

# Exploring Nominal Agreement in Early Medieval Charters through the LLDB Database: Methods and a Pilot Sample from Tuscany (785–795)

ELISA D'ARGENIO<sup>1,2\*</sup> 

<sup>1</sup> Ca' Foscari University of Venice, Italy

<sup>2</sup> ELTE Research Centre for Linguistics, Hungary

## RESEARCH ARTICLE

Received: January 2, 2026 • Accepted: February 14, 2026

© 2026 The Author(s)



### ABSTRACT

This paper explores the LLDB database as a resource for investigating nominal agreement in eighth-century Latin charters, based on a pilot sample of 48 documents drafted in the Lucca area (Tuscany) between 785 and 795. Methodologically, mismatches are identified through the LLDB coding system and then checked qualitatively, especially where agreement interacts with other types of deviation. The analysis shows that nominal agreement mismatches are unevenly distributed across the texts and tend to cluster in high-frequency formulaic sequences. A second recurrent factor is their interaction with very common phonographic phenomena and with lexemes showing case-form rigidification in medieval sources. The paper concludes that nominal agreement variation may contribute to assessing scribal linguistic competence within a predominantly memory-based drafting process, provided that frequency and formulaic routinisation are explicitly taken into account.

### KEYWORDS

nominal agreement, eighth-century Latin charters, LLDB database, formulaic language, documentary Latin

\* Corresponding author. E-mail: elisa.dargenio@unive.it

## 1. RESEARCH QUESTION

Within the framework of the DiLaDi project,<sup>1</sup> this paper explores the potential of the *Computerized Historical Linguistic Database of the Latin Inscriptions of the Imperial Age* (henceforth LLDB database) – specifically its section devoted to the eighth-century Latin private-law charters –<sup>2</sup> as a resource for investigating grammatical agreement.<sup>3</sup> This is a preliminary study and draws on a small, well-delimited sample, with the aim of testing methods and identifying challenges for analysis.

Our central question is whether the realisation or non-realisation of nominal agreement between a controller (the nominal head) and targets (modifiers such as adjectives, appositions, nominal forms of verbs) can serve as an indicator of notarial linguistic competence, with particular attention to the dynamics of the scribes' writing process.<sup>4</sup>

In practical terms, scribes draft charters by drawing on an overarching documentary template (e.g., *chartula*, the most frequent type, *notitia*, *praeceptum*)<sup>5</sup> and on transaction-specific sub-templates (e.g., *chartula venditionis*, *chartula donationis*). These templates are organised around fixed textual structures (opening and closing formulae, penalty and safeguard clauses), i.e. the document's so-called formulaic sections. They are subsequently adapted to the case at hand by inserting details supplied orally by the parties, including personal names and kinship, place names, the object of the transaction and its attributes (given that the transactions often concern landed property: extent, boundaries, appurtenances), and payment conditions (price or exchanges in kind), which constitute the so-called free sections of the document.<sup>6</sup> This raises the question of whether the realisation of agreement is sensitive to the contrast between reproducing formulae and composing these freer segments.

Although the overall process of charter production is reasonably well understood, several aspects remain opaque, above all the material form and modes of transmission of the formulae themselves. Were such formulae stabilised in written formulary collections, or were they

<sup>1</sup>“Digital Latin Dialectology (DiLaDi): Tracing Linguistic Variation in the Light of Ancient and Early Medieval Sources” (HORIZON-ERC-2022-ADG project no. 101098102, PI Béla Adamik; <https://pric.unive.it/projects/diladi/home>).

<sup>2</sup>Beginning in 2023, the LLDB database (<http://lldb.elte.hu/>) expanded to include data derived mainly from eighth-century charters of private-law published in the first series of *Chartae Latinae Antiquiores* (henceforth ChLA). For methodological details on linguistic annotation, see ADAMIK (2016).

<sup>3</sup>For a comprehensive account of agreement patterns in Latin, see PINKSTER (2015) 1243–1301.

<sup>4</sup>For a discussion of irregular subject-predicate agreement patterns in a documentary corpus from ninth-century Southern Italy, see D'ARGENIO (2017).

<sup>5</sup>For a general introduction to medieval documentary types and their constituent parts, see BRESSLAU (1912–1931=1998) and PRATESI (1987).

<sup>6</sup>The linguistic contrast between formulaic and free sections was discussed by SABATINI (1965) and SABATINI (1968). See also SORNICOLA (2017) for a critical re-examination of these concepts.

primarily conveyed through scribal training and professional practice?<sup>7</sup> This, in turn, raises a corollary question about the drafting process that Korciakangas explicitly expresses as a continuum between two idealised poles: (i) direct copying from physical models and (ii) reproduction from memory after training.<sup>8</sup> Focusing on Italian private charters (especially from modern-day Tuscany, 8th–10th centuries), he argues that the dense interleaving of short formulaic strings with case-specific insertions, together with frequent micro-variation in otherwise equivalent wording, makes sustained “mechanical” copying an inefficient default strategy. Instead, it points to composition through the recombination of memorised formulaic chunks, supported by the scribe’s linguistic competence.<sup>9</sup> Korciakangas further notes that the existence (or absence) of surviving formularies is not decisive for the copying hypothesis: where formularies are attested (e.g., in France), they seem to function mainly as idea banks and memory aids rather than as templates for verbatim transfer.<sup>10</sup> Error patterns likewise suggest primarily memory-based production in private charters, since clear mechanical copy errors are rare, whereas occasional copying is most plausible for less frequent document types or in contexts where the authoritative phrasing mattered.<sup>11</sup> Finally, drawing on psycholinguistic models of working and long-term memory, Korciakangas proposes that scribes may have consulted models selectively – especially to secure the correct sequence of formula elements – while relying chiefly on recall in routine drafting.<sup>12</sup>

On account of drafting that is predominantly memory-based, variation in nominal agreement may serve as a potentially informative indicator of notarial linguistic competence, insofar as it reflects morphosyntactic choices computed online during composition rather than sequences reproduced verbatim from written models.

## 2. THE LLDB DATABASE DATA FOR INVESTIGATING NOMINAL AGREEMENT

Before outlining how the LLDB database may be used to examine nominal agreement in medieval charters, it is useful first to specify what counts as a “datum” in the project and then briefly to sketch the data-registration workflow. In LLDB database terms, «data are defined as

<sup>7</sup>In the absence of substantial direct evidence, the modalities of early medieval notarial training and apprenticeship remain difficult to reconstruct with confidence. The same uncertainty surrounds the transmission and circulation of documentary formulae, whose stability across time and place often spans large groups of texts. Although the use of general formularies has been hypothesised, [COSTAMAGNA \(1975\)](#) 215–216, among others, has argued that notaries may have relied primarily on previously produced charters, using them as practical models during drafting. In addition, as [SORNICOLA \(2026\)](#) 275–279 notes, since it is unclear whether the Roman tradition of *scholae tabellionum* persisted into early medieval Italy, it is necessary to distinguish well-structured ecclesiastical milieux (e.g. cathedral contexts with established schools of writers, such as Lucca and Benevento) from less organised settings, where scribes may have had only rudimentary training and relied on formularies in a largely mechanical way, not always fully understanding them.

<sup>8</sup>[KORCIAKANGAS \(2022\)](#) 4.

<sup>9</sup>[KORCIAKANGAS \(2022\)](#) 11–14.

<sup>10</sup>[KORCIAKANGAS \(2022\)](#) 14–15.

<sup>11</sup>[KORCIAKANGAS \(2022\)](#) 15–18.

<sup>12</sup>[KORCIAKANGAS \(2022\)](#) 20–23.

linguistic phenomena that can be isolated at text level (in terms of surface manifestation, such phenomena can be phonetic [orthographic], morphological, syntactic, lexicographical, or lexico-semantic) and that deviate from what is called the “classical” norm. For the data collector, qualifying a phenomenon as datum is based on mere difference, as defined above, irrespective of whether that difference corresponds to his/her expectations in terms of diachronic processes that he/she may be aware of, whether it is due to linguistic or extra-linguistic (i.e., graphic or technical etc.) factors». <sup>13</sup>

Within this framework, the database is designed to process and analyse textual “errors”, i.e., deviations from the classical norm, <sup>14</sup> in ancient and early medieval primary documents by standardising their annotation and rendering them searchable and comparable across user-defined subcorpora. For each document in the first series of the *Chartae Latinae Antiquiores*, the text as printed in the edition is entered into the platform and paired with a facing normalised “Classical Latin” version, which serves as the baseline against which deviations are identified and coded. Data are recorded via structured data forms: in addition to core metadata (the charter’s date and place of issue, scribe’s name, etc.), each deviation is classified according to a code system organised into six macro-categories, each further subdivided into finer-grained subcodes: *vocalismus* (e.g., *é* > *I*); *consonantismus* (e.g., (voc.)-b-(voc.) > *V*); *nominalia* (e.g., dative/ablative used instead of genitive); *verbalia* (e.g., change in the form of the perfect passive); *syntactica et lexica* (e.g., *quod*, *quia*, *ut* etc. used instead of an infinitive/accusative + infinitive; very rare lexeme); *errores non grammatici* (e.g., superfluous letters). Each data form requires the entry of a code, but – where deemed appropriate – an additional alternative code may also be selected: for instance, a structure such as *componere casa* <sup>15</sup> (= *componere casam*) may legitimately be double-coded, since it can be analysed along two dimensions: (i) case confusion (macro-category *nominalia*: nominative/ablative in place of an accusative in direct-object function) and (ii) final consonant loss (macro-category *consonantismus*: -m >  $\emptyset$ ). <sup>16</sup>

It is important to clarify that, within the LLDB coding architecture, the (main) code and the alternative code are not formally ranked: the system does not encode an explicit distinction between a “preferred” and a “disfavoured” interpretation, and the alternative code merely provides a way to register a competing – yet plausible – analysis of the same phenomenon. <sup>17</sup> Accordingly, the two codes may represent either genuinely co-equal options or an unequal pairing reflecting different degrees of confidence (cf. § 3) – an asymmetry that the database itself does not represent. For this reason, the main vs alternative ordering cannot be used as evidence of the data collectors’ interpretative preference, even if one may reasonably suppose

<sup>13</sup>LLDB guidelines for data collection, available at: [https://lldb.elte.hu/admin/doc\\_guidelines.php](https://lldb.elte.hu/admin/doc_guidelines.php).

<sup>14</sup>The “classical” norm may be understood as grammatically correct Latin.

<sup>15</sup>Data form reference number: LLDB-166878.

<sup>16</sup>As is well known, final -m omission is very common in substandard registers and in Late Latin. For this phenomenon in the Latin of inscriptions, see VÄÄNÄNEN (1966) 71–77 and ADAMIK (2020).

<sup>17</sup>Cf. SZLOVICSAK (2024). Working on gender confusions in inscriptional LLDB material, the scholar highlights the value of the database’s dual-coding system for capturing competing analyses and stresses that assigning an interpretation to the main vs alternative code field does not imply a preference. He further argues that ignoring alternative codes may yield misleading temporal or spatial patterns and therefore explicitly controls for alternative-code effects using statistical modelling to test whether previous results remain robust once competing interpretations are taken into account.

that, in the event of asymmetry, collectors tend to place their favoured analysis in the main code field.

Consistent with Korciakangas's remarks on the relative rarity of material copying errors,<sup>18</sup> the LLDB database evidence points in the same direction. A total of 27,029 data forms drawn from 312 documents in the first series of the *Chartae Latinae Antiquiores* indicates that only c. 2% of recorded deviations are non-grammatical material slips – predominantly single-letter omissions, insertions, or transpositions – whereas the remainder belong to linguistic domains (see Fig. 1).<sup>19</sup> Crucially, these non-grammatical errors are not “copy” errors in the strict palaeographical sense: the loss or addition of a single letter is more plausibly attributed to momentary lapses during drafting (or internal dictation) than to line-of-sight reproduction from an exemplar, and this differs qualitatively from classic copying phenomena such as homeoteleuton.

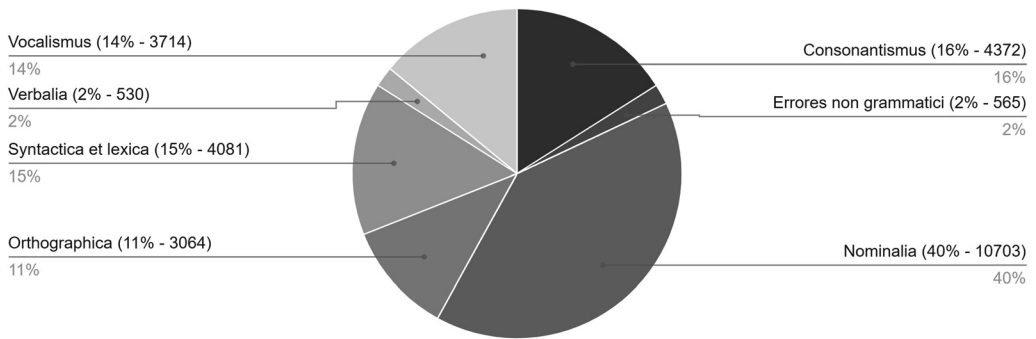


Fig. 1. Distribution of deviation types in the LLDB database, ChLA section<sup>20</sup>

In what follows, nominal agreement mismatches are defined as those instances in which the controller and one or more targets fail to share the same feature values (gender, number, and/or case). By contrast, uniformly inflected sequences are excluded, even when the morphology is non-classical or the case choice is non-standard, provided that concord holds within the noun phrase (henceforth NP). The instances below illustrate this distinction:

<sup>18</sup>KORCIKANGAS (2022) 15–18.

<sup>19</sup>As the LLDB database is continuously updated, the data presented in this study refer to the state of the dataset on 19 April 2025.

<sup>20</sup>The *orthographica* category covers deviation types that are strictly orthographic in nature, i.e. vocalic and consonantal spellings that do not entail a morphological or syntactic change, such as: «litterae Graecae, g > C, qu > CV, H > Ø, *aspiratio vitiosa*, ch > C, ph > P, th > T, PH~ F, c > K, k > C, x > SX/CS/XS/XSS/XX, i (=j/) > II, ae > E, é > AE, é: > AE, ae > E, e > AE, e: > AE, ae/ae > AI, ae/ae > AEI, ai > AEI/AE, i: > II, e: > EE, a: > AA, o: > OO, u: > VV (the colon or “:” after a vowel indicates that the vowel is long; the accent above a vowel as in e.g. “é” indicates that the vowel is stressed). Purely orthographic phenomena here include not only spelling patterns based on different (substandard) practices (which were to be avoided in standard orthography), such as CV instead of QV, CS instead of X, or geminating vowels as VV to denote long u, but also which testify to linguistic changes that have already taken place (are no longer active), such as not writing H (H > Ø), writing it in the wrong place (*aspiratio vitiosa*), or confusing AE and E» (ADAMIK (2021) 5, fn. 15).

(1) *Signum + mano Rachiprando filio*<sup>21</sup> (= *Signum + manus Rachiprandi filii*)

(2) *ipsos pauperos*<sup>22</sup> (= *ipsos pauperes*)

Although the construction in (1) is non-classical (with the proper name and its apposition expected in the genitive), the controller *Rachiprando* and its target *filio* are inflected in the same way (apparent masculine singular dative/ablative). Agreement, therefore, holds with respect to the overt features; this is not counted as an agreement mismatch in this study. In (2), the noun exhibits declensional reshaping, a metaplasm from the third to the second declension (plural accusative in *-os* instead of the expected *-es*), yet the noun and its determiner *ipsos* share the relevant feature values (masculine accusative plural). Agreement is satisfied, so this kind of structure is excluded from the dataset.

To identify instances of nominal agreement mismatches in the database, it is necessary first to determine which annotation codes may capture them. This step is not fully automatic. On the one hand, certain deviation types – such as purely orthographic phenomena and consonantal developments that do not affect nominal morphology because of their type or position – can be excluded straightforwardly. On the other hand, some codes are used specifically to flag agreement mismatches, for example, in number or gender.<sup>23</sup> Relevant categories include singular instead of plural and masculine instead of feminine, under which the database records sequences such as *Ego () parentes = Ego () parens* (LLDB-157014)<sup>24</sup> and *ipse basilica = ipsa basilica* (LLDB-176210).

Case-related mismatches are more difficult to handle, because the analysis cannot be restricted to agreement alone: the well-formedness of the controller's case must also be taken into account. Consider a sequence such as *in mea sit potestatem*<sup>25</sup> (= *in mea sit potestate*). If one looks only at internal agreement, the controller appears in the accusative (*potestatem*), while the modifier shows a form without final *-m* (*mea*), i.e., superficially a nominative/ablative-like ending. At the same time, however, the preposition *in* requires the ablative in this context. The database therefore records the deviation in the controller *potestatem*, assigning the main code accusative instead of ablative, and an alternative code  $\emptyset > -m$  (i.e. spurious addition of word-final *-m*).

Comparable logic applies to a construction such as *hanc paginam roborata esse videntur* (= *haec pagina roborata esse videtur*). In the LLDB database, two separate data forms are entered: one for *hanc* (LLDB-162198) and one for *paginam* (LLDB-162197), both coded as accusative instead of nominative; for *paginam*, an alternative code is also provided for the addition of final *-m*. No separate data form is entered for *roborata*, since its morphology is compatible with the “intended” nominative feminine singular agreement. The extracted data are therefore exported to Excel for manual review and additional annotation.

<sup>21</sup>LLDB-150558.

<sup>22</sup>LLDB-160166.

<sup>23</sup>For the LLDB database coding of gender confusions, see SZLOVICSAK (2024).

<sup>24</sup>In the LLDB database, omissions within quoted text made by the data collectors are conventionally indicated by round brackets.

<sup>25</sup>LLDB-165235.

### 3. EXTRACTING NOMINAL AGREEMENT MISMATCHES FROM THE LLDB DATABASE: APPLICATION TO A PILOT SELECTED SAMPLE

With these preliminary remarks in place, the analysis was conducted on a small sample of 48 charters produced in the area of present-day Tuscany – primarily in and around Lucca – between 785 and 795, and preserved in the Lucca Diocesan Archive.<sup>26</sup> This archive is particularly well known and historically significant, both for the sheer quantity of original and very early documents it has transmitted and for the structured context in which they were produced.<sup>27</sup> In particular, scholars have hypothesised the existence of a local *scriptorium*<sup>28</sup> and an associated school, which has in turn allowed the reconstruction of *magister-discipulus* relationships among individual scribes.<sup>29</sup>

The systematic annotation of all deviations in this corpus yielded 7,588 data forms in the database. A first step is to examine their distribution across the six major categories of the LLDB coding system. As shown in Fig. 2, the resulting profile, characterised by a comparatively stronger concentration of deviations in the nominal domain, followed by phonetic phenomena, broadly corresponds to the overall distribution observed for the full set of documents currently processed in the database (cf. Fig. 1).

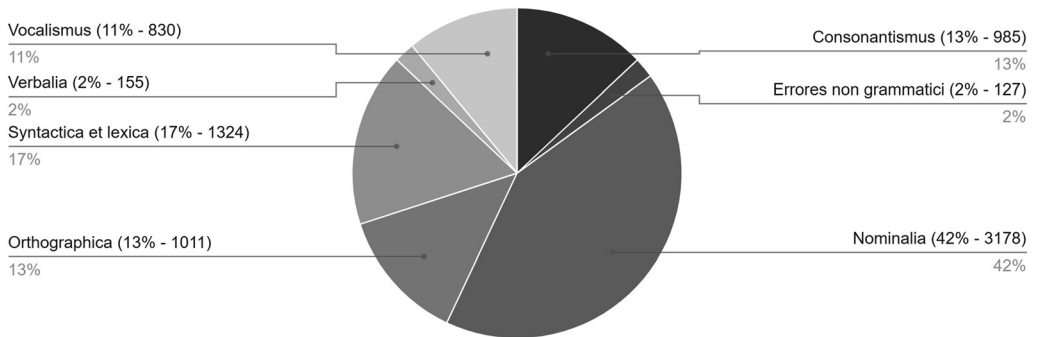


Fig. 2. Distribution of deviation types in the corpus

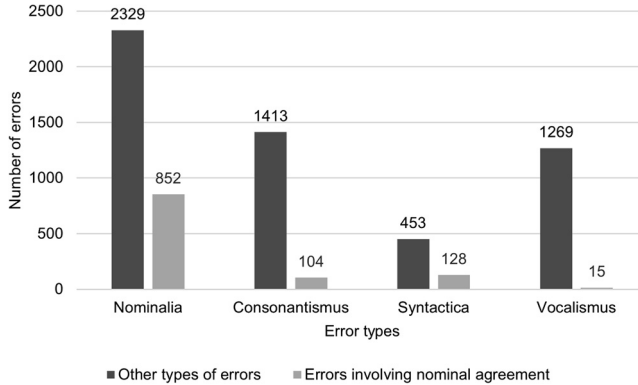
<sup>26</sup>The corpus comprises all documents published in ChLA 38 and, from ChLA 39, documents 1126–1127, 1129–1133, 1136, and 1139–1144, 1146–1149.

<sup>27</sup>MAGISTRALE (2010); TODROS (2010).

<sup>28</sup>Evidence for documentary production on site is often connected with the *Codex Lucensis 490*, compiled under Bishop Giovanni I (796–816) and involving roughly forty hands; beyond this celebrated manuscript, the existence and organisation of a local *scriptorium* remain difficult to document directly.

<sup>29</sup>Building on palaeographical criteria, scholarship has reconstructed a school and networks of scribes in Lucca; see SCHIAPARELLI (1924); PETRUCCI (1973); CATUREGLI (1959). More recently, GALDI'S (2025) linguistic analysis has added a further line of evidence, strengthening the case for a cohesive school by focusing on shared patterns of language use and the circulation of formulae.

Based on the manually refined dataset, nominal agreement mismatches account for 14% of all deviations. Figure 3 contrasts the number of deviations involving nominal agreement with other deviations recorded under the same primary codes, showing where agreement-related phenomena concentrate and how prominent they are within each category.<sup>30</sup>



**Fig. 3.** Errors involving nominal agreement versus other deviations in the corpus, grouped by main LLDB code types

The analysis necessarily relies on the main code for two reasons. First, in a large proportion of cases the primary code is the only code provided. Second, the primary code remains the most consistent anchor for aggregation and quantification, plausibly also reflecting the analysis that the data collector considered more likely when two interpretative options are not truly on a par. In this respect, the small proportion of agreement errors classified under *vocalismus* is particularly instructive: NPs such as *ipsi () Petrus = ipse () Petrus* (LLDB-169594) are only dubiously classified as agreement mismatches. A more straightforward account is a common vocalic confusion (code:  $e > i$ ), with a secondary (and less compelling) possibility of nominative-dative confusion.

### 3.1. Nominal agreement mismatches and their interface with *consonantismus* codes

With regard to the *consonantismus* category, nominal agreement-related mismatches in the corpus are restricted to the omission or overextension of final *-m* and *-s*.<sup>31</sup> This is particularly clear in instances where there is no independent reason to posit case substitution. For example, in *ad ipsa casas* (LLDB-177173) the mismatch is most straightforwardly accounted for as the omission of final *-s* on the modifier (*ipsa* for expected *ipsas*), without any compelling motivation to invoke a reanalysis in terms of nominative/ablative instead of accusative and singular instead

<sup>30</sup>The comparison is restricted to the codes that are relevant to errors involving nominal agreement. For this reason, deviations affecting verbal morphology, the lexicon, and non-grammatical (material) errors are excluded.

<sup>31</sup>For a discussion of the omission of final *-s* in Latin inscriptions, see HERMAN (1987=2006) and ADAMIK (2017).

of plural. A similar interpretation applies to *Constat me Ellaru*<sup>32</sup> (= *Constat me Hilarum*), where the relevant deviation concerns exclusively the omission of final *-m*.

At the same time, and in line with the discussion above on double coding, consonantal phenomena may coexist with case-related interpretations in the annotation. Accordingly, a *consonantismus* code may appear with a *nominalia* code entered as an alternative, particularly in cases where a residual analysis in terms of case or gender confusion is also considered plausible:

(3) *manus mea subscripsi* = *manu mea subscripsi*; code:  $\emptyset > -s$ ; alternative code: nominative/genitive instead of ablative (LLDB-155123)

(4) *omnes res illa quas* = *omnes res illas quas*; code:  $-s > \emptyset$ ; alternative code: neuter instead of feminine (LLDB-149059).<sup>33</sup>

### 3.2. Nominal agreement mismatches and their interface with *syntactica* codes

With regard to the *syntactica* codes, the data forms record instances of the singular used in place of the plural (110 occurrences; an alternative code is provided in only five cases), which are by far the dominant pattern, alongside fewer instances of the reverse configuration, i.e. plural in place of singular (18 occurrences; with an alternative code in only two cases).

As for singular-for-plural, the most frequent context is the dating formula (44 occurrences):

(5) *Regnante d(om)n(o) n(ostr)o Carulo rege Francorum et Langubardorum, anno regni eius quo Langobardiam coepit duodecimo, et filio eius d(om)n(o) n(ostr)o Pipino rege, anno regni eius quinto, quinto kal(endas) septembris, ind(ictione) octava* (ChLA 38, 1099, 1–2)<sup>34</sup>

This is a relatively long opening section, realised as an ablative absolute construction and functioning – after the invocation to God – as a fixed introductory frame for the *chartula* documentary type, in which a participle in the ablative singular is construed with two coordinated, complex NPs.<sup>35</sup>

In this case, the use of the singular cannot straightforwardly be treated as a deviation from grammatically correct Latin. Indeed, as Pinkster observes, with regard to grammatical agreement in person/number of the verb with a compound subject, «[i]n Classical Latin, the majority of verbs are plural when (i) the subject is compound and consists of two or more singular members that refer to animate beings, and (ii) the entire subject precedes the verb. In all other situations agreement with the nearest member is much more common, especially with inanimate subjects [...]. Latin, therefore, presents a broad spectrum of possibilities. There is, moreover,

<sup>32</sup>LLDB-173458.

<sup>33</sup>For *res* associated with the neuter, cf. § 3.3.

<sup>34</sup>In citations from ChLA, references follow this format: volume number, document number, line number(s). The editor's expansions, printed in italics in ChLA, are given in round brackets.

<sup>35</sup>The formula is absent from documentary types other than the *chartula*, such as the *notitia brevis*, where dating is expressed through a different formula placed in the closing section of the document: *Act(um) est hoc anno dominorum nostrorum Caruli et Pipini regum quintodecimo et octavo, septimodecimo kal(endas) augustas, ind(ictione) undecima* (ChLA 38, 1121, 30–32). For some observations on the dating formulae attested in the chartulae drafted by the scribes of Lucca, cf. GALDI (2025) 562.

considerable variation between authors and within individual authors, so it is better to speak of more or less strong tendencies than of rules». <sup>36</sup> Pinkster further notes that «if the verb precedes a compound subject, it usually agrees with the first (and nearest) member of that subject», while a following apposition may nevertheless appear in the plural. <sup>37</sup>

Similar considerations apply to nominal modifiers: «[w]hen a nominal modifier modifies two or more coordinated nouns or noun phrases it rarely agrees with the compound head as a whole [...]. It is much more common for the modifier to agree with only one of the members of the compound noun phrase, almost always with the nearest one». <sup>38</sup> In our corpus, this is illustrated by an anaphoric dating formula occurring in the final section of some documents: *regn(um) et ind(ictione) s(upra)s(crip)ta* (LLDB-155650); *die et tempore, regnum et indictio s(upra)s(crip)ta* (LLDB-150630). Here, the agreement target aligns in number and gender with the nearest conjunct, despite «the rule that masculine takes precedence over feminine». <sup>39</sup>

Within the LLDB framework and for the purposes of the present study, the plural in formulae such as (5) is treated as the expected option on two grounds. First, on semantic grounds, since the two coordinated members are animate entities and jointly refer to a shared circumstance. Second, on empirical grounds, because if one considers the entire eighth-century documentation of the Lucca Diocesan Archive (ChLA 30–40) and, for comparison, the charters from the monastery of *S. Salvatore al Monte Amiata* preserved in the Siena State Archive (ChLA 23), this type of absolute construction is at least occasionally attested with plural agreement.

The first instances in the Lucca documents in which a present participle in the ablative is construed with two coordinated NPs are reported below. In earlier charters, by contrast, the controller is consistently a single ruler (Liutprand, Aistulf), so the opening with *regnante* is entirely coherent.

(6) *Regnante d(om)no n(ostr)o Desiderio rege, anno regni eius D(e)o propitio tertio, idem et d(om)n(o) n(ostr)o [Adelch]is rex filio eius primo, mense octubrio, ind(ictione) tertiadecima* (ChLA 32, 954, 1–2)

(7) *Regnante d(om)n(o) n(ostr)o Desiderio et Adelghis regibus, anno regni eorum tertio et primo, m(ense) octubrio ind(ictione) .XIII.* (ChLA 32, 955, 1–2)

In (6), the presence of *idem* (= *item*) <sup>40</sup> can be taken to semantically “absorb” the repetition of *regnante*, making the singular participle somewhat less unexpected. In (7), by contrast, *regnantibus* would appear more strongly expected, given the plural *regibus* and *eorum*, which encourage an interpretation of the compound head as a whole. These constructions, even with minor variation, <sup>41</sup> alternate across charters, and the pattern illustrated in (6) also occurs without *item*. Conversely, the pattern in (7) may, albeit rarely, include further plural elements (e.g., *Regnante d(om)n(i) n(ostr)i*, ChLA 33, 980, 1; ChLA 34, 996, 1).

<sup>36</sup>PINKSTER (2015) 1247.

<sup>37</sup>PINKSTER (2015) 1251.

<sup>38</sup>PINKSTER (2015) 1273.

<sup>39</sup>PINKSTER (2015) 1273.

<sup>40</sup>The interpretation, however, remains uncertain: the form may also be the neuter of the pronoun. Cf. GALDI (2025) 562.

<sup>41</sup>A further type is: *Regnante d(om)n(o) n(ostr)o Desiderio rege et filio eius idem d(om)n(o) n(ostr)o Adelchis rege, anno regni eor(um) duodecimo et nono* (ChLA 34, 1001, 1).

Comparable formulae occur in the Siena charters as well, where, however, constructions with a plural participle are also attested:

(8) *regnantes d(om)n(is) nostris Desiderio et Adelgis praecell(entissimis) regibus, anno regni eor(um) septimo et quarto, quintadec(ima) diae mensi magii, ind(ictione) prima* (ChLA 23, 739, 1–2)

(9) *regnantibus d(om)n(is) n(ostris) Desiderio et Adelchis fil(io) eius precell(entissimis) regibus anno regni eor(um) Deo ausiliante nono et septimo, mense octubrio, ind(ictione) quarta* (ChLA 23, 743, 1–3)

In 774, Charlemagne appears for the first time in the dating clause of the Lucca charters, as the sole ruler:

(10) *Regnante d(om)n(o) n(ostr)o Carulo rege Francor(um) et Langobardor(um), anno regni eius quo coepit Langobardia(m) primo, septimo decimo kalendas augustas, p(er) ind(ictionem) duodecimam* (ChLA 36, 1047, 1–2)

The earliest surviving dating clauses in which Pippin is also mentioned are cited below; in (11) the regnal-year computation is given separately for each ruler, whereas in (12) it is presented as a unified count:

(11) *Regnante d(om)n(o) n(ostr)o Carulo rege Francorum et Langobardorum, anno regni eius quo La[ngobardia(m)] coepit octavo et filio eius d(om)n(o) n(ostr)o Pippino rege, anno regni eius primo, quinto decimo ka[lendas...] ind(ictione) quarta* (ChLA 37, 1077, 1–3)

(12) *Regnante d(om)n(o) n(ostr)o Carulo rege Francorum et Langobardorum et d(om)n(o) n(ostr)o Pipino idem rex filio eius, anno regni eorum nono et secundo, m(ense) augusto p(er) ind(ictione) quinta* (ChLA 37, 1082, 1–3)

With only minor variation, irrelevant to the issue at hand, these formulae recur throughout the subsequent documentation. The following instances constitute exceptions:

(13) *Regnantes dom(ni) n(os)tri Carulo et Pipino fil(io) eius veris excelle(ntissimi) reges Francor(um) seo et Langobardor(um), anno regni eoru(m) in Aetalia octavo et primo, sub d(ie) octavo kal(endas) iunias, ind(ictione) quinta* (ChLA 37, 1081, 1–4)

(14) *Regnantes domni nostri Carulo et Pipino filio eius viris excelle(ntissimi) reges Francorum seo et Langobardorum, anno regni eorum in Ethalia decimo et tertio, mense genua[rio], indict(i)o septima* (ChLA 37, 1090, 1–3)

(15) *Regnantes domni nostri Carulo et Pipino filio eius viris excellentissimi reges Francorum seo et Langobardorum, anno regni eorum in Ethalia decimo et tertio, menses genuario, indictio septima* (ChLA 37, 1092, 1–3)

(16) *Regnantes domn(is) no(st)r(is) Carulo et Pipino fil(io) eius viris excell(entissimis) regis Francor(u) m seo et Langobardor(u)m, anno regni eor(um) in D(e)i n(omin)e quod Langobardia cepit in Ethalia tertiodecimo et sexto, pridie k(a)l(endas) septembris, ind(ictione) nona* (ChLA 38, 1104, 1–3)

It is noteworthy that attestations (14)–(16) were produced by the same scribe, *Aupert presbiter*. This supports an interpretation in terms of formulaic fixation, either anchored in a specific notarial milieu or reflecting individual scribal habits. In (13), the prominent role of the commissioner (the *gastaldus* of Volterra, designated *vir magnificus* in the *signum manus*) may be relevant as well and may plausibly correlate with a higher level of the scribe's written proficiency and/or more advanced training.

More specifically, four features of different kinds characterise this dating formula: (i) the nominative/accusative form of the participle; (ii) the immediate adjacency of the two heads in the coordinated NPs; (iii) the (comparatively infrequent) mention of Italy; and (iv) the fact that the relevant documents were drafted in localities other than Lucca (Volterra, Paterno, Populonia). Taken together, these traits suggest that animacy and the linear ordering of targets and controllers are unlikely to be the main factors influencing agreement realisation. A more plausible conditioning factor, at least in relative terms, may be the distance between the coordinated nominal heads: reduced distance, and especially strict adjacency, appears to favour plural-oriented readings of the coordination. This hypothesis is consistent with the evidence from the Siena charters: in the plural-participle formulae (e.g., (8)–(9)), the heads Desiderius and Adelchis are immediately adjacent, a configuration that plausibly favours plural agreement on the participle and may also encourage plural rendering in preceding appositions such as *domini(s) nostri(s)*.

Against this background, the singular is best explained as the effect of a crystallised dating template: a formula originally framed for a single ruler tends to be retained even when the political configuration involves two rulers, circulating with considerable stability within the notarial environment of Lucca. Comparable “inertia” can be observed in other highly frequent choices in the same formula (e.g. *cepit* being almost systematically realised as *coepit*).<sup>42</sup> This may point to the fixation of local patterns: where plural participial forms do occur, they appear more readily associated with notarial contexts outside Lucca, suggesting a distribution shaped by local scribal networks and practices rather than by grammatical constraints alone.

### 3.3. Nominal agreement mismatches and their interface with *nominalia* codes

As shown in Fig. 3, the codes most frequently associated with nominal agreement mismatches fall within the *nominalia* category. The discussion below is limited to codes with more than 40 attestations.

The most frequent code is accusative for ablative. An interesting context concerns the witness-subscription formula, e.g. *Ego () in hanc cartula () subscripsi* (for expected *in hac chartula*).<sup>43</sup> This context is noteworthy because it lies, at least in principle, outside the scribe’s direct responsibility, since the subscription is in the witness’s own hand; at the same time, it is possible that the witness wrote under the scribe’s guidance or dictation, and that earlier subscriptions in the same charter may function as immediate models – here in the literal sense of direct copying – for those added later.<sup>44</sup>

Other cases occur in highly routinised formulaic contexts where the morphosyntactic reading plausibly coexists with a competing phonographic analysis; accordingly, the database records  $\emptyset > -m$  as an alternative code. This applies, for instance, to expressions for circumstances typically mentioned in notarial documents, referring to: (i) the stated motivation for

<sup>42</sup>Significantly, across the entire LLDB database (ChLA section), the code *e/é: > OE* records only instances of *cepit > coepit*, all of which come from the Lucca charters.

<sup>43</sup>Here I do not consider the alternative, grammatically correct Latin rendering with the dative: *huic chartulae subscripsi*. Cf. *OLD*, s.v. ‘*subscribo*’.

<sup>44</sup>Despite their brevity and stereotyped wording, autograph subscriptions may nonetheless warrant separate consideration, since their conditions of production differ from those of the main text and may affect nominal agreement.

the transaction (*pro anima meam*, LLDB-166315), (ii) legal or mentioned procedures (*canonico ordinem*, LLDB-176198; *in predicto ordinem*, LLDB-151417), (iii) the extent of control retained over the property (*in mea sit potestatem*, LLDB-150899), and (iv) modes of reparation in the event of infringement (*in ferquidem loco*, LLDB-150510). For the last example, Arcamone interprets *ferquidem*, a Germanism, as an accusative singular, in comparison with *ferquido* (ablative) and *ferquidis* (ablative plural), deriving these forms from an unattested nominative *\*ferquidus* or *\*ferquidis*.<sup>45</sup> On this basis, the frequently attested form *ferquide* (*in ferquide loco*) in our corpus may be analysed as an ablative, which in turn makes it possible to treat the alternation *in ferquidem loco* vs *in ferquide loco* as cases of, respectively, lack of agreement and agreement. Semantically, from an etymological sense roughly ‘aforementioned/already mentioned’, *ferquid-* develops the meaning ‘corresponding, equivalent’ in the *Leges Langobardorum*, where it is glossed as *id est simile*, and in notarial Latin, where it occurs in the formula *componere in ferquidem loco, sub stimulatione* with the sense ‘for an equivalent compensation’.<sup>46</sup> As Stoffella observes, «in the central decades of the eighth century *ferquidem* was sporadically used as an alternative to the Latin word *eodem*, in the second part of the eighth century this Lombard word became recurrent in the area of Lucca within a formula used in final clauses of sales or exchanges written by both lay and ecclesiastic scribes».<sup>47</sup> In the quantitative data reported by Stoffella for early medieval Tuscan private charters (720–790),<sup>48</sup> the distribution is balanced: agreement forms, traceable to *in ferquide/ferquido loco* and *in ferquidem locum*, amount to 25 occurrences, matching the 25 cases of non-agreement attributable to *in ferquidem loco* and *in ferquido/ferquide locum*. Even where the formal realisation varies, these data point to the role of Lucca’s cathedral-school milieu, whose pupils « must have contributed significantly to the definition and diffusion of the formula».<sup>49</sup>

The second most frequent annotation in this set is dative/ablative for nominative. In most cases, it co-occurs with the alternative code *-us > O* (i.e. *-s > Ø + -u > O*). Two recurrent configurations can be distinguished, and both may also occur in autograph witness subscriptions:

- (i) Nominative personal pronoun and/or proper name + apposition with a second-declension noun ending in *-o*: *Petrus germano* (LLDB-169595); *Ego Verbonus pr(es)b(ite)r* (LLDB-150893).
- (ii) Nominative personal pronoun + second-declension proper name ending in *-o*: *Ego Domnuccio* (LLDB-155063); *ego Vito* (LLDB-172372)

From a morphosyntactic perspective, the *-o* ending is compatible with an Italo-Romance-type morphology and may be interpreted as a reflex of case neutralisation (i.e. a single case form extending across syntactic functions), especially with NPs used to name and qualify the

<sup>45</sup>ARCAMONE (1969) 267.

<sup>46</sup>FRANCOVICH ONESTI (1999) 79–80.

<sup>47</sup>STOFFELLA (2021) 319.

<sup>48</sup>STOFFELLA (2021) 319–322.

<sup>49</sup>STOFFELLA (2021) 329.

participants in the transaction, a context that is particularly susceptible to the pressure of everyday language usage.<sup>50</sup>

The third most frequent code is nominative/ablative for accusative in the function of direct object (62 occurrences) and is often associated with the omission of final *-m*. From a syntactic point of view, the phenomenon may affect either the head of the NP (e.g., *Ipsam () petia () recepit*, LLDB-158270) or the modifier (e.g., *reddere () ipsa cautionem*, LLDB-166026). Quantitatively, the code is assigned to the modifier in 46 instances, whereas in the remaining 16 cases it applies to the head. As for linear order, the modifier precedes the head in 42 cases and follows it in 20. Although the sample is too small to support generalisations, the imbalance is at least compatible with a weak tendency for targets to be implicated more frequently than controllers. A plausible contributing factor is the loss of final *-m*, which can erode overt accusative marking, especially in short, high-frequency forms. Since the targets in question are often possessives and demonstratives, their repeated, short forms may be especially susceptible to this phonological weakening, increasing the likelihood that accusative morphology becomes less transparent in the target position.

Two further codes are nominative for ablative and accusative for nominative. As regards the former, the great majority of instances again come from the dating clause, where, within the ablative absolute frame, appositional titles attached to Charlemagne, Pippin, or both surface in the nominative, both singular (e.g. *rex, patricius*) and plural (e.g. *virī, domni*); these also appear with the alternative code ablative absolute with a mixed nominative.

Within the accusative-for-nominative code, which in several cases co-occurs with the alternative code  $\emptyset > -m$ , the most recurrent pattern involves elements occurring in the so-called *firmitas* clause, whose function is to guarantee that the document, as drafted and in its contents, remains stable and legally valid over time: e.g., *hec cartulam omni tempore in pr(e)d(ic)to ordinem firma(m) et stabile(m) permaneat* (ChLA 39, 1141, 21–22); *hec () paginam in s(upra) s(crip)to ordinem omni tempore in sua p(er)maneat firmitatem* (ChLA 39, 1147, 30–31).

Finally, it is worth mentioning the NP *omnia res*, often accompanied by an additional modifier such as *mea*, *illa*, or *ipsa*. In its most frequent documentary setting, this NP is governed by prepositions such as *cum*, *de*, and *in*, or by verbs such as *trado*, *do*, *compono*, within clauses specifying that the object of the transaction belongs to a party or is transferred/compensated together with all its appurtenances. In these contexts, both a collective reading ('the whole, the totality') and a distributive reading ('each individual item within a set', the latter is the meaning continued in Italian *ogni*) may be implicated. As a consequence, the number interpretation of these constructions can also be ambiguous, especially since, in the same formulaic sequences across other documents, *omnibus rebus* (in prepositional contexts) and *omnes res* (in verbal contexts) are more frequent, although attestations of the singular *omni re* and *omnem rem* are also found. Moreover, as Stotz notes in connection with the rigid use of a case form, «[d]ie Nom./Acc. neuter. pl.-Form *omnia* ist seit der Spätanike vielfach für einen (kollektiven) Sg.-Begriff (etwa: ‚das Ganze‘) genommen worden, wie denn die Neutra, was die Kongruenz betrifft, vielfach mit Sg.-Formen verbunden waren [...]. Zu *hec omnia rem* ist die häufige Verknüpfung von *res* und einem Begriff im (Neutr.) Pl. zu stellen. In langob. Urkunden etwa begegnen

<sup>50</sup>On the emergence of a "single case" form in notarial documents, see SABATINI (1965); for Langobard charters, see VALENTINI (2018) 128–148, which also discusses greater syntactic flexibility in appositions, in the numerous noun phrases referring to the parties in legal transactions.

Formulierungen wie: *cum homnia ris (res) mea, cum omnia legnamen [...]*. Bei *res, lignamen* usf. kann an kollektive Sg.-Begriffe gedacht werden [...]. Eingewirkt haben kann stellenweise die bei vielen Adjektiven auf *-is/-e* auftretende Tendenz, die Formen auf *-ia* als Nom./Abl. sg. fem.-Formen „umzunutzen“». <sup>51</sup>

Since within a given document the form *omnia* does not alternate with either *omnibus/omni* or *omnes/omnem*, the phrase can be taken – at least for some scribes – as a rigid, crystallised pattern whose potential structural ambiguity is pragmatically neutralised by its recurrence and functional transparency (‘property transferred with its appurtenances’), and thus does not affect the clause’s primary pragmatic-legal force.

#### 4. CONCLUSIONS

This pilot study suggests that the ChLA section of the LLDB database can be productively exploited for the study of nominal agreement, provided that the identification of relevant cases through the LLDB code system is supplemented with targeted qualitative control, especially where agreement interfaces with other deviation domains. Three points may be noted.

First, nominal agreement mismatches tend to cluster in high-frequency formulaic sequences, rather than being evenly distributed across the textual fabric of the charters. This concentration points to the role of routinisation: in highly recurrent sequences, mismatches may persist because the communicative and legal function of the clause is stable and predictable. As a result, nominal agreement outcomes in these environments are less likely to mirror on-line morpho-syntactic computation, and more likely to reflect the pull of crystallised patterns.

Second, many nominal agreement deviations show a strong interaction with very frequent phonographic phenomena. <sup>52</sup> This interaction is particularly visible in short, high-frequency targets – especially possessives and demonstratives – whose reduced phonological form makes case/number marking vulnerable to erosion, quite apart from agreement proper. Moreover, nominal agreement mismatches sometimes involve lexemes that in medieval sources display the rigidified use of a single case form. Since such rigidification is widely attested and can also occur in notarial documents where the Latin case system is relatively well preserved, <sup>53</sup> it cannot be treated as a diagnostic of scribal linguistic competence, even when it may co-occur with – and at times contribute to – agreement mismatches.

Third, some non-agreement patterns show signs of diffusion and stabilisation within specific notarial milieus. This suggests that nominal agreement outcomes may reflect documentary habitus and local practice.

Overall, nominal agreement variation may still be informative about scribes’ linguistic competence in a predominantly memory-based drafting process, provided that frequency effects and formulaic routinisation are explicitly taken into account. Future work should address two potentially promising factors: (i) an analysis of scribes who drafted multiple charters, in order to

<sup>51</sup> STOTZ (1998) 111–112.

<sup>52</sup> Significantly, subject-predicate agreement mismatches as well (e.g., *Sign(um) + m(anus) Gudip(er)ti, qui () rogavi*, LLDB-171452; *ego () voluit*, LLDB-162139), involving first-person forms used for third-person and, more often, the reverse, are associated with omission or overextension of a word-final consonant, in this case *-t*; cf. D’ARGENIO (2017).

<sup>53</sup> Cf. SORNICOLA (2012) 57–59.

assess the weight of individual linguistic habits and the extent to which fixed chunks are reused without being fully integrated into the local syntax; (ii) a closer examination of those mismatches that do not interact with highly frequent phonographic deviations and occur in less formulaic segments, since these cases are more likely to reflect on-line morphosyntactic computation during drafting. Along this line, it would be worth assessing whether nominal agreement patterns differ across documentary types – in particular, between *chartulae* vs *notitiae*. Since *notitiae*, that record various stages of a trial, were produced less frequently, they are likely to be less routinised in scribal practice; their specific textual organisation (including testimonies in direct speech) has a different balance between formulaic language and freer segments than *chartulae*, with potential consequences for where and how agreement mismatches surface.<sup>54</sup>

Finally, expanding the sample to other areas may help to assess how strongly local notarial practice shapes not only formulae in a broad sense, but also the micro-structures that compose them.<sup>55</sup>

## ACKNOWLEDGMENTS

This paper was prepared within the framework of the HORIZON-ERC-2022-ADG project no. 101098102, entitled “Digital Latin Dialectology (DiLaDi): Tracing Linguistic Variation in the Light of Ancient and Early Medieval Sources”, led by Béla Adamik (<https://pric.unive.it/projects/diladi/home>). Funded by the European Union. Views and opinions expressed are however those of the author only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

## REFERENCES

- ADAMIK, B. (2016). Computerized Historical Linguistic Database of the Latin Inscriptions of the Imperial Age: Search and Charting Modules. In: GRADVOHL, E. – SZABÓ, Á. (eds.), *From Polites to Magos: Studia Gyorgy Nemeth Sexagenario Dedicata*. University of Debrecen, Debrecen–Budapest, pp. 13–27.
- ADAMIK, B. (2017). The Problem of the Omission of Word-final -s as Evidenced in Latin Inscriptions. *Graeco-Latina Brunensia*, 22(2): 5–21, <https://doi.org/10.5817/GLB2017-2-1>
- ADAMIK, B. (2020). On the Loss of Final -m: Phonological or Morphosyntactic Change? *Acta Antiqua Academiae Scientiarum Hungaricae*, 59(1–4): 97–108, <https://doi.org/10.1556/068.2019.59.1-4.11>
- ADAMIK, B. (2021). Romanization and Latinization of the Roman Empire in the light of data in the Computerized Historical Linguistic Database of Latin Inscriptions of the Imperial Age. *Journal of Latin Linguistics*, 20(1): 1–19, <https://doi.org/10.1515/joll-2021-2016>

<sup>54</sup>Cf. D’ARGENIO (2026) for a linguistic perspective on *notitiae*, with particular attention to direct speech and to scribal practice, also assessed against *chartulae* written by the same scribe.

<sup>55</sup>On the comparability of notarial documents across early medieval Italian writing centres, and on how local writing traditions, socio-historical settings, and notarial-group features can shape formulaic practices and linguistic patterns, see D’ARGENIO–VECCHIA (2026).

- ARCAMONE, M.G. (1969). Glossario delle parole di origine germanica. In: TIRELLI CARLI, M. (ed.), *Carte dell'Archivio Capitolare di Pisa*, 4(1101–1120). Edizioni di Storia e Letteratura, Roma, pp. 264–272.
- BRESSLAU, H. (1912–1931=1998). *Manuale di diplomatica per la Germania e l'Italia*. Pubblicazioni degli Archivi di Stato, Roma.
- CATUREGLI, N. (1959). La corsiva lucchese del secolo VIII. *Atti dell'Accademia lucchese di scienze, lettere e arti*, 5(10): 125–150.
- ChLA 23, PETRUCCI, A. – TJÄDER, J.-O. (eds.) (1985). *Chartae Latinae Antiquiores*. Part XXIII, Italy IV. Urs Graf Verlag, Dietikon–Zurich.
- ChLA 32, NICOLAJ, G. (ed.) (1989). *Chartae Latinae Antiquiores*. Part XXXII, Italy XIII. Urs Graf Verlag, Dietikon–Zurich.
- ChLA 33, MAGISTRALE, F. (ed.) (1989). *Chartae Latinae Antiquiores*. Part XXXIII, Italy XIV. Urs Graf Verlag, Dietikon–Zurich.
- ChLA 34, SUPINO MARTINI, P. (ed.) (1989). *Chartae Latinae Antiquiores*. Part XXXIV, Italy XV. Urs Graf Verlag, Dietikon–Zurich.
- ChLA 36, NICOLAJ, G. (ed.) (1990). *Chartae Latinae Antiquiores*. Part XXXVI, Italy XVII. Urs Graf Verlag, Dietikon–Zurich.
- ChLA 37, SUPINO MARTINI, P. (ed.) (1990). *Chartae Latinae Antiquiores*. Part XXXVII, Italy XVIII. Urs Graf Verlag, Dietikon–Zurich.
- ChLA 38, PALMA, M. (ed., in collaboration with BIANCHI, F.) (1990). *Chartae Latinae Antiquiores*. Part XXXVIII, Italy XIX. Urs Graf Verlag, Dietikon–Zurich.
- ChLA 39, MAGISTRALE, F. (ed.) (1991). *Chartae Latinae Antiquiores*. Part XXXIX, Italy XX. Urs Graf Verlag, Dietikon–Zurich.
- COSTAMAGNA, G. (1975). L'alto medioevo. In: AMELOTTI, M. – COSTAMAGNA, G. (eds.), *Alle origini del notariato italiano*. Consiglio Nazionale del notariato, Roma, pp. 147–314.
- D'ARGENIO, E. (2017). L'accordo tra soggetto e predicato: due fenomeni di irregolarità. In: SORNICOLA, R. – D'ARGENIO, E. – GRECO, P. (eds.), *Sistemi, norme, scritture: La lingua delle più antiche carte cavensi*. Giannini, Napoli, pp. 203–211.
- D'ARGENIO, E. (2026). Spoken Language in Medieval Legal Charters? Revisiting some 8th-century *notitia*. *Acta Linguistica Academica*, 73(1): 19–40, <https://doi.org/10.1556/2062.2025.01017>
- D'ARGENIO, E. – VECCHIA, C. (2026). Comparing Language, Style, and Formulary in Notarial Documents in Early Medieval Italy. In: MOLINELLI, P. – GHEZZI, C. (eds.), *Multilingualism and Writing Practices in Defining Communities. Italy in Late Antiquity and the Middle Ages*. De Gruyter, Berlin, pp. 329–351.
- FRANCOVICH ONESTI, N. (1999). *Vestigia longobarde in Italia (568-774)*. Lessico e antroponomia. Artemide, Roma.
- GALDI, G. (2025). Osservazioni sulla lingua e sul formulario dei rogati lucchesi (VIII-IX secolo). In: D'ARGENIO, E. – DELLE DONNE, R. – SORNICOLA, R. (eds.), *Lingue, scritture e società nell'Italia longobarda: Un percorso di sociolinguistica storica*. FedOA Press, Napoli, pp. 551–570.
- OLD = GLARE, P.G.W. (ed.) (2012). *Oxford Latin Dictionary*, 2nd ed. Oxford University Press, Oxford.
- HERMAN, J. (1987=2006). La disparition de -s et la morphologie dialectale du latin parlé. In: HERMAN, J. (ed.), *Du latin aux langues Romanes, II: Nouvelles études de linguistique historique* (réun. S. Kiss). Niemeyer, Tübingen, pp. 33–42.
- KORKIAKANGAS, T. (2022). From Memory or Formulary: How Were Medieval Documentary Formulae Reproduced? *Mirator*, 22(1): 4–24, <https://doi.org/10.54334/mirator.v22i1.119760>
- MAGISTRALE, F. (2010). Le pergamene dell'Archivio Arcivescovile di Lucca (secoli VIII-IX): l'esperienza delle *Chartae Latinae Antiquiores*. In: PAGANO, S. – PIATTI, P. (eds.), *Il patrimonio documentario della chiesa*

- di Lucca. Prospettive di ricerca.* Atti del Convegno internazionale di studi (Lucca, Archivio arcivescovile, 14–15 novembre 2008). Sismel, Firenze, pp. 73–92.
- PETRUCCI, A. (1973). Scrittura e libro nella Tuscia altomedievale (secoli VIII-IX). In: *Atti del 5° congresso internazionale di studi sull'alto medioevo (Lucca, 3-7 ottobre 1971)*. CISAM, Spoleto, pp. 628–643.
- PINKSTER, A. (2015). *The Oxford Latin Syntax*, I. Oxford University Press, Oxford.
- PRATESI, A. (1987). *Genesi e forme del documento medievale*, 2nd ed. Jouvence, Roma.
- SABATINI, F. (1965). Esigenze di realismo e dislocazione morfologica in testi preromanzi. *Rivista di cultura classica e medioevale*, 7: 972–998.
- SABATINI, F. (1968). Dalla “scripta latina rustica” alle “scriptae” romanze. *Studi Medievali*, s. III, 9(1): 320–358.
- SCHIAPARELLI, L. (1924). *Il codice 490 della Biblioteca Capitolare di Lucca e la scuola scrittoria lucchese (sec. VIII-IX). Contributi allo studio della minuscola precarolina in Italia*. Pompeo Sansaini, Roma.
- SZLOVICSÁK, B. (2024). Gender Confusions and Other Linguistic Changes: A Provisional Description of Vulgar Latin Phenomena. *Acta Antiqua Academiae Scientiarum Hungaricae*, 63(4): 575–594, <https://doi.org/10.1556/068.2024.00123>
- SORNICOLA, R. (2012). *Bilinguismo e diglossia dei territori bizantini e longobardi del Mezzogiorno: le testimonianze dei documenti del IX e X secolo*. Giannini, Napoli.
- SORNICOLA, R. (2017). «Transizione» e «transizioni» dal latino al romanzo: il progetto di analisi linguistica dei documenti cavensi del IX secolo. In: SORNICOLA, R. – D'ARGENIO, E. – GRECO, P. (eds.), *Sistemi, norme, scritture: La lingua delle più antiche carte cavensi*. Giannini, Napoli, pp. 13–25.
- SORNICOLA, R. (2026). When historical linguistics matters to historical sociolinguistics. Italian legal documents from the Early Middle Ages and their scribes. In: MOLINELLI, P. – GHEZZI, C. (eds.), *Multilingualism and Writing Practices in Defining Communities. Italy in Late Antiquity and the Middle Ages*. De Gruyter, Berlin, pp. 271–307, <https://doi.org/10.1515/9783112233214-013>
- STOFFELLA, M. (2021). Traces of Bilingualism in Early Medieval Northern Italy: The Evidence from Eighth- and Ninth-Century Private Charters. In: GALLAGHER, R. – ROBERTS, E. – TINTI, F. (eds.), *The Languages of Early Medieval Charters: Latin, Germanic Vernaculars, and the Written Word*. Brill, Leiden–Boston, pp. 296–341.
- STOTZ, P. (1998). *Handbuch zur lateinischen Sprache des Mittelalters, IV: Formenlehre Syntax und Stilistik*. Beck, München.
- TODROS, G. (2010). Brevi considerazioni sul patrimonio archivistico dell'Arcidiocesi lucchese. In: PAGANO, S. – PIATTI, P. (eds.), *Il patrimonio documentario della chiesa di Lucca. Prospettive di ricerca.* Atti del Convegno internazionale di studi (Lucca, Archivio arcivescovile, 14-15 novembre 2008). SISMEL, Firenze, pp. xi–xiv.
- VÄÄNÄNEN, V. (1966). *Le latin vulgaire des inscriptions pompéiennes*, 3rd ed. Akademie Verlag, Berlin.
- VALENTINI, C. (2018). *L'evoluzione della codifica del genitivo dal tipo sintetico al tipo analitico nelle carte del Codice diplomatico longobardo*. Firenze University Press, Firenze, <https://doi.org/10.36253/978-88-6453-827-3>