it is followed by the abbreviation for the author in the Beuron *Kirchenschriftstellerliste*.<sup>38</sup> This repertoire explicates the standard source abbreviations used in the monumental Beuron critical edition of the *Vetus Latina*, and as such often provides more detailed information about late antique Christian sources. In turn, the *Kirchenschriftstellerliste* provides the corresponding *TLL* abbreviations where possible. This system of mutual cross-reference is a form of authority control that guarantees long-term compatibility, and it also helps the reader find additional information about the relevant author or work.

The last kind of metadata, and for classicists perhaps the most valuable, is the list of reference editions provided in the rightmost column (*editiones*). The list is up-to-date until 1990; after 1990 only sporadic additions have been made, principally of editions that have altered the *TLL* source abbreviation or *locus*. Accordingly, many significant, post-1990 editions are currently absent from this list, but they are expected to be added soon. The most recent edition before the “cf.” (if there is a “cf.”) typically provides the capitulation and numeration of citations; other important editions may be added after a “cf.”. Altogether there are over 3,100 unique editions listed in this column.

The use of multiple editions is essential for *TLL* practice: while it is necessary to use a single reference edition for capitulation, the text as quoted in the *TLL* articles is meant to be the most plausible text, regardless of whether it is found in an edition or not. This sometimes means making a choice among tex-

---

34 Krömer (2003) discusses some issues relating to dating.

35 The more precise dating is presumably based on the historical situation described by Sidonius. Note also that the *Index* specifies parenthetically that the poem is *vix genuinum*. This lampoon on the emperor, posted up on the palace wall, seems unlikely to have been written by a consular, despite Sidonius’s report. Nevertheless, the *signum* remains “ABLAB.” rather than “PS. ABLAB.”.

36 On this innovation introduced by the *Index* second edition, see Krömer / van Leijenhorst (1990), v.

37 On the exemplary use of authority control in the *Perseus Catalog*, see Babeu (2019).

38 The most recent, fifth edition: Gryson (2007). H. J. Frede, who worked on the third and fourth editions (1981 and 1995 respectively), was in close contact with the editors of the *TLL Index* while both works were under revision.

tual variants, citing conjectures published elsewhere, or even suggesting a new emendation where necessary.

### 3.4 Graeca

One would not expect to consult an index of Latin sources for information about Greek texts, but there is a surprising amount of information in the *TLL Index* for Hellenists.<sup>39</sup> First, there are a few Greek-language texts scattered within the *Index* as primary sources occasionally cited by the *TLL*: notably, the sixth-century lexicographer Hesychius (“HESYCH. lex.”) and Iohannes Lydus (“LYD.”), a sixth-century scholar who preserves many etymologies and discussions of Latin words.<sup>40</sup> Secondly, and more significantly, there are over two hundred Greek texts referenced in the *editiones* column whenever an ancient Latin text was either translated from Greek or possibly translated into Greek. These are preceded by the abbreviation “gr.” when the Greek text is known to be source or “cf. gr.” when the Greek is potentially later or the precise relationship has not been sufficiently established. The systematic collection and evaluation of relevant Greek texts was one of the achievements of the second edition of the *Index*, and it offers a tantalizing possibility for future research into ancient bilingualism or even large-scale digital language alignment of Greek and Latin texts.

### 3.5 Visualizing the *Index*

Finally, to give an idea of the scale of the ancient Latin corpus depicted by the *Index*, I provide some graphic representations of the number of citeable works across time. By “citeable work”, I simply mean a unique combination of *signum*, *opus*, and *opus adiectum*. Altogether I count 2,880 of these, divided out over ca. 1,035 *signa*.

The results need to be treated with caution. First, a ‘citeable work’ varies significantly in scale: for instance, the entire *CIL* (*Corpus inscriptionum Latinarum*) counts as one citeable work, as does a two-line epigram attributed to Ablabius (“ABLAB. epigr.”). Second, the *Index* does not always provide a date, and where it does, it may be approximate and/or uncertain.<sup>41</sup> To convert these dates into centuries I have treated uncertain cases as certain and converted any Arabic number or Roman numeral into a corresponding century (using the lower of two alternatives where provided). Third, some of the entries in the *Index* provide only a few sample works rather than an exhaustive listing of all known works.<sup>42</sup> Thus the number of works in the *Index* is less than the actual number of works cited in the *TLL*.

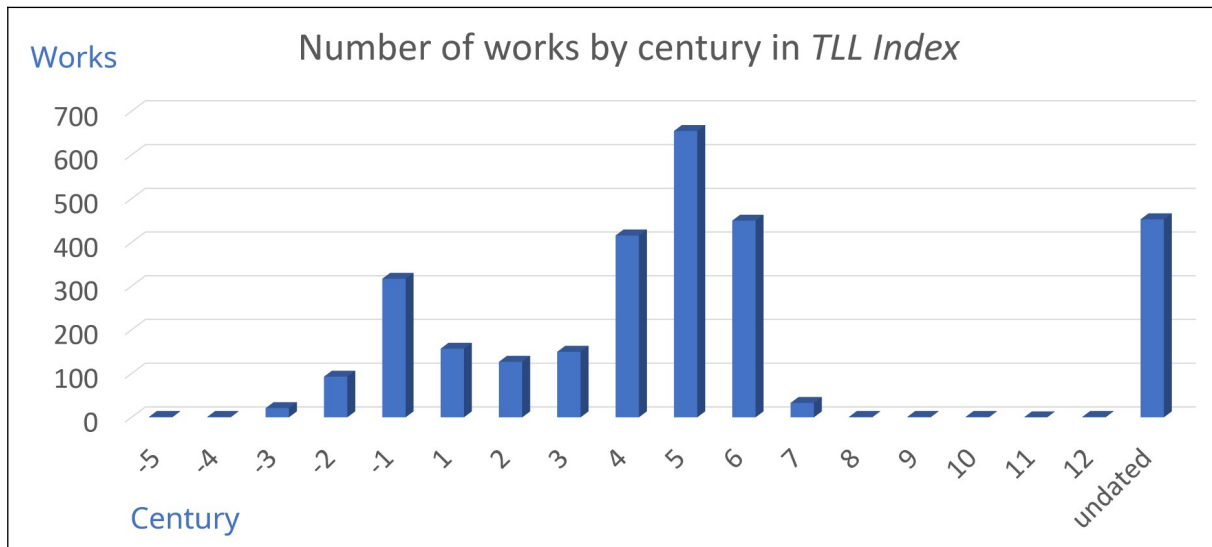
---

39 On Hellenists using the *TLL* see Vogt (1995).

40 Though not mentioned in the *Index*, Polybius (“POLYB.”) is cited in the article *praefectus* (X 2, 624, 20; van Leijenhorst 1985).

41 There are about 453 citeable without dates, either on their own or derivable from their parent *signum*. These can mostly be assumed to be late.

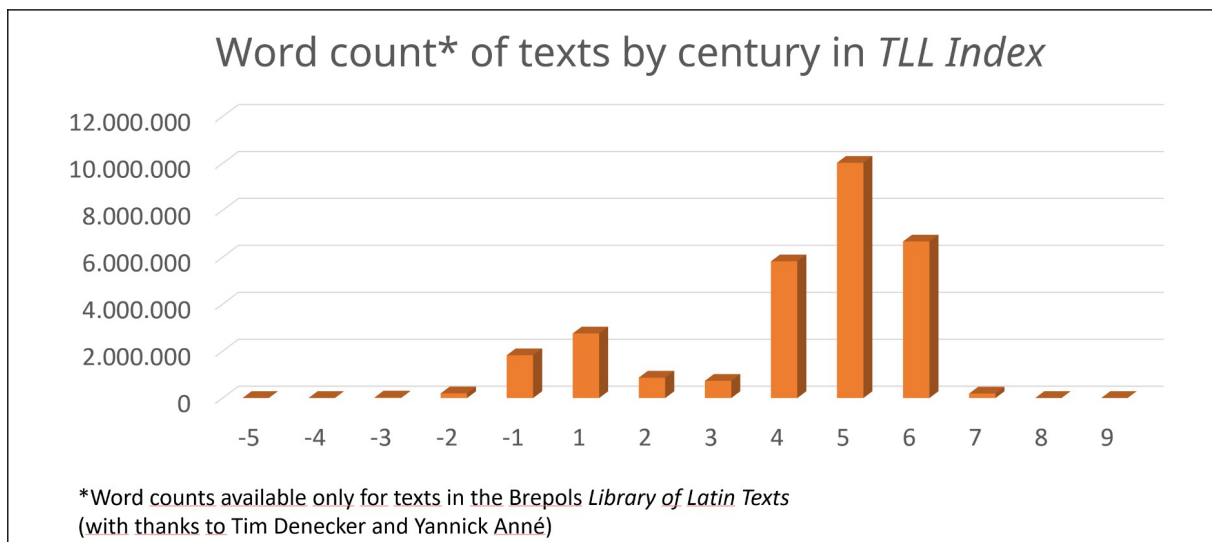
42 For instance, under “DECRET. decur.” only some sample decurional decrees are given; listing all such texts would take up too much space. Other such *Mustereinträge* can be found by searching the *Index* for *similiter*.



**Fig. 5: Citable works in the TLL Index by century.**

It is no surprise to see an abundance of material for late antiquity: if everything before 200 CE is considered ‘classical’, there are nearly 2.5 times as many works that survive from late antiquity. What did surprise me was to see that the contents did not drop off completely after 600 CE: the *Index* in fact includes 33 works from the seventh century, and one each from the eighth (PAVL. FEST.), ninth (BASILIC.), tenth (CONSTANT. PORPH.), and twelfth centuries (*Osbern.*). I encourage readers who are curious to look these up in the *Index*.

A more effective representation would involve word count. Since the *Index* does not include any word counts, I have relied on the figures generously provided by Brepols’ *Library of Latin Texts (LLT)*. Their database includes *TLL* source abbreviations wherever possible, which makes alignment easier. But not all the texts in the *Index* are in *LLT*, especially inscriptions and other non-literary documents. Moreover, even when a text is in both corpora, it is often difficult to align them automatically since each database may have made different choices about authenticity, text division, and what constitutes an independent work.



**Fig. 6: Word count of citable works in the TLL by century.**

The resulting distribution remains somewhat similar, with approximately three times as many words from late antiquity as from the classical period. The most dramatic difference is the relative drop in the first century BCE. This certainly has to do with the number of fragmentary texts from this period: very many citable works but with a very low word-count. Given the inconsistencies in the data, these fig-

ures should be regarded as an *amuse-bouche* for what can be derived from future digitization rather than a reliable representation of the *TLL* textual corpus.

### 4. Future Prospects

The *Index* is currently being migrated to a new relational database as part of a thorough digitalization initiative within the *TLL* that aims to increase access and functionality. Currently work on the *Index* has focused on updating the digital architecture rather than making changes to the content. It is still too early to share details about implementation. The most urgent priority is integrating the *Index* with a new digital workflow for producing *TLL* articles in XML so that lexicographers can quickly cite the correct source abbreviations and ensure their accuracy. Nevertheless, this transformation also opens the prospect for significant improvements for external users and the linked open data community at large.

By way of conclusion I will sketch here some of the main areas of future work as well as some open questions and desiderata. Not all of these features will be possible to implement in the near future, or perhaps at all, but discussing these issues in this public forum can generate awareness and excitement among the larger scholarly public and solicit feedback from interested parties at a formative moment of development.

#### 4.1 Searchability and Complex Queries

A central desideratum is improving the searchability of the *Index* online. Currently a full-text search of the HTML table is the only way of querying the data, which is unwieldy when looking for short snippets of text. In a new relational data structure, it will be possible to perform targeted queries for text in specific elements: for instance, to find all *signa* containing the string “AVG” or to identify all texts certainly (or possibly) composed after the year 410. Indeed it should be possible to combine these queries into a complex search in order to generate a list of Augustine’s works composed after 410.

For these searches to work effectively, some of the orthographic conventions of the *TLL* must be taken into account automatically: for instance, to recognize that an upper-case “V” in “AVG” can correspond to a lower-case “u” or “v” depending on the context. More generally, the natural-language dates provided in the *aetas* column for each author and work will need to be converted into a machine readable format, such as the Extended Date Time Format (EDTF) specification of the Library of Congress.

Furthermore, it will be necessary for some ambiguous elements in current source abbreviations to be clearly sorted into the relevant categories (*signum*, *opus*, *opus adiectum*, and *locus*, assuming the analysis proposed above is viable). As we have seen (§3.2), the presentation of the sources in the *Index* does not always make clear how to analyze a compound citation such as “ACAC. epist. Ver. 4 p. 5, 22”. Fortunately, the Digital Workgroup has already gone through the entire *Index* and made preliminary assignments, but more internal testing and proofreading will be necessary to guarantee consistency across the data.

Once the *Index* in its current state has been successfully migrated to a new database, it will be possible to consider adding additional categories of information, such as tags to identify the genres of individual works, along the lines of the *TLG Canon*, or even information about the modality of transmission, which is currently implicit in some but not all of the source abbreviations (e.g. the *signum* “PAP.” or the abbreviation “frg.” in the *locus*). It would also be desirable for all this data to be accessible via an Application Programming Interface (API) to facilitate re-use.

## 4.2 Linked Open Data

As the most complete and up-to-date repository of information about ancient Latin texts, the *Index* can play an important role as a digital clearing-house, supplementing and refining the data already made available by trailblazers in Classics linked open data, such as *Perseus*, *Trismegistos*, and *Wikidata*. On the one hand, this would require linking *Index* entries outward with the relevant records in external databases. On the other hand, it would require making *Index* entries externally linkable by assigning persistent identifiers to all entries at the *signum* and *opus* levels.

An open question remains the form such persistent identifiers should take. The *TLG Canon* provides a useful model: in their case this is a bipartite structure consisting of a four-digit ‘author’ ID and three-digit ‘work’ ID (e.g. 0012.001 represents Homer’s *Iliad* and 3123.001 represents Theodorus Daphnopates’s *Epistulae*). The author IDs also encode some chronological information: IDs in the range 1–2000 are early/classical, the 3000 range is mostly for Byzantine authors.<sup>43</sup> Since *TLL* names are occasionally tripartite (when there is an *opus adiectum*), this complication would have to be met either by flattening out the hierarchy, and thereby losing information about textual affinities, or by tolerating three positions. Furthermore, the identifiers must be flexible enough to accommodate the possibility that authorship attributions and datings may change in future scholarship.

In addition to linking to other Latin digital catalogs, the most important targets for linkage will be open access Latin editions. This will facilitate quick access to relevant texts and could eventually provide the basis for translating *TLL* source citations articles automatically into hyperlinks. Currently there is a growing amount of ancient Latin material available online, thanks to collections such as *Corpus Corporum*,<sup>44</sup> *digilibLT*,<sup>45</sup> the *Digital Latin Library*,<sup>46</sup> and the *Perseus Digital Library*.<sup>47</sup> But in order to be in the public domain, many of the texts are not the most up-to-date. Meanwhile, through a pilot project led by Massimo Cè, a former *TLL* researcher, it was possible to map approximately 30% of the editions referenced in the *Index* to PDFs in the public domain (i.e. out of European copyright or available in open access). It remains a great desideratum to see more high-quality editions in open access, ideally with a critical apparatus, or at least with an XML structure compatible with the *TLL* source citations.

## 4.3 Source Citations via Canonical Text Services (CTS)

A standardized, machine readable format for *TLL* source citations is necessary not just for composing future *TLL* articles in XML, but to encourage wider adoption of these reference standards, and in the future to create links from citations to digital texts.

For encoding a complete citation, including the *locus*, the Canonical Text Services (CTS) model developed by the *Homer Multitext* project provides an attractive solution.<sup>48</sup> It has already been adapted to the citation of Latin texts in the *Perseus Catalog* using the Packard Humanities Institute (PHI) and Stoa identifiers.<sup>49</sup> A CTS reference is a Universal Resource Name (URN) with the following basic structure: urn:cts:namespace:work:part. For instance, using the greekLit namespace,<sup>50</sup> a citation of the first line of the *Iliad* can be represented as: urn:cts:greekLit:tlg0012.tlg001:1.1. CTS allows for the in-

43 Pantelia (2022), xxiii n. 4.

44 <https://mlat.uzh.ch/> (last access 23.04.2025).

45 <https://digiliblt.uniupo.it/> (last access 23.04.2025).

46 <https://digitallatin.org/> (last access 23.04.2025).

47 <https://www.perseus.tufts.edu/> (last access 23.04.2025).

48 Blackwell / Smith (2019).

49 <https://catalog.perseus.org/> (last access 22.04.2025); see Babeu (2019), 55.

50 Maintained by the Center for Hellenic Studies: <https://github.com/chs-tg/greekLit> (last access 23.04.2025).

clusion of an optional ‘version’, such as the representation of the text in a specific manuscript or edition, as well as the citation of the ‘part’ according to complex textual hierarchies, such as ‘1.proposition.20’ in Euclid’s *Elements*.

For the most part CTS syntax can be easily adapted to *TLL* source citations. Indeed, one can even provide users with a choice between numerical identifiers or human-readable names, such as the following two representations of VERG. Aen. 12, 952:

```
urn:cts:tll:sign04691.op0003:12.952
urn:cts:tll:Verg.Aen:12.952
```

Nevertheless, there are potential conflicts and limitations of CTS syntax for which the *TLL* would have to find custom solutions. First, on the *TLL* side, there is the problem of homonymous *signa* discussed above (§2.3). Ideally these should be distinguished in a machine readable way. Second, some of the more complex *TLL* source citations discussed above (§2.4) require recursion or the coordination of multiple citations. Some new notation might have to be developed to represent these relationships. For instance, to cite a letter within Cicero’s epistolary corpus as having been written by Brutus (fig. 2), one might imagine the following syntax: “urn:cts:tll:Brut||urn:cts:tll:Cic.fam.11.4”. Here the double pipes (||) represent the relationship ‘is a description, synonym, or specification for’.

An even more complex case is the Latin Oribasius (e.g. “ORIBAS. eup. 2, 1 A 24 La p. 485”). Here there are three different layers (‘versions’ or ‘manifestations’?) that can and sometimes must be distinguished after the work level: an individual recension of a work, a manuscript containing a recension, and a modern edition reporting the manuscript or recension. Moreover, the *TLL* cites in parallel both the logical capitulation of the work and its pagination in the only available edition. Perhaps it would have to look something like:

```
urn:cts:tll:Oribas.eup.La.La:2.1.A.24||urn:cts:tll:Oribas.eup.Molinier:pag.485
```

The URN before the pipes is a citation of the manuscript La (Laon Ms. 424, s. ix),<sup>51</sup> which belongs to the eponymous recension La; the URN after the pipes reports the same text according to Molinier’s edition (1873–1876). Two citations are needed here because Molinier’s pagination – still necessary for readers to find the text and compare it with that of other recensions – is not fully coterminous with or nestable within the logical capitulation.

More generally, citations in the *TLL* do something subtly different than what canonical citations are meant to do. The ‘reference’ edition is used as a *basis* for the capitulation and text, but the lexicographer ultimately quotes the text using all available evidence as they believe it to be correct (see §3.3). In other words, the lexicographer tries to print the Latin Oribasius, not Molinier’s version of it. Accordingly, perhaps there needs to be a flag (a big exclamation point?) included as part of an URN, so that the user knows that the edition is referenced, but not necessarily quoted verbatim.

### 4.4 Envoi

After over 130 years of work, the *TLL* prepares to enter a new phase of digital life. The source citations and *Index* are only a small part of this, but they sit at a crucial meeting place where article production interacts with texts, names of works are encoded and decoded, dates are assigned, and external users come with queries of their own. There is still much work to do, and many of the developments will be gradual. Nevertheless, it will be possible, not only because of the dedication of the *TLL* team, but in part thanks to the digital trailblazers who have shown the way.

---

51 <https://bibliotheque-numerique.ville-laon.fr/idviewer/1468/7> (last access 22.04.2025).

## List of Abbreviations

- OLD*: P. G. W. Glare (ed.), Oxford Latin Dictionary, 2<sup>nd</sup> edition, Oxford 2012.
- TLG*: Thesaurus linguae Graecae. <http://stephanus.tlg.uci.edu> (last access 04.04.2025)
- TLL*: Thesaurus linguae Latinae, Leipzig / Berlin 1900–. <https://tll-open.badw.de/> (last access 04.04.2025)

## References

- Andrieu (1940): J. Andrieu, Étude critique sur les sigles de personnages et les rubriques de scène dans les anciennes éditions de Térence, Paris 1940.
- Audollent (1904): A. M. H. Audollent, Defixionum tabellae quotquot innotuerunt, tam in Graecis Orientis quam in totius Occidentis partibus praeter Atticas in Corpore inscriptionum atticarum editas, Paris 1904.
- Babeu (2019): A. Babeu, The Perseus Catalog: of FRBR, Finding Aids, Linked Data, and Open Greek and Latin, in: M. Berti (ed.), Digital Classical Philology: Ancient Greek and Latin in the Digital Revolution, Berlin / Boston 2019, 53–72.
- Baratta (2019): G. Baratta, Benest, malest: Archeologia di un gioco tardo-repubblicano, Barcelona 2019.
- Berti (2024): M. Berti, Digital Canons and Catalogs of Fragmentary Literature, in: F. Neuerburg et al. (eds.), Fragmente einer fragmentierten Welt: Zur Problematik des Umgangs mit Fragmenten in der gegenwärtigen klassisch-philologischen Forschung, Berlin / Boston 2024, 217–236.
- Blackwell / Smith (2019): C. W. Blackwell / N. Smith, The CITE Architecture: a Conceptual and Practical Overview, in: M. Berti (ed.), Digital Classical Philology. Ancient Greek and Latin in the Digital Revolution, Berlin / Boston 2019, 73–93.
- Bögel et al. (1996): D. Krömer / M. Flieger (ed.), Thesaurus-Geschichten: Beiträge zu einer Historia Thesauri linguae Latinae von Theodor Bögel, Stuttgart / Leipzig 1996.
- Corbeill (2007): A. Corbeill, “Going Forward”: A Diachronic Analysis of the Thesaurus linguae Latinae, American Journal of Philology 128 (2007), 469–496.
- Di Stefano (2011): A. Di Stefano (ed.), Arusiani Messi: Exempla elocutionum, Hildesheim 2011.
- Doignon (1993): J. Doignon, Review of Krömer / van Leijenhorst (1990), Revue d’Études Augustiniennes et Patristiques 39 (1993), 225–226.
- Flow (2015): C. Flow, Thesaurus Matters: Frames for the Study of Latin Lexicography, in: C. Stray and G. Whitaker (ed.), Classics in Practice: Studies in the History of Scholarship, BICS Supplement 128, London 2015, 33–73.
- Flury (1995): P. Flury, Vom Tintenfaß zum Computer, in: D. Krömer (ed.), Wie die Blätter am Baum, so wechseln die Wörter: 100 Jahre Thesaurus linguae Latinae, Stuttgart 1995, 29–56.
- Grabowski (2024): A. Grabowski, The Craft of History: Turning History into a Discipline in the Twelfth and Thirteenth Centuries, Turnhout 2024.
- Grondeux (2015): A. Grondeux, Le traitement des ‘autorités’ dans le Liber glossarum (s. VIII), Eruditio antiqua. revue électronique de l’érudition gréco-latine 7 (2015), 71–95.
- Grondeux (2023): A. Grondeux, Comment définir un ‘dictionnaire latin’? Du Liber glossarum (VII<sup>e</sup> s.) à l’Elementarium de Papias (XI<sup>e</sup> s.), Histoire Épistémologie Langage 45/2 (2023), 15–33.

- Gryson (1999–2004): R. Gryson, *Altlateinische Handschriften: Manuscrits vieux latins: répertoire descriptif*, Freiburg 1999–2004.
- Gryson (2007): R. Gryson, *Répertoire general des auteurs ecclésiastiques latins de l’antiquité et du haut moyen âge*, 5<sup>th</sup> edition, Freiburg 2007.
- Kajanto (1966): I. Kajanto, *Supernomina: A Study in Latin Epigraphy*, Helsinki 1966.
- Keudel (1990): U. Keudel, *Thesaurus linguae Latinae: Praemonenda de rationibus et usu operis*, Leipzig 1990, <https://tll.degruyter.com/help> (last access 04.04.2025).
- Krömer (2003): D. Krömer, *Don’t Trust the Label: Zur Datierung lateinischer Texte*, in: H. Solin / M. Leiwo / H. Halla-aho (ed.), *Latin vulgare – latin tardif: Actes du Vie colloque international sur le latin vulgare et tardif*, Hildesheim et al. 2003, 183–189.
- Krömer / van Leijenhorst (1990): D. Krömer / C. G. van Leijenhorst, *Thesaurus linguae Latinae: Index librorum scriptorum inscriptionum ex quibus exempla afferuntur*, 2<sup>nd</sup> edition, Teubner 1990.
- Marchionni (2015): R. Marchionni, *Latein als Sprache des Thesaurus linguae Latinae*, *Akademie Aktuell* 53 (2015), 54–57, [https://badw.de/fileadmin/pub/akademieAktuell/2015/53/0215\\_13\\_Marchionni\\_V05.pdf](https://badw.de/fileadmin/pub/akademieAktuell/2015/53/0215_13_Marchionni_V05.pdf) (last access 24.04.2025).
- Molinier (1873–1876): A. Molinier (ed.) in: U. C. Bussemaker / C. Daremberg (ed.), *Œuvres d’Oribase*, voll. 5 and 6, Paris 1873–1876.
- Morin (1917): G. Morin (ed.), *Sancti Aureli Augustini tractatus sive sermones inediti ex codice Guelferbyitano 4096*, Kempten / Munich 1917.
- Oomes (1968): F. Oomes, *Nomina ominosa*, in: *Lemmata: Donum natalicium Guilelmo Ehlers sexagenario a sodalibus Thesauri linguae Latinae oblatum*, München 1968, 221–231.
- Pantelia (2022): M. C. Pantelia, *Thesaurus Linguae Graecae: A Bibliographic Guide to the Canon of Greek Authors and Works*, Oakland (CA) 2022.
- Ribbeck (1897): O. Ribbeck (ed.), *Scaenicae Romanorum poesis fragmenta*, third edition, Leipzig 1897.
- Scappaticcio (2019): M. C. Scappaticcio, *Testi latini su papiro e lessicografia. In margine ad un contributo possibile al Thesaurus Linguae Latinae*, *BStudLat* 49 (2019), 689–698.
- Vollmer (1904): F. Vollmer, *Thesaurus linguae Latinae: Index librorum scriptorum inscriptionum ex quibus exempla adferuntur*, Leipzig 1904.
- Vogt (1995): E. Vogt, *Ein Gräzist benutzt den Thesaurus*, in: D. Krömer (ed.), *Wie die Blätter am Baum, so wechseln die Wörter: 100 Jahre Thesaurus linguae latinae*, Stuttgart / Leipzig 1995, 99–109.

## Author Contact Information<sup>52</sup>

Dr. Adam Gitner

Thesaurus linguae Latinae / Bayerische Akademie der Wissenschaften

Alfons-Goppel-Str. 11

80539 München

E-mail: [agitner@thesaurus.badw.de](mailto:agitner@thesaurus.badw.de)

---

<sup>52</sup> The rights pertaining to content, text, graphics, and images, unless otherwise noted, are reserved by the author. This contribution is licensed under CC BY-SA 4.0.