



## Translation and grammaticalization in 17th-Century Neapolitan: A comparative Romance analysis

Riley VanMeter \*

**Abstract.** This study explores grammaticalization in 17th-century Neapolitan through a comparative analysis with Italian, focusing on perfect auxiliary selection and possessive adjective position. The research aims to position these features of Neapolitan along the Romance Grammaticalization Cline and assess the utility of translation corpora for comparative grammaticalization analysis. The corpus used for analysis includes the Neapolitan literary text *Lo Cunto de li Cunti* and its Italian translation. Through a quantitative and qualitative approach, the study examines how Neapolitan's preference for certain auxiliary forms and syntactic structures suggests a more advanced stage of grammaticalization compared to Italian. Specifically, Neapolitan shows a tendency toward reduced variability in auxiliary selection, favoring *avere* over *essere*, and a marked preference for postnominal possessive adjective position. The findings support the hypothesis that 17th-century Neapolitan exhibited more grammaticalized structures than Italian in certain respects. Additionally, the study demonstrates the viability of using translation as a methodological tool in historical linguistics to analyze language change in synchrony. The results offer new insights into the grammaticalization process in Romance languages and raise questions about the influence of translation practices on linguistic analysis.

**Keywords.** grammaticalization; Neapolitan; Italian; comparative linguistics; perfect auxiliary; possession; translation

**1. Introduction.** The goal of the present research is twofold: first, to position 17th-century Neapolitan along the Romance Grammaticalization Cline with reference to Italian, and second, to propose the viability of translation corpora for historical linguistics and comparative grammaticalization research. The first objective will be accomplished through the quantitative analysis of both perfect auxiliary selection and possessive adjective position. The second through methodological testing and qualitative analysis.

The study is structured as such: section 2 will give an overview of grammaticalization as a phenomenon, how it has been studied, and present its relevance to Neapolitan and Italian. Subsequently, section 3 will touch on the present corpus and methodology. Section 4 will present the results. Finally, sections 5 and 6 will discuss the implications behind the results and the novelties of this methodology.

**2. Previous literature.** Grammaticalization is a foundational theory and subject of research within the study of language variation and change. It is also a theory that is closely related to comparative/contrastive linguistics. As defined by Traugott (2003: 645), grammaticalization is

---

\* I wish to thank Mairi McLaughlin and The Sociolinguistics Lab at Berkeley (SLaB) for their support and insight throughout this project. I am most grateful to Claudia Romanelli and the Italian section at The University of Alabama for planting the seeds of this project in me and offering unwavering support for my study of Italian linguistics. Authors: Riley VanMeter, The University of California, Berkeley (rileyvanmeter@berkeley.edu).

“the process whereby lexical material in highly constrained pragmatic and morphosyntactic contexts is assigned grammatical function, and once grammatical, is assigned increasingly grammatical, operator-like function”. Within Romance, the poster child for grammaticalization is the morphosyntactic future.

The morphosyntactic future of Italian and the dialects of the Italian peninsula originally derives from the vulgar Latin combination of the infinitive plus a conjugation of the Latin verb HABERE ‘to have’ (Varvaro 2013: 32–35). Thus, through grammaticalization, *cantare ho*, which expressed the obligation to sing or ‘I am to sing’, has become *canterò*, which carries the semantics of futurity. The verb HABERE therefore lost its lexical semantics of possession in favor of the grammatical meaning of futurity and ultimately became a bound suffix within this specific context.

2.1. DEFINITIONS AND PARAMETERS. Narrog and Heine (2021) recognize four parameters for identifying and reconstructing grammaticalization. The first is *context extension*. This means that a lexical structure acquires new uses in new contexts, giving rise to new meanings and with a consequently greater frequency (Narrog & Heine 2021: 57). Their second metric is *semantic bleaching*, the loss of lexical semantic properties that a linguistic expression experiences during the grammaticalization process (2021: 67). The third is *deategorization*, the loss of morphological and syntactic aspects both internally (e.g. loss of ability to be inflected) and externally (e.g. increased dependency on another form) (2021: 72). Lastly, there is *erosion*, which is characteristic of late stage grammaticalization, when a grammaticalizing word/construction loses phonetic substance (2021: 78). These indicators of grammaticalization were designated by Narrog & Heine after considering the various debates by Lehmann (2015 [1982]), Heine and Kuteva (2007), Hopper (1991), among others. However, this research will focus on the paradigmatic and syntagmatic parameters outlined by Lehmann (2015 [1982]): *paradigmaticity*, *syntagmatic variability*, and *paradigmatic variability*—all of which contribute to *obligatorification*.

*Paradigmaticity* is defined by Lehmann as “the formal and semantic integration both of a paradigm as a whole and of a single subcategory into the paradigm of its generic category” (2015 [1982]: 141). The most salient aspect of *paradigmaticity* is the size of the paradigm (Lehmann 2015 [1982]: 141). It then follows that this in turn affects *paradigmatic variability*, which is “the freedom with which the language user chooses a sign” (Lehmann 2015 [1982]: 146). It is observed that, in the grammaticalization process, *paradigmaticity* is generally reduced, resulting in a correspondent reduction in *paradigmatic variability* (Lehmann 2015 [1982]: 147). *Syntagmatic variability* is then defined as “the ease with which a sign can be shifted around in its context” (Lehmann 2015 [1982]: 167–170). This positional freedom decreases as grammatical status increases. For example, once a grammaticalizing form becomes an auxiliary, it loses the positional variation/freedom it previously had. It is this depletion of choice on the part of the language-user through reduction in *paradigmatic* and *syntagmatic variability* which results in the *obligatorification* of a structure. This is the process of becoming obligatory, a hallmark of the grammaticalization process (Narrog & Heine 2021). It is these structural synchronic parameters that allow researchers to not only quantify levels of grammaticalization, but also make comparisons.

2.2. GRAMMATICALIZATION CLINES. The unidirectionality of grammaticalization is often considered by linguists to be a statistically universal trend (Haspelmath 2004). It is this

directional constraint which allows for the quantification of the differing grammaticalization of a structure over time. This is an adaptation of Lehmann (2004: 156) who says:

Consider now the methodological situation of someone who wants to argue that a certain constellation of data is empirical evidence for grammaticalization. Such evidence must take the following form:

1. There are two historical stages of language L, earlier L1 and later L2.
2. L1 has form F1 and L2 has form F2, such that F2 is diachronically identical with F1.
3. F2 is more grammatical than F1.

A cline of grammaticalization can therefore be created such as that in Figure 1, adapted from Norde (2020: 6).

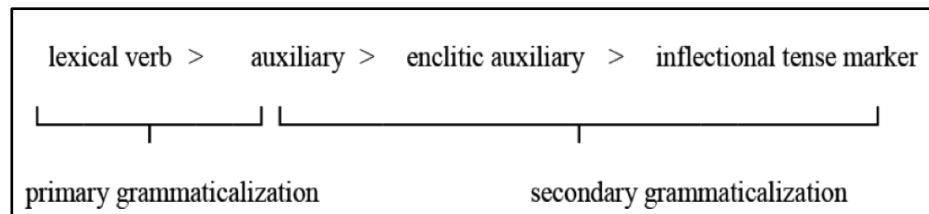


Figure 1. Visualization of a possible grammaticalization cline

Although Lehmann (2004: 3) states that “In order to use some data as empirical evidence for grammaticalization, it must be historical data”, this methodology does not account for genetically related languages that are at different stages of development. It is this methodology that I here propose to apply to the positioning of Neapolitan. Specifically, one can use this approach by considering Latin as the 'L' language, with Italian and Neapolitan both representing different stages.

This has been variously done previously with the Romance Grammaticalization Cline (RGC), pioneered by Lamiroy (1999) and followed by Lamiroy and De Mulder (2012), De Mulder and Lamiroy (2012), Lamiroy and Pineda (2017), and Miola (2017), among others. Lamiroy and De Mulder (2012) found French to be the most holistically grammaticalized Romance language with Italian following and Spanish after. This positioning was based on the features of auxiliaries, the simple past, existential sentences, mood, and demonstratives. Methodologies used to develop the RGC vary with some linguists preferring to look at only one specific feature's position between languages and with other linguists examining a conglomeration of features to place entire languages more holistically on the continuum. Recent studies of the first type include Poplack et al. (2018) who look at the cline of just the subjunctive mood and Burgo (2010) who studies the present perfect usage for the preterite. Examples of the second more holistic type of study include the previously mentioned Lamiroy and De Mulder (2012) as well as Miola (2017) who positioned Piedmontese as more grammaticalized than French based on various features. This study follows the first approach and intends to do the same with Neapolitan, looking at just two features, specifically past auxiliary selection and possessive adjective position relative to their usage in Italian. Data are taken from a 17th-century literary text titled *Lo Cunto de li Cunti* (Basile & Petrini 1976 [1634]), and its translation (Basile & Croce 1925). These data are complemented with comparison relative to contemporary Neapolitan and Italian.

**3. Corpus and methodology.** The present study positions the Neapolitan language on the RGC. Neapolitan is the southern Italo-Romance language with the largest number of native users,

exceeding 7.5 million (UNESCO 2010). Despite this significant population, the language has been relatively ignored in scholarly literature (Ryan 2018). This research therefore aims to contribute to a greater understanding of this understudied language. I will first describe the corpus and then the theoretical underpinnings used for analysis. I will subsequently outline the two features investigated and the methodology in corpus analysis.

3.1. CORPUS. The corpus used for analysis was the 1635 *Lo Cunto de li Cunti* by Giambattista Basile. The book is written completely in early literary Neapolitan and is a collection of 50 fictional fairy tales within the narrative structure of frame stories. *Lo Cunto de li Cunti* has been rather neglected within the Italian literary sphere, only having seen a revival in interest in the last century, evidenced by the lack of a popular Italian translation until Benedetto Croce's translation in 1925. Some of the most influential and well-known stories which find their origin in *Lo Cunto de li Cunti* include Cinderella (*La Gatta Cennerentola*), Puss in Boots (*Gagliuso*), and Sleeping Beauty (*Sole, Luna e Talia*).

The edition and translation used are the Mario Petrini (Basile & Petrini 1976 [1634]) Neapolitan philological edition and the acclaimed Benedetto Croce (1925) translation. These works were chosen as the Petrini book is the only full Neapolitan edition of the text since the 19th century, and the Croce translation because it is the first and most acclaimed Italian translation being both "eminently readable" and "impeccably annotated" (Canepa 1999: 29).

The source text and translation combined produced a corpus of approximately 200,000 words. The books are open access, having been digitized by Hathi Trust and Internet Archive, and are downloadable in a .txt format, making them available for use in concordancing and text analysis software such as the here-used AntConc (Anthony 2023). A small amount of coding was required to format the .txt files in such a way suitable for analysis. This primarily included the conjoinment of hyphenated words at the end and beginning of lines to be sure all words were correctly represented and searchable.

A visit was also made to the Biblioteca Nazionale Universitaria di Torino to consult some of the earliest 1788 print versions of the Neapolitan text (Basile 1788 [1634]) which themselves were used in the curation of the Petrini (Basile & Petrini 1976 [1634]) Neapolitan edition. This was done to confirm that the Petrini edition accurately and faithfully represents the grammar from the earliest texts. Petrini (Basile & Petrini 1976 [1634]) claims to have modified and standardized orthography while keeping word order and morphology consistent. Luckily, I found this proclaimed transcription methodology to be faithful and reliable. Petrini's orthographic standardization of the text therefore does not affect the analysis of the here-studied morphosyntactic variables and only helps to improve the book's reliability and searchability as a corpus for morphosyntactic analysis.

For the data on contemporary Standard Italian, comparative positioning relied on prior descriptive analyses (Berruto 2017; Ledgeway 2009; Maiden 1995, Maiden & Robustelli 2013). Introspection was also used to complement these analyses. To mitigate potential bias, introspective insights were carefully cross-checked against the descriptive sources, ensuring a balanced and reliable comparative framework.

3.2. QUANTIFYING GRAMMATICALIZATION. As mentioned above, the parameters of *paradigmaticity*, *paradigmatic variability*, and *syntagmatic variability* are the theoretical grounding behind the present quantification of grammaticalization. I argue these paradigmatic and syntagmatic parameters to be the most straightforward in application when comparing two parallel texts. This is because semantic analyses can be challenging and sometimes impossible

through text analysis due to the inherent subjectivity in interpreting meaning across historical texts. Similarly, a phonetic erosion analysis, which would need to track sound changes, presents practical difficulties in written text analysis since the phonetic realization of sounds is often not directly reflected in orthography until the latest stages of grammaticalization. Given these limitations, a focus on syntagmatic positions and pragmatic frequencies of structures offers a more straightforward and reliable method for analyzing grammaticalization in historical texts, where changes in form and structure can be more easily quantified and compared. Ultimately, these are the only parameters that can easily be quantified based on comparative frequencies, word orders, and sizes of a structure rather than requiring a more complex variable constraint analysis which may require the reliance on semantic assumptions.

### 3.3. AUXILIARY USAGE.

3.3.1. NEAPOLITAN. As in Italian, there are two ways to express perfectivity in Neapolitan: the *passato remoto* (preterite) and the *passato prossimo* (present perfect). In Neapolitan, the distinction between the two is maintained in speech in a fashion more similar to standard *literary* Italian (Ledgeway 2009: 439) or Spanish. While the preterite is used for singular perfect actions in the past without any reference to duration, the present perfect is reserved for recent perfective actions in the past with relevance in the present moment of enunciation. For this reason, the *passato prossimo* has a much higher frequency in speech as compared to writing.

The preterite is formed morphologically on the verb, inflecting for person and number. The present perfect, on the other hand, is formed with a grammaticalized auxiliary plus the past participle. Neapolitan seemingly only ever displayed two principal auxiliaries. As Ledgeway states, “The evidence provided to us by early Neapolitan texts reveals a distribution of the two auxiliaries *avere* and *essere* according to a model of alternation that is also largely valid for modern Italian” (*translation mine*) (Ledgeway 2009: 591). Based on Standard Italian, this means that the *essere* ‘be’ auxiliary was used mostly for unaccusative intransitive participle verbs, reflexive participle verbs, and reciprocal constructions. *Avere* ‘to have’ was used for all others.

Although historically these auxiliary verbs were both used, modern Neapolitan is developing only one verb, *avè(re)*, used in the active perfective form, or past perfect (Ledgeway 2009: 591). Cennamo (2008: 127) contends that there were early signs of the *avè(re)* auxiliary encroaching upon the functional domains of the *essè(re)* auxiliary already in the 14th and 15th centuries. Contemporarily, *avè(re)* has been so grammaticalized that another verb has emerged as the expression of possession: *tenè(re)* ‘to have’. These two verbs, *avè(re)* and *essè(re)*, are the only two historically attested auxiliaries in Neapolitan texts when forming the *passato prossimo* construction (Ledgeway 2009: 591). An example from the source text is:

- (1) Neapolitan (Basile & Petrini 1976 [1634])
- |  |            |                 |               |               |           |     |
|--|------------|-----------------|---------------|---------------|-----------|-----|
| Vui  | stesse     | v'              | <b>av-ite</b> | <b>f-atto</b> | la causa  | ... |
| 2PL  | yourselves | 2PL-ACC         | have-2PL.PRES | do-PTCP       | the cause |     |
| ‘You yourselves have done the process ...’ |            |                 |               |               |           |     |
| <b>av-ite</b>                              |            | <b>ferm-ato</b> | lo decreto    |               |           |     |
| have-2PL.PRES                              |            | sign-PTCP       | the decree    |               |           |     |
| ‘you have signed the decree’               |            |                 |               |               |           |     |

Whereas Italian has grammaticalized various verbs as auxiliaries beyond *essere* and *avere* such as the cross-linguistically rare *venire* (to come) as a passive auxiliary, the Neapolitan of the present source text does not display any further options beyond *essere* and *avere*.

Unmarked past participle formation in Neapolitan is done by adding the morphemes *-ato* to the *-à* class of infinitive verbs, *-uto* to the *-è* class, and *-ito* to *-ì*. However, the *-ì* infinitive verb class is rather unsystematic, often defaulting to *-uto* or *-ətə* as in the case of *veni(re)* (*venutə*) ‘to come’ and *dici(re)* (*dicitə*) ‘to say’ (Ledgeway 2009: 49). Furthermore, in all verb classes, the final vowel in the past participle is almost always pronounced as a shwa resulting in forms like /par.ˈla.tə/ and /man.ˈdʒa.tə/. In the Petrini (Basile & Petrini 1976 [1634]) edition of *Lo Cunto de li Cunti*, however, the final vowel is always written regardless of pronunciation.

Historically, in the *passato prossimo*, the past participle needed to agree with the direct object it was referencing in gender and number, no matter the auxiliary used and no matter if the direct object referenced was explicit or not. This is because of the construction’s grammaticalization trajectory. In the early stage, it involved possessing a noun that was the object of a verb, such as ‘I have the printed papers’ (with ‘printed’ obligatorily agreeing in number and gender with ‘papers’). In the late stage, this participle adjective was mapped onto the past tense, as in ‘I have printed the papers’. Today, in Italian, this gender and number agreement is only obligatory with the *essere* auxiliary. Ledgeway (2009: 852–855) notes that this agreement has been very highly reduced in modern Neapolitan but is not rare. *Lo Cunto de li Cunti* shows instances of agreement (2) and instances of lack thereof (3):

(2) Neapolitan (Basile & Petrini 1976 [1634])

O	che	mal-a	<b>jornata</b>	<b>a-ggio</b>	<b>f-att-a</b>
O	COMP	bad-F	day.F	have-1SG.PRES	do-PTCP-F

‘Oh I have made it such a bad day’

(3) Neapolitan (Basile & Petrini 1976 [1634])

Chi	m’	<b>h-a</b>	<b>f-att-o</b>	<b>sta</b>	<b>varva</b>
Who	1SG.ACC	have-3SG.PRES	do-PTCP-M	DET.F	chin.F

‘Who gave me this beard?’

3.3.2. ITALIAN. Modern Italian also has the *passato remoto* formed morphologically on the verb, inflecting for person and number. For the *passato prossimo*, modern Italian has two verbs grammaticalized as perfect auxiliaries: *essere* for unaccusative intransitive and reflexive participle verbs and *avere* for all the others. The correspondent examples in the translation from example 1 above are:

(4) Italian (Basile & Croce 1925)

Voi	stesse	vi	<b>si-ete</b>	<b>f-atto</b>	il processo	...
2PL	yourselves	2PL-ACC	be-2PL.PRES	do-PTCP	the process	

‘You yourselves have done the process ...’

<b>av-ete</b>	<b>firm-ato</b>	il decreto
have-2PL.PRES	sign-PTCP	the decree

‘you have signed the decree’

Past participle formation in Italian is most frequently done by adding the morphemes *-ato*, *-uto* and *-ito* for *-are*, *-ere*, and *-ire* verb classes respectively. However, there exists vast allomorphy in many high-frequency verbs such as *mettere* (*messo*) ‘to put’ and *leggere* (*letto*) ‘to read’.

Historically, as in earlier Neapolitan, when using the *passato prossimo*, the past participle needed to agree with the direct object it was referencing in gender and number, no matter the

auxiliary used. In modern Italian, however, this gender and number agreement with the object is obligatory only with verbs which take the *essere* auxiliary (Maiden 1995).

3.3.3. HYPOTHESIS. The expansion of the *passato prossimo* into previously *passato remoto* preferring contexts, desemanticization of the auxiliaries, loss of gender agreement on participle, and the more recent reduction in auxiliary alternation show that, in both languages, the creation of the *passato prossimo* was clearly driven by a process of grammaticalization. The hypothesis being tested here is that the Neapolitan source text will show a reduced alternation in auxiliaries compared to the Italian translation, with *avere* appearing more often and in what are otherwise *essere*-requiring contexts in the Italian. This would show a greater level of grammaticalization in the perfect auxiliary paradigm. This is because, according to Lehmann’s parameter of *paradigmatic variability*, this movement towards a single auxiliary reduces the choices available to a speaker as well as the size of a grammatical paradigm, resulting in the *obligatorification* of the use of *avere*.

### 3.4. POSSESSIVE ADJECTIVE USAGE.

3.4.1. NEAPOLITAN. The most pragmatically common fashion of expressing possession in Neapolitan is via the use of a determiner possessive. In this case, the identity of the possessor is marked on the determiner possessive, the head of the determiner phrase. The determiner possessive indexes the person feature of the possessor, as well as the gender and number of the possessed. For example:

(5) Neapolitan (Basile & Petrini 1976 [1634])

<b>bene</b>	<b>mi-o</b>	non conviene lo farete venire a	<b>lo</b>	<b>palazzo</b>	<b>mi-o</b>
good	POSS-1.M.SG	not desirable the let you come to	DET	palace.M	POSS-1.M.SG
<i>‘My beloved, it is not decorous to let you come to my palace’</i>					

The same possessive adjectives also function as possessive pronouns. Historically, the possessive adjective could appear in either a prenominal or postnominal position (Ledgeway 2009: 230), but contemporary Neapolitan has a high preference for the postnominal position<sup>1</sup>. Additionally, the use of a definite article in possessive determiner phrases was more common historically, though its presence is now variable in contemporary Neapolitan (Ledgeway 2009: 230).

3.4.2. ITALIAN. In Italian, the same determiner possessive is used but with greater variability in prenominal/postnominal position and with a greater use of a determiner:

(6) Italian (Basile & Croce 1925)

<b>bene</b>	<b>mi-o</b>	non è decoroso farti entrare	<b>a-l</b>	<b>mio</b>	<b>palazzo</b>
good	POSS-1.M.SG	not is decorous to let you to enter	to-DET	POSS-1.M-SG	palace.M
<i>‘My beloved, it is not decorous to let you come to my palace’</i>					

In Italian, the postnominal use is especially common when used as a term of address or endearment for an individual, functioning as a type of diminutive, so one could expect Italian to prefer this postnominal position when the possessed has human qualities (Rosenbach 2008). In Italian,

---

<sup>1</sup> Some upper-southern Italo-Romance languages display enclitic possessives with kinship nouns and in other cases of inalienable possession (Russo 2021). This has been observed in medieval Neapolitan texts but with a “rather limited distribution” (Ledgeway 2009: 269). Enclitic possessives are not attested in the present corpus.

postnominal possessive adjective position is also perceived as poetic (sometimes *overly* poetic) when used in speech.

3.4.3. HYPOTHESIS. The hypothesis is that the Neapolitan source text will show a reduced alternation between prenominal and postnominal possessive adjectives relative to Italian with an overall preference for postnominal. It is hypothesized that both Italian and Neapolitan will prefer the postnominal position more frequently with human referents as the possessed. This would result in the conclusion of a higher level of grammaticalization in Neapolitan because, according to Lehmann's (2015 [1982]) grammaticalization parameter of *syntagmatic variability*, this preference for a single postnominal position in Neapolitan reduces the choices a speaker has in word order, resulting in the *obligatorification* of the use of the possessive adjective in a postnominal position.

### 3.5. CORPUS ANALYSIS.

3.5.1. PERFECT AUXILIARY. To quantify a movement towards *avè(re)* in places where *esse(re)* used to be required, all uses of these verbs with a past participle were extracted. This entailed creating a full preliminary paradigm of all conjugations of both *avè(re)* and *essè(re)* based on Ledgeway (2009) and then fleshing out the paradigm with missing forms and checking all alternatives with further text analysis. Being a largely unstandardized written language for so long, particularly in old texts, many orthographic variations are present. For example, in the 2nd person singular future simple conjugation of *esse(re)*; *sarràje*, *sarraie*, *sarrai*, *sarai*, *saraie*, *sarài*, among others, are all attested across Neapolitan texts<sup>2</sup>. The 1st and 2nd person plural were not attested as the work is largely written in third person and quotes are most commonly in the 1st person singular form.

Using AntConc, a free concordance software for .txt files, (Anthony 2023) all attested verb forms were checked to be sure to include all conjugations within the source text in the concordance. Once all verb forms had been extracted, all forms not used with a past participle were eliminated. The AntConc concordance produced around 3500 tokens from just the Neapolitan source. For the analysis, 250 random tokens were selected from the Neapolitan source text which had a matching translation in the Italian text. The translation had to use a parallel structure, i.e. the Italian also had to use an auxiliary + past participle. Furthermore, the translation could not use a particle verb with different qualities than those of the source text which would alter auxiliary expression; for example, if Neapolitan used a transitive verb, an intransitive verb used in Italian would result in the token being excluded. A change in voice between source and translation also disqualified the token. These token exclusion criteria were upheld to ensure that the grammaticalization process in both languages could be accurately compared without introducing structural biases, given grammaticalization is so dependent on construction and context (Traugott 2003). The auxiliary verb used in both languages — *avere* or *essere* — was then coded.

3.5.2. POSSESSIVE ADJECTIVE POSITION. A similar methodology was followed for the study of possessive adjective position. A paradigm of all possessive adjectives was formed based on Ledgeway (2009) then subsequently fortified by and cross-checked by combing through the text

---

<sup>2</sup> While Petrini's transcription criteria were mainly focused on standardizing orthography, he does claim to have maintained the numerous variations in the spelling of the same word when no specific spelling was most common or correct. This includes the 3SG FUT *sarà/sarrà* as well as other conjugations of *esse(re)* and *ave(re)*, making it imperative that the paradigm was completely filled before analysis.

with AntConc (Anthony 2023). The concordance produced around 500 tokens of the use of a possessive adjective.

250 random possessive adjectives were then selected from the Neapolitan text and combed through the translation to identify and compare the same 250 possessives in the Italian. Again, the translation needed to be parallel with the possessed, person quality, and number remaining consistent between source and translation. The adjective was then coded as postnominal or prenominal, along with whether the possessed referent had [+human] qualities or not. This included not only names of specific people but also examples like (6) *bene mio* where *bene* ‘my beloved’ refers to a character in the story.

3.5.3. STATISTICS. The chosen statistical test for both analyses was a chi-squared test of association as is common in these types of historical change analyses with one categorical dependent variable (Gries 2013: 151).

#### 4. Results

##### 4.1. PERFECT AUXILIARY.

4.1.1. QUANTITATIVE ANALYSIS. Table 1 shows the results of the perfect auxiliary analysis, showing that, while Italian has an *essere/avere* ratio moving towards the expected direction, the difference between Neapolitan and Italian is not statistically significant. As expected, Neapolitan shows a stronger preference for the *avere* auxiliary over Italian’s ratio, but more data from these texts would be needed to confirm this preference as statistically significant during the era in which the source text was written.

	Italian	Neapolitan
Essere	115	100
Avere	135	150

**p = .175422**

Table 1. Results of perfect auxiliary analysis

With this data, it is only possible to situate the source and translation relative to one another on a grammaticalization cline, with the 17th-century Neapolitan at the same position as the Italian translation:

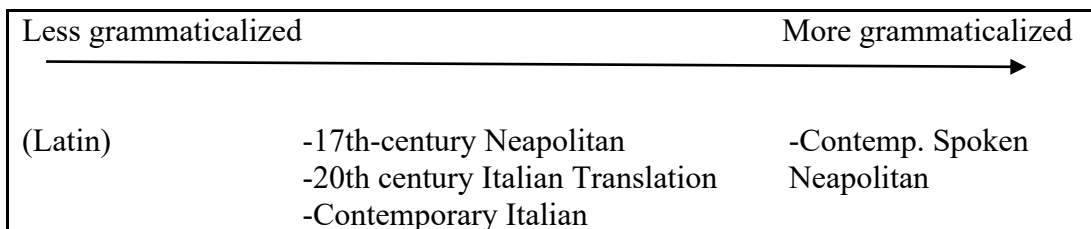


Figure 2. Grammaticalization cline of perfect auxiliary selection

Although the data is unable to statistically confirm the claim by Cennamo (2008) that this move toward one auxiliary was observed as early as the 14th and 15th centuries, these results could rather be taken in reference to periodization of the change.

4.1.2. QUALITATIVE ANALYSIS. These results suggest that, rather than showing a clear divergence between Neapolitan and Italian, the observed state in Neapolitan may be understood as part of a broader periodization of grammatical evolution. Periodization refers to the process of identifying specific stages in the development of a linguistic feature over time. This concept is central to research by Ayres-Bennett and Caron (2016), who demonstrated how grammatical changes in French could be traced across distinct historical periods based on linguistic evidence from translations. Similarly, the change observed in Neapolitan may be in an intermediate stage, where alternations in auxiliary selection are certainly present but have not yet statistically diverged from Italian's pattern. Therefore, the observed alternation may reflect a transitional phase of development within the larger continuum of change leading to today's single auxiliary system.

Further qualitative analysis shows that quite frequently, tokens were excluded from the parallel text comparisons due to voice differences. Oftentimes, the Italian translation would take up a passive voice where there was none in the Neapolitan. The reverse, passive voice in Neapolitan and active in Italian, was attested with considerably lower frequency. However, the most common cause of token exclusion was the use of *passato remoto*. For example, particularly when the Neapolitan used the *trapassato prossimo* (pluperfect), the Italian translation preferred the *passato remoto*. This importantly does not mean the Italian used the *passato remoto* more, however, as analysis from Italian to Neapolitan was not performed.

Tokens that were included in coding where the languages differed frequently was the construction of a progressive passive with clitic in both languages. For example:

(7) Neapolitan (Basile & Petrini 1976 [1634]: 409)

<b>ave-nno-se</b>	<b>lev-ato</b>	la	robba
have-PROG-REFL.3SG	take off-PTCP	DET	clothing
<i>'having taken the clothing off herself'</i>			

(8) Italian (Basile & Croce 1925: 416)

<b>esse-ndo-si</b>	<b>tol-ta</b>	la	roba
be-PROG-REFL.3SG	take off-PTCP	DET	clothing
<i>'having taken the clothing off herself'</i>			

In these cases, the Italian always preferred *essere* as expected whereas Neapolitan variably used both *avere* and *essere*. Historically, the presence of a reflexive verb and clitic would demand an *essere* auxiliary in both languages, but the Neapolitan of *Lo Cunto de li Cunti* was showing movement towards the *avere* auxiliary in previously *essere*-requiring contexts. However, this is where the stark differences end, with the Neapolitan and Italian by and large following similar auxiliary alternation patterns in this period as predicted by Ledgeway (2009: 439) and in contrast to Cennamo (2008).

## 4.2. POSSESSIVE ADJECTIVE POSITION.

4.2.1. QUANTITATIVE ANALYSIS. Table 2 shows the total possessive adjective positions used between the languages without consideration for human characteristics of the possessed.

	Italian	Neapolitan
Prenominal	95	5
Postnominal	155	245

**p < .00001**

Table 2. Overall results of possessive adjective position analysis

The Neapolitan source text almost categorically prefers postnominal position while the Italian is more mixed. Ultimately, if possession were to become a further grammaticalized bound affix in a language, it would first have to find an obligatory syntactic position, something the Neapolitan has nearly already accomplished in this text. The apparent *obligatorification* of the postnominal position in the Neapolitan text due to the loss of *syntagmatic variability* (Lehmann 2015 [1982]) regarding word order variation allows positioning of this feature along the RGC as such:

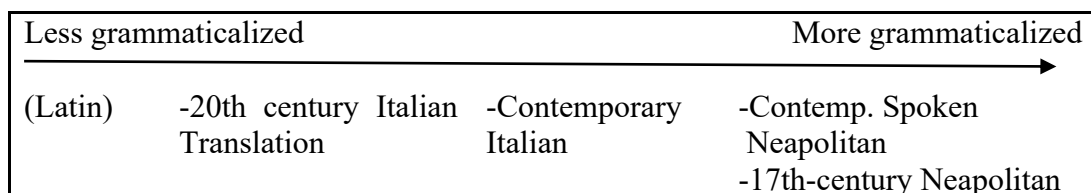


Figure 3. Grammaticalization cline of overall possessive adjective position

Turning toward the [ $\pm$ human] semantics of the possessed, the results are presented in Tables 3 and 4:

	Italian	Neapolitan
Prenominal [-human]	74	2
Postnominal [-human]	89	161

Yate's chi-squared: 86.493  
**p < .00001**

Table 3. Results of possessive adjective position analysis with [-human] possessed

	Italian	Neapolitan
Prenominal [+human]	21	3
Postnominal [+human]	66	84

Yate's chi-squared: 13.968  
**p < .0001**

Table 4. Results of possessive adjective position analysis with [+human] possessed

The results therefore show that this separation of pre/postnominal position in Italian is ever so slightly less pronounced when the possessed is [+human], but the difference between Italian and Neapolitan is still highly statistically significant. This adds further evidence to the observation that the grammaticalization process does not affect all usages of a construction at once but is rather conditioned by a wide number of pragmatic and constructional factors (cf.

Traugott 2003). However, it still shows strong evidence that this diversion in grammaticalization was very strong in this text and during this period, affecting the entire paradigm as well as specific paradigmatic situations.

4.2.2. QUALITATIVE ANALYSIS. It is hard to discern frequent places where the included tokens differed between the languages without a more in-depth variable constraint analysis. However, in general, tokens in Italian which preferred prenominal positioning were rather low on the animacy scale while still ranging from both concrete objects such as “crown” and “shoe” to abstract concepts such as “hopes” and “wants”.

The most common cause of exclusion of a token was the use of the preposition *di* in the Italian translation. *Di* is Italian’s “one construction fits all” (Aikhenvald 2012: 4) equivalent to the English ‘of’ as in *la casa di lei* ‘the house of her (her house)’. The variable exclusions of these tokens emphasizes what Aikhenvald calls the “special status of ‘me’ and ‘you’” (2012:10). In Italian, the use of *di* for first- and second-person singular possessives in a fashion that passes all of Aikhenvald’s language-internal tests for possession is, at best, pragmatically uncommon, and at worst, ungrammatical, e.g. *la casa di io* ‘\*the house of I’. Seeing as the paradigm size for possession in the first- and second-person singular is reduced, *paradigmatic variability* and therefore the grammaticalization trajectory of the construction are all affected. This is again, very pragmatics and construction specific, *à la* Traugott (2003).

**5. Discussion.** An overall analysis of the data reveals that, in some respects, namely possessive adjective position, 17th-Century Neapolitan was already more grammaticalized than the 20th century Italian translation. The results for the perfect auxiliary analysis were inconclusive in differentiating the source text from the translation.

As far as I am aware, this is the first study to have looked at Neapolitan with reference to a grammaticalization cline. It is also the first to produce a cline with data significantly coming from a parallel text and translation corpora comparison.

5.1. IMPLICATIONS AND LIMITATIONS. This study has implications for how the expression of competing structures with differing grammaticalizations across languages is researched, both in and out of translation. This includes correcting worries about using “mere text frequency” (De Troij & Van De Velde 2020: 1) in the study of grammaticalization, with translation decreasing the sometimes-misleading measure of frequency in grammaticalization research. This study also presents a viable methodology for researching grammaticalization synchronically, challenging the widely diachronic view of the phenomena.

This research also introduces implications back onto translation studies such as questions to consider like how the translator’s idiolect, language history, and bilingualism affect the translation and therefore expression of differing grammaticalized features. For example, Petrini notes in his Neapolitan edition that, “Basile’s language was accused of being Italianized; and in fact, if we consult our texts, we see that many variants of the critical editions are precisely in the direction of dialectal hyper-correctness without a shortage of inverse cases” (*translation mine*) (Basile & Petrini 1976 [1634]: 607). This is simultaneously a confounding factor but could also prove as a fruitful object of study. For example, if Croce was influenced by the original Neapolitan when translating into Italian and erred on the side of Neapolitan-like structures in the translation, there could be a greater tendency to use postnominal possessive adjectives as in the source text.

This then reflects back on questions of how grammaticalization theory can provide insight into translation practices. For example, maybe a translator wants their Italian translation of *Lo*

*Cunto de li Cunti* to have a ‘Neapolitan flare’ and will err on the side of differently grammaticalized structures more similar to what we see in the source text. Otherwise, being aware of the different grammaticalizations, they could do exactly the opposite and err on the side of structures less grammaticalized such as those in Italian.

The greatest limitation at the present time is the reliance on a single text, however this was done intentionally here to show the viability of the technique. The study's quantitative analysis is also based on a smaller sample size of a total 500 tokens for both perfect auxiliary and possessive adjective usage. A larger and more diverse data set would enhance the reliability and generalizability of the findings.

5.2. TRANSLATION AS A METHODOLOGY. Novel alongside the results was the methodology in using a source text and translation as parallel corpora for a comparative grammaticalization analysis. Many previous studies have focused on the implications that corpus-based contrastive linguistics can have in the field of translation studies, i.e. a corpus-based contrastive analysis approach applied to translation studies (Ramón García 2021; Shang 2022); however, the present study has done exactly the opposite. Ramón García (2021) sustains that contrastive linguistics can be considered as the primary and most essential foundation in translation theory, providing the transition between theory and practice (404); the present study has seen it in a slightly modified but still compatible way: that is, translation can be considered among the purest of sources for application in contrastive linguistics. However, for clarity, what is new about the present approach is not the use of translation for language change analysis but rather applying it to grammaticalization and the RGC. This technique was rather inspired by the methodologies in contrastive linguistics employed by researchers such as Ayres-Bennett and Caron (2016) and Enrique-Arias (2016) who study grammatical change in French and Spanish respectively through the usage of successive translations of a singular text.

Some of the advantages of this method are that one can bypass methodological concerns highlighted by both Szmrecsanyi (2016) and Enrique-Arias (2016), including potential ‘environmental’ content changes which affect the faithfulness of comparability. This is because as Granger (2010: 5) states, translations “are an ideal resource for establishing equivalence between languages since they convey the same semantic content”. Furthermore, this methodology could present itself as viable if one had access to a single text in a dead or low-resource language from the period one wants to analyze; the text would just need to be long enough, and a sufficient translation would need to exist to provide enough content for a quantitative and qualitative analysis.

Disadvantages to this methodology include the laborious nature of having to compare two different languages and corpora along with the need for an in-depth knowledge of both languages’ structures. This methodology also introduces complications. For example, if the translator is a bilingual individual, it is possible that the source language shines through in the translation, giving a false reading of the different grammaticalization levels and making the two languages appear closer to each other than they actually are (McLaughlin 2011; 2013). The methodology is also inherently limited due to comparison with a single text, making extrapolation difficult. However, this caveat could be avoided if multiple texts and translations by multiple authors/translators from the same period were to be analyzed and quantified together.

There is also the need to consider any changes that may have been made to the source text. For example, Petrini (Basile & Petrini 1976) found it necessary to modify various aspects of the Neapolitan source to make it more standardized, due to the lack of standardization of

orthography in the seventeenth-century. Luckily, the vast majority of these modifications were merely orthographic.

5.3. FUTURE DIRECTIONS. Future directions include the analysis of a broader range of texts and translations to validate these findings and explore additional grammatical features with more variable constraints. Replicability studies could also be done on different structures, languages, and genres of literature to confirm the viability of this methodology. Even the use of a separate additional translation of *Lo Cunto de li Cunti* for comparison with the here-found results could give an idea of the margin of error for the usage of a single text and translation as well as demonstrate how much impact the style of the translator has on differing grammaticalization. The present results could also be supplemented with analysis of the distinct standardization histories of Italian and Neapolitan and how they may have shaped grammaticalization patterns. Previous research shows that standardization practices are commonly “aimed at halting, regulating or redirecting the ongoing linguistic changes that can be identified as grammaticalization” (Laitinen 2004: 249). Italian’s slower grammaticalization could be attributed to its early standardization based on Tuscan during the Renaissance. Further research is needed to confirm this hypothesis.

**6. Conclusion.** This study has presented a comparative analysis of grammaticalization between 17th-century Neapolitan and Italian, employing a parallel text and translation-corpora methodology. By analyzing perfect auxiliary selection and possessive adjective position, the research has highlighted how the Neapolitan text exhibited more advanced stages of grammaticalization in specific areas compared to its Italian counterpart. The findings support the argument that historical texts and their translations can serve as valuable tools in tracing synchronic and diachronic linguistic changes, especially within the framework of grammaticalization. By employing this methodology, the potential of this approach for revealing historical linguistic changes and synchronic grammaticalization phenomena has also been demonstrated. Additionally, the study emphasizes the importance of considering the translator’s linguistic background in comparative linguistic studies, as their choices can influence the representation of grammaticalization in translated texts. Although more research is needed to generalize these findings, this research contributes to both historical linguistics and translation studies, offering a promising avenue for further investigation into grammaticalization theory and language change across Romance languages.

## References

- Aikhenvald, Alexandra Y. 2012. Possession and ownership: A cross-linguistic perspective. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds.), *Possession and ownership*, 1–64. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199660223.003.0001>.
- Anthony, Laurence. 2023. AntConc (version 4.2.4) [computer software]. Tokyo: Waseda University. <https://www.laurenceanthony.net/software>.
- Ayres-Bennett, Wendy & Philippe Caron. 2016. Periodization, translation, prescription and the emergence of classical French. *Transactions of the Philological Society* 114(3). 339–390. <https://doi.org/10.1111/1467-968X.12081>.
- Basile, Giambattista. 1788 [1634]. *Il Pentamerone del caualier Giovan Battista Basile ouero Lo cunto de li cunte trattenimento de li peccerille di Gian Alesio Abbattutis, tomo 1.* (G.-M. Porcelli, ed.).
- Basile, Giambattista & Benedetto Croce. 1925. *Il Pentamerone: ossia, La fiaba delle fiabe / Tra*

- dotta dall'antico dialetto napoletano e corredata di note storiche da Benedetto Croce.*  
Laterza.
- Basile, Giambattista & Mario Petri. 1976 [1634]. *Lo cunto de li cunti: ovvero, Lo trattenemiento de peccerille; Le muse napoletane e Le lettere.* G. Laterza.
- Berruto, Gaetano. 2017. What is changing in Italian today? Phenomena of restandardization in syntax and morphology: An overview. In Marco Cerruti, Claudia Crocco & Simone Marzo (eds.), *Towards a new standard*, 31–60. Berlin: De Gruyter.  
<https://doi.org/10.1515/9781614518839-002>.
- Burgo, Clara. 2010. A case of grammaticalization in the use of the perfect for the preterite in Bilbao Spanish. *Studies in Hispanic and Lusophone Linguistics* 3(2). 301–328.  
<https://doi.org/10.1515/shll-2010-1077>.
- Canepa, Nancy L. 1999. *From court to forest: Giambattista Basile's Lo cunto de li cunti and the birth of the literary fairy tale.* Detroit, MI: Wayne State University Press.
- Cennamo, Michela. 2008. The rise and development of analytic perfects in Italo-Romance. In Thráinn Eythórsson (ed.), *Grammatical change and linguistic theory: The Rosendal papers* (Linguistik Aktuell/Linguistics Today 113), 115–142. Amsterdam: John Benjamins.  
<https://doi.org/10.1075/la.113.05cen>.
- De Mulder, Walter & Béatrice Lamiroy. 2012. Gradualness of grammaticalization in Romance: The position of French, Spanish and Italian. In Kristin Davidse, Tine Breban, Lieselotte Brems & Tanja Mortelmans (eds.), *Grammaticalization and language change: New reflections* (Studies in Language Companion Series 130), 199–226. Amsterdam: John Benjamins.  
<https://doi.org/10.1075/slcs.130.08mul>.
- De Troij, Robbert & Freek Van de Velde. 2020. Beyond mere text frequency: Assessing subtle grammaticalization by different quantitative measures. A case study on the Dutch *soort* construction. *Languages* 5(55). 1–17. <https://doi.org/10.3390/languages5040055>.
- Enrique-Arias, Andrés. 2016. Dos problemas en el uso de corpus diacrónicos del español: perspectiva y comparabilidad. *Scriptum Digital* 1. 85–106.
- Granger, Sylviane. 2010. Comparable and translation corpora in cross-linguistic research: Design, analysis and applications. *Journal of Shanghai Jiaotong University* 2. 14–21.
- Gries, Stefan Th. 2013. *Statistics for linguistics with R: A practical introduction*, 2nd edn. Berlin: De Gruyter Mouton.
- Haspelmath, Martin. 2004. On directionality in language change with particular reference to grammaticalization. In Olga Fischer, Muriel Norde & Harry Perridon (eds.), *Up and down the cline: The nature of grammaticalization* (Typological Studies in Language 59), 17–44. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.59.03has>.
- Heine, Bernd & Tania Kuteva. 2007. *The genesis of grammar: A reconstruction.* Oxford: Oxford University Press.
- Hopper, Paul J. 1991. On some principles of grammaticization. In Elizabeth Closs Traugott & Bernd Heine (eds.), *Approaches to grammaticalization* (Typological Studies in Language 19:1), 17–36. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.19.1.04hop>.
- Laitinen, Lea. 2004. Grammaticalization and standardization. In Olga Fischer, Muriel Norde & Harry Perridon (eds.), *Up and down the cline: The nature of grammaticalization* (Typological Studies in Language 59), 247–262. Amsterdam: John Benjamins.  
<https://doi.org/10.1075/tsl.59.13lai>.
- Lamiroy, Béatrice. 1999. Auxiliaires, langues romanes et grammaticalisation: Les auxiliaires: délimitation, grammaticalisation et analyse. *Langages* 33(135). 33–45.

- Lamiroy, Béatrice & Walter De Mulder. 2012. Degrees of grammaticalization across languages. In Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of grammaticalization*, 1st edn., 302–317. Oxford: Oxford University Press.  
<https://doi.org/10.1093/oxfordhb/9780199586783.013.0024>.
- Lamiroy, Béatrice & Anna Pineda. 2017. Grammaticalization across Romance languages and the pace of language change: The position of Catalan. *Linguisticae Investigationes* 40(2). 304–331. <https://doi.org/10.1075/li.00007.lam>.
- Ledgeway, Adam N. 2009. *Grammatica diacronica del napoletano*. Berlin: Walter de Gruyter – Max Niemeyer Verlag. <https://doi.org/10.1515/9783484971288>.
- Lehmann, Christian. 2004. Theory and method in grammaticalization. *Zeitschrift für Germanistische Linguistik* 32(2). 152–187. <https://doi.org/10.1515/zfgl.2004.32.2.152>.
- Lehmann, Christian. 2015 [1982]. *Thoughts on grammaticalization*, 3rd edn. Berlin: Language Science Press.
- Maiden, Martin. 1995. *A linguistic history of Italian*. London: Longman.
- Maiden, Martin & Cecilia Robustelli. 2013. *A reference grammar of modern Italian*, 2nd edn. London: Routledge.
- McLaughlin, Mairi. 2011. *Syntactic borrowing in contemporary French: A linguistic analysis of news translation*. London: Legenda/Modern Humanities Research Association and Maney.
- McLaughlin, Mairi. 2013. News translation as a source of syntactic borrowing in Italian. *The Italianist* 33(3). 443–463. <https://doi.org/10.1179/0261434013Z.00000000055>.
- Miola, Emanuele. 2017. The position of Piedmontese on the Romance grammaticalization cline. *Folia Linguistica* 51(1). 133–167. <https://doi.org/10.1515/flin-2017-0004>.
- Narrog, Heiko & Bernd Heine. 2021. *Grammaticalization*, 1st edn. Oxford: Oxford University Press.
- Norde, Muriel. 2020. Grammaticalization in morphology. In *Oxford research encyclopedia of linguistics*. Oxford: Oxford University Press.  
<https://doi.org/10.1093/acrefore/9780199384655.013.733>.
- Poplack, Shana, Rena Torres Cacoullos, Nathalie Dion, Rosane de Andrade Berlinck, Salvatore Digesto, Dora Lacasse & Jonathan Steuck. 2018. Variation and grammaticalization in Romance: A cross-linguistic study of the subjunctive. In Wendy Ayres-Bennett & Janice Carruthers (eds.), *Manual of Romance sociolinguistics*, 217–252. Berlin: De Gruyter.  
<https://doi.org/10.1515/9783110365955-009>.
- Ramón García, Noelia. 2021. Contrastive linguistics and translation studies interconnected: The corpus-based approach. *Linguistica Antverpiensia, New Series – Themes in Translation Studies* 1. 393–406. <https://doi.org/10.52034/lanstts.v1i.27>.
- Rosenbach, Anette. 2008. Animacy and grammatical variation: Findings from English genitive variation. *Lingua* 118(2). 151–171. <https://doi.org/10.1016/j.lingua.2007.02.002>.
- Russo, Michela. 2021. The possessive enclitics with kinship nouns in Italo-Romance and the possessive determiners in the Francoprovençal of Faeto: Two syntactic linearizations? *Zeitschrift für Romanische Philologie* 137(1). 217–251.  
<https://doi.org/10.1515/zrp-2021-0007>.
- Ryan, John M. 2018. Embracing Neapolitan as a language which is key to the reconstruction of early Romance. *Theory and Practice in Language Studies* 8(11). 1377.  
<https://doi.org/10.17507/tpls.0811.01>.
- Shang, Xin. 2022. When contrastive analysis meets translation studies: A historical perspective.

- Target: International Journal of Translation Studies*. 186–214.  
<https://doi.org/10.1075/target.20027.sha>.
- Szmrecsanyi, Benedikt. 2016. About text frequencies in historical linguistics: Disentangling environmental and grammatical change. *Corpus Linguistics and Linguistic Theory* 12(1). 153–171. <https://doi.org/10.1515/cllt-2015-0068>.
- Traugott, Elizabeth Closs. 2003. Constructions in grammaticalization. In Brian D. Joseph & Richard D. Janda (eds.), *The handbook of historical linguistics*, 1st edn., 624–647. Malden, MA: Wiley. <https://doi.org/10.1002/9780470756393.ch20>.
- UNESCO. 2010. *Atlas of the world's languages in danger* (Christopher Moseley, ed.), 3rd edn. Paris: UNESCO.
- Varvaro, Alberto. 2013. Latin and the making of the Romance languages. In Martin Maiden, John Charles Smith & Adam Ledgeway (eds.), *The Cambridge history of the Romance languages*, 1st edn., 6–56. Cambridge: Cambridge University Press.  
<https://doi.org/10.1017/CHO9781139019996.002>.