Linguistic Typology at the Crossroads 5-2 (2025): 224-288 ISSN 2785-0943

Governor-driven subjunctive selection: a variationist study from Latin to Romance

SALVIO DIGESTO
CARLETON UNIVERSITY - DEPARTMENT OF FRENCH (OTTAWA, CANADA)

Submitted: 05/06/2024 Revised version: 24/10/2024
Accepted: 07/08/2025 Published: 25/02/2026

Abstract

A key parameter to measure (dis)continuity between Romance languages and the ancestor language, Latin, is mood selection, especially the use of the subjunctive as opposed to the indicative according to syntactic environments and semantic meanings supposedly conveyed. This study explores the trajectory of mood selection in *that*-completive clauses from Latin to modern Romance languages, with a particular focus on Italian. It challenges the assumption that subjunctive selection is semantically motivated, highlighting recent variationist findings that identify the main clause verb's lexical identity as the subjunctive's major predictor. By employing a variationist sociolinguistic approach, this research delineates the evolution of subjunctive selection, revealing that contemporary patterns in Romance languages may reflect a continuation of lexicalization processes initiated in Vulgar Latin, rather than a recent desemanticization phenomenon. This analysis contributes a nuanced understanding of subjunctive selection, offering new perspectives on its function and evolution across Romance languages.

Keywords: subjunctive mood; variationist sociolinguistics; lexicalization; Romance languages; Vulgar Latin; completive clauses.

1. Tracing the conditioning of subjunctive selection

Languages are typically classified according to the presence or absence of a given structural feature, in some cases taking account the feature's frequency or its

distribution. Typologists have examined the subjunctive-indicative distinction in complementation, a well-attested phenomenon in several language families. Romance languages, such as Italian (1), present a dedicated verbal inflectional paradigm inherited from their ancestor language, Latin.[1]

(1) a. *Credo che tutto ritorn**a***$_{[3SG.PRES.IND]}$. (C.511.264)[2]
'I believe that everything comes back.'
b. *Ma può darsi che ritorn**i***$_{[3SG.PRES.SUBJ]}$, *guarda*. (L.16.23)
'But she may come back, look.'

The Italian subjunctive is said to be used to express an attitude of uncertainty and the speaker's attitude towards the truth of the embedded proposition, as opposed to the indicative that is said to convey a more assertive statement (e.g., Whaley 1997; Schachter & Shopen 2007). Subjunctive morphology also supposedly conveys other types of meanings as well, such as indefinite, uncontrolled, or uncertain readings, highlighting its non-assertive nature (Terrell & Hooper 1974; Hooper 1975; Lindschouw 2011). According to Hooper (1975: 123), "complements that are assertions, that is, complements to assertive verbs, have indicative verb forms. […] complements to non-assertive verbs, and (imperative) complements to volitional verbs, all have subjunctive verb forms," suggesting a clear-cut semantic division between subjunctive and indicative contexts. It is said to be a mood fluctuating "between opinion and perception" (author's translation, Binazzi 2015: 259); projecting "a modality of uncertainty and doubt" of the event in question (author's translation, Simone 1993: 80); or representing "the intense and emotive degree, the particular and the personal, the doubt and the unreal, the unexpected and the surprising, the desired and the feared, the extraordinary and the exceptional" (author's translation, Dorigo 1951: 322). In the syntactic literature, the mechanism of mood selection is considered to be "implemented at a distance" (author's translation, Shlonsky 2006: 83), i.e., the main verb determines the type of complementizer and embedded mood, operating by checking mood features (Poletto 2000). This view represents a general assumption in the literature: the governor's semantic

[1] Italian: ita; Indo-European; Romance – Latin: lat; Indo-European; Italic.

[2] The codes identify the corpus, the numeric speaker code, and the line number at which the utterance occurred. Examples from two corpora will be shown throughout this paper, including the C-ORAL-ROM (2005) (C) and the *Lessico di frequenza dell'Italiano Parlato* (1993) (L).

characteristics determine the selection of the 'appropriate' mood for the embedded clause and governors, with *volition*, *emotion*, and *doubt* consistently deemed categorical subjunctive-selecting contexts. This "hereditary feature" that is the Romance subjunctive (author's translation, Cressot 1947: 139) would have inherited from the ancestor language the set of meanings ascribed to it in the Latin era, such as dubitative and hypothetical (Cressot 1947), or uncertain and irrealis (Leone 1949: 13; see also Mellet 1994; Perotti 1996). Italian, in particular, is acknowledged as having "retained from Latin [subjunctive] not only clear formal marking, but also substantial semantic motivation" (Harris & Vincent 1997: 303). However, especially within the Romance language family, whether subjunctive use is triggered primarily by semantic factors remains a matter of some debate in the literature.

1.1. Semantic and pragmatic considerations

While we cannot do full justice to the vast literature on the subjunctive, this section focuses on major trends and influential accounts of its usage and variability, with particular attention to Italian. Key cross-linguistic and foundational studies that have shaped these accounts are also referenced where relevant.

Earlier approaches have typically framed the indicative/subjunctive opposition in terms of a *realis* versus *irrealis* contrast, whereas more recent proposals rely on more refined interpretive parameters (Giorgi & Pianesi 1997: 201). These newer perspectives make the analyst's task more complex, as mood selection becomes closely linked to psychological or attitudinal factors, i.e., the speaker's mental stance at the moment of utterance (e.g., Costantini 2011: 39; Giannakidou & Mari 2015; Portner & Rubinstein 2013). For instance, Wandruszka's analysis of subjunctive mood in completive clauses outlines a hierarchy of subjunctive use based on three semantic groups: volitive, dubitative and factive predicates.

According to Wandruszka (1998: 416), volitive predicates such as those with *volere* 'to want'*, sperare* 'to hope'*, pretendere* 'to pretend', etc. are those that convey the will or intention of the subject. The core semantic property of such predicates is their orientation toward a state of affairs that is not yet realized but is instead desired, prevented, or otherwise projected. As a result, volitive complement clauses do not assert facts; rather, they represent the speaker's stance toward a hypothetical outcome. Wandruszka explicitly notes that a volitive modalized clause is fundamentally distinct from an assertion, and it is precisely this non-assertive nature

that makes the subjunctive particularly salient in such contexts. The subjunctive, in these cases, serves to signal that the embedded proposition is not presented as part of the actual world but as part of the speaker's projected or wished-for reality.

Dubitative predicates (Wandruszka 1998: 418) are characterized as belonging to the broader domain of epistemic modality, i.e., the modality of knowing and believing, where the subject evaluates the truth value or plausibility of a proposition. These predicates range across a continuum, from the speaker's certainty about the non-existence of a state of affairs to varying degrees of uncertainty regarding its existence. What distinguishes this class is not a binary division between true and false, but rather the speaker's degree of commitment to the embedded assertion. While a certain degree of epistemic uncertainty is often sufficient to license the subjunctive, Wandruszka notes that it is not a necessary condition, citing cases such as *non dubito che* 'I don't doubt that' or *sono sicuro che* 'I am sure that', which can still select the subjunctive despite expressing high epistemic commitment. The notion of dubitative or epistemic subjunctive is understood therefore in a broad sense, encompassing predicates like *non credo/credo* 'I don't think/I think', *è possible* 'it is possible', *dubito* 'I doubt', *suppongo* 'I suppose', *pare* 'it seems', and *sono sicuro che* 'I am sure that'. Dubitative or epistemic subjunctive usage is classified across a range of structural types. First, verbs such as *dubitare* 'to doubt' express an absence of belief or knowledge (*non sapere se p* 'not to know if *p*'), and typically select the subjunctive, even under negation (*non dubito che sia intelligente* 'I don't doubt s/he is intelligent'), due to the semantic alignment with conviction rather than factuality. Second, verbs that negate a proposition (*negare che p* 'to deny that *p*') also favour the subjunctive, since they question the truth of the embedded clause. Even when negated themselves (*non negare che p* 'not to deny that *p*'), these predicates retain subjunctive selection. Third, predicates expressing uncertain belief or assumption (*credere* 'to believe', *pensare* 'to think', *ritenere* 'to consider') select the subjunctive when the embedded proposition is not presented as factual. In the presence of negation, interrogative and conditional clauses, this pattern is reinforced (1998: 435-436), as these environments weaken epistemic commitment. Additionally, verbs of communication such as *dire* 'to say' and impersonal constructions such as *dicono che* 'they/people say that', *si dice che* 'it is said that' can trigger subjunctive when they imply reported belief rather than assertion (in a similar vein, Chinellato 2001). Similarly, verbs such as *parere* 'to seem' and *sembrare* 'to seem' tend to select the subjunctive when they convey inference rather than fact (*pare che sia stanco* 'it seems that he is tired'), while modal predicates

like *è possibile che* 'it is possible that', *può darsi che* 'it might be that' nearly categorically require the subjunctive.

Adjectival and nominal predicates indexing epistemic stance also influence mood. Adjectives like *possible* 'possible', *probabile* 'probable', *improbabile* 'improbable' co-occur with the subjunctive, whereas factive adjectives (*certo* 'certain', *evidente* 'evident', *sicuro* 'sure') are more compatible with the indicative. Epistemic nouns such as *dubbio* 'doubt', *opinione* 'opinion', *ipotesi* 'hypothesis', and expressions like *l'idea che* 'the idea that', *l'impressione che* 'the impression/feeling that', favour subjunctive morphology by designating non-factual propositions.

Wandruszka highlights a third major class: factive predicates (which he also refers to as *tematici* or *fattivi di valutazione*; 1998: 418-419). These include evaluative and emotional predicates where the embedded proposition is not in doubt: expressions such as *dispiacersi* 'to regret/to be sorry' or *essere felice* 'to be happy' presuppose the truth of the subordinate clause. What licenses the subjunctive in these cases is not the speaker's uncertainty, but rather the presuppositional status of the embedded content. As Wandruszka observes, one can only experience joy or regret about a state of affairs that is assumed to be real. This presupposition is thus a necessary precondition for the matrix predicate. Moreover, unlike epistemic predicates, factive-evaluative predicates maintain the same truth-value assumptions even when embedded under negation, interrogation, or conditional constructions. The consistent use of the subjunctive in such contexts reflects its core semantic function: to signal that the clause is not part of the speaker's primary communicative act but rather a backgrounded, non-assertive element of discourse.

Within a formal semantic-syntactic framework, Giorgi and Pianesi (1997) aligning in many respects with Wandruszka's discourse-functional approach. Both accounts converge on the idea that subjunctive mood arises when the embedded proposition is not presented as an assertion within the common ground. Subjunctive is consistently licensed by volitional, desiderative, directive, and evaluative predicates, all of which impose modal or counterfactual conditions on the truth of the embedded proposition. In contrast, factives (called *true factives* by Hooper 1975) like *dispiacersi* 'to regret/to be sorry' select the subjunctive due to their emotional-evaluative content and causal structure (Farkas 1992), whereas semifactives such as *sapere* 'to know' select the indicative, given their lack of a modal component. This contrast aligns with Wandruszka's classification of factives and further clarifies why *sapere* 'to know'

resists subjunctive even under negation unless additional semantic or temporal factors intervene, i.e., negation and past tense such as in *non sapevo che* 'I didn't know that'.

In contexts of belief predicates (e.g., *credere* 'to believe', *pensare* 'to think', *ritenere* 'to consider'), Giorgi and Pianesi observe that subjunctive mood is generally licenced. As also noted by Wandruszka (1998), the presence of negative and interrogative operators as well as conditional clauses tend to increase the likelihood of using a subjunctive. Giorgi and Pianesi interpret this pattern in terms of conversational backgrounds and doxastic alternatives: the subjunctive arises when the embedded proposition is not anchored in the speaker's or subject's common ground, but rather evaluated across possible worlds or non-actual scenarios. This conception overlaps with Wandruszka's treatment of epistemic predicates but provides a more formalized grounding. Additionally, they emphasize mood variation under *dire* 'to say' and other *verba dicendi*. While such verbs generally select indicative as the modal base consists of shared facts, and the ordering source is null, operators such as negation or interrogatives can license the subjunctive even under *dire* 'to say', as in *Mario non ha detto che Giuseppe sia impazzito* 'Mario didn't say that Giuseppe has gone mad'. This would demonstrate that mood choice under *verba dicendi* is not purely lexical but sensitive to clause-external conditions.

Although less recent, Schmitt Jensen's (1970) analysis of subjunctive selection attempts to reconcile the competing paradigms in mood selection analysis by returning to structural description while not entirely abandoning semantic considerations. Rejecting the exclusivity of either a purely semantic or purely syntactic model, Schmitt Jensen argues for a dual-layered approach that grounds itself first in observable syntactic structure, i.e., the position and function of elements within the clause, and subsequently seeks interpretive clarity through semantic regularities. Based on a meticulous analysis of a diverse corpus of contemporary Italian, which includes literary texts, essays, journalistic prose, and film dialogues from the 20th century, as well as grammars, Schmitt Jensen identifies four distinct groups of governing predicates (1970: 125-126): a) those that typically govern the subjunctive; b) those that govern the indicative; c) those allowing mood alternation signaling semantic opposition; d) those allowing alternation without any semantic contrast. These types are not derived from theoretical principles but from corpus-driven observation. Only after establishing these distributional categories, he identifies recurrent semantic tendencies within each group, treating these not as defining criteria but as emergent patterns. Within Group A, comprising predicates

that predominantly govern the subjunctive, three main semantic domains recur (1970: 232-233): volition and necessity (*volere* 'to want', *desiderare* 'to desire', *ordinare* 'to command/to order'), doubt and uncertainty (*dubitare* 'to doubt', *temere* 'to fear', *essere incerto* 'to be uncertain'), and subjective or affective evaluation (*essere contento* 'to be happy', *dispiacersi* 'to be sorry/to regret', *è strano che* 'it is strange that'). These predicates typically introduce a non-factual or projected state of affairs and reflect the subject's psychological stance or emotional response rather than a commitment to the truth of the embedded proposition. Group B, by contrast, includes predicates that govern the indicative and presuppose the factuality of the embedded clause (1970: 234). These include epistemically strong verbs such as *sapere* 'to know', *vedere* 'to see', and *essere sicuro* 'to be sure', along with nominal and adjectival expressions like *la certezza* 'the certainty' or *è chiaro che* 'it is clear that', all of which align the embedded content with the speaker's or subject's belief in its reality. Group C consists of predicates that allow both moods, with mood alternation reflecting a semantic contrast (1970: 235-237). For verbs such as *credere* 'to believe', *pensare* 'to think', *dire* 'to say', or expressions like *ho l'impressione che* 'I have the impression/feeling that', the subjunctive marks the embedded proposition as epistemically or pragmatically distanced from the speaker's assertional base, while the indicative aligns it with the common ground. Mood choice in these contexts is said to be often sensitive to discourse factors such as negation, interrogation, or conditional operators, which modulate speaker commitment. Finally, Group D includes cases in which both moods are found without a clearly identifiable semantic opposition (1970: 238). Here, unlike previously discussed accounts, predicates such as *è possibile che* 'it is possible that', *può darsi che* 'it might be that', and impersonal expressions like *si dice che* 'it is said that' exhibit mood alternation without consistent interpretive consequences, raising the possibility that such variation reflects register, regional norms, or stylistic preference rather than semantic contrast. Another interesting aspect raised by Schmitt Jensen's analysis is that the subjunctive functions as a subordination marker, signaling a tighter syntactic and semantic dependency between the matrix and embedded clause. In contrast, the indicative, by virtue of its assertive force, would grant the embedded clause a higher degree of autonomy, effectively reducing the strength of the subordinating link. However, this notion of increased dependency under subjunctive is somewhat problematic, since the author does not provide independent syntactic or semantic diagnostics beyond mood itself to justify such a

hierarchy of integration. This raises the question of whether mood alone can index clause dependency or whether other structural features must be considered.

It is worth noting that across descriptive, formal, and also prescriptive accounts of the subjunctive mood, stylistic and register-based variation in mood choice is acknowledged, though often only marginally. In Wandruszka's analysis, informal, colloquial, or familiar registers are mentioned as contexts in which the indicative may surface in place of the expected subjunctive, even where the semantics would typically require it, for instance, in volitive constructions. Giorgi and Pianesi similarly note intralinguistic variation, such as alternation with *credere* 'to believe', but do not examine register effects, focusing instead on modal semantics within the standard language. Schmitt Jensen, more explicitly, identifies a category of predicates that allow alternation without consistent semantic opposition, and cautiously suggests that stylistic variation, register, and genre may play a role. Yet none of these accounts systematically investigate such alternation beyond intuition or textual observation. These departures from the expected subjunctive are often dismissed as mere matters of style, but the very ease with which they occur raises deeper questions about the force and interpretive necessity of mood marking. If the subjunctive indeed signals epistemic or attitudinal stance, it is unclear why such signaling should be suspended in informal speech. This calls into question theoretical models based primarily on introspective judgments, which may fail to capture the structured nature of vernacular variation. Spontaneous speech data, in fact, can often challenge some widespread assumptions on the conditioning of subjunctive selection (Digesto 2019: 69). For instance, as reported above, a *verbum dicendi* such as *dire* 'to say' may combine with evaluative features, thereby shifting the meaning from that of a purely communicative verb to one that places the proposition within the speaker's opinion or point of view (see Wandruszka 1998; Chinellato 2001, amongst others). This would signal non-engagement with the truth of the embedded proposition (e.g., *dicono* 'they say', *si dice* 'it is said'), as shown in (2):

(2) *Marlon Brando è Marlon Brando. Dicono che* ***sia***$_{[3SG.PRES.SUBJ]}$ *omosessuale.* (C.201.402)
'Marlon Brando is Marlon Brando. People say that he's homosexual.'

However, one might argue that the same indeterminacy can be conveyed when the speaker opts for the indicative counterpart as well (3):

(3) *Qualcuno dice che ci si mett**e**$_{[3SG.PRES.SUBJ]}$ meno.* (C.207.218)
'Some people say that it takes less time.'

Likewise, if we assume that subjunctive morphology signals greater uncertainty, this assumption is challenged by examples such as (4) below. In (4b), the interrogative context suggests that the speaker is uncertain about the truth of the proposition in the complement clause, i.e., which neighbourhood the referent lives in. Nonetheless, the indicative is selected:

(4) a. *Mi sembra che **sia**$_{[3SG.PRES.SUBJ]}$ una felce.* (C.505.141)
'It seems to be that it is a fern.'
b. *Mi sembra che abit**i**$_{[3SG.PRES.IND]}$* a Santa Lucia, *è possibile?* (L.314.28)
'It seems to me that she lives in Santa Lucia, is it possible?'

1.2. Subjunctive mood in quantitative research

Relatively few studies have examined subjunctive selection in Italian through a quantitative lens. Among those that do, two core questions tend to recur: whether the subjunctive is declining, and whether variation in mood choice is semantically governed, particularly by the lexical-semantic class of the matrix predicate. Several studies focus on predicates of volition (Bonomi 1993; Loengarov 2006), hope (Bonomi 1993; Lombardi Vallauri 2003; Santulli 2009), and fear (Bonomi 1993; Santulli 2009), under which the subjunctive is assumed to be categorical, while alternation is expected with predicates of opinion (Bonomi 1993; Lombardi Vallauri 2003; Santulli 2009; Veland 1991). In many cases, the scope is limited to a small set of governors, for example most often *credere* 'to believe', *pensare* 'to think', and *ritenere* 'to consider', which are frequently cited as showing variability and decline. Veland (1991) finds that these verbs of opinion favour the subjunctive, while Bonomi (1993) reports near-categorical selection under opinion (93%) and volition (90%) predicates, with more variability under judgment/evaluative predicates (60%). Gatta (2002) suggests that when the indicative is selected, speakers compensate through other modalizing strategies, such as conditional sentences or expressions like *a mio parere* 'in my opinion', *secondo me* 'to me', and *la mia sensazione* 'my feeling'. Veland (1991: 219) also reports that adverbials encoding [+certainty] can disfavour the subjunctive. Santulli (2009) argues that the subjunctive remains productive in speech, though is

more less likely to occur in first-person contexts (e.g., *penso* 'I think'). Lombardi Vallauri (2003) tests broader semantic classes and finds that the subjunctive is favoured under hope predicates, but disfavoured with verbs of communication.

These studies reveal several interesting patterns, but their divergent data types and methodological choices hinder comparability. While some authors argue for ongoing productivity (Bonomi 1993; Santulli 2009; Veland 1991), others suggest a shift toward the indicative and interpret this as a change in progress (Lombardi Vallauri 2003; Schneider 1999). Yet, none of these claims is supported by diachronic evidence. Voghera (1993) points to low overall rates in speech, which she interprets as part of a broader decline of the subjunctive. Overall rates vary widely, from 2% (Voghera 2001) to over 80% (Santulli 2009), but are based on different corpora and non-uniform criteria, precluding direct comparison.

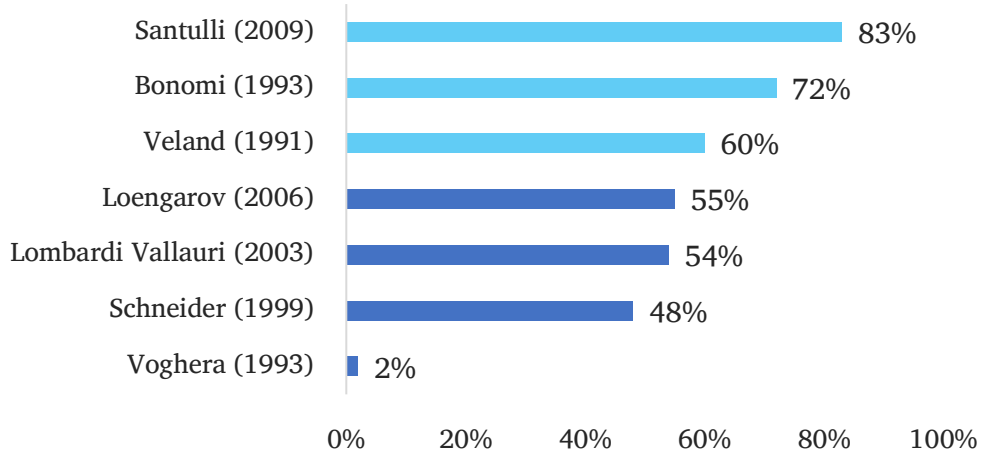


Figure 1: Comparison of the rate of subjunctive selection in previous quantitative research. With the exception of Voghera (1993), overall rates were calculated on the basis of the results provided by the authors in their publications. Results of the studies that made use of LIP corpus are in dark blue.

In terms of testing semantic conditioning and the contribution of some lexical types, some studies impose some restrictions by extracting only embedded forms of *essere* and *avere* (Lombardi Vallauri 2003; Loengarov 2006) or restrict subject person to either first (Santulli 2009) or third (Loengarov 2006). These a priori constraints may obscure the full range of variation, violating the Principle of Accountability (Labov

1972), since the subjunctive might occur with other governors and contexts in the selected data[3].

The data sources are similarly heterogeneous. In addition to spontaneous speech corpora, studies have drawn on newspapers (Bonomi 1993; Santulli 2009; Veland 1991), literary texts (Soliman 2002), text messages (Santulli 2009), online forums (Loengarov 2006; Santulli 2009), and television talk shows (Gatta 2002). Even studies using the same dataset, such as the LIP corpus, differ in their selection criteria, excluded contexts, and analytical procedures, whether they report raw counts or rates. As a result, a single governor may be found to favour the subjunctive in one study and disfavour it in another. For example, *credere* 'to believe' is reported as favouring in Veland (1991) and disfavouring in Bonomi (1993); *ritenere* 'to consider' appears to favour in Santulli (2009) and disfavour in Schneider (1999).

1.3. Towards an account of the lexicalization of the subjunctive in Romance

Following Poplack's seminal work on subjunctive variability (Poplack 1990; 1992; also Poplack et al. 2013), recent variationist sociolinguistic research has addressed the above-mentioned issues and examined the use of the subjunctive in running discourse, particularly in complement clauses, highlighting a trend that is shared across the varieties of Romance languages examined (e.g., Italian, Canadian French, Parisian French, Brazilian Portuguese, Mexican Spanish) and uncovering a systematic selection process guided more by lexical governors than semantic roles (Digesto 2021; Kastronic 2016; Poplack et al. 2018). This pattern suggests a path of development divergent from prescriptive norms and from common assumptions that the subjunctive is used in discourse in contexts that convey meanings such as uncertainty, doubt, volition, necessity, etc., revealing instead a trajectory towards *lexical routinization*—a process where the selection of the subjunctive mood becomes progressively restricted to a select array of lexical triggers (Poplack et al. 2018). This lexicalization in discourse challenges the traditional understanding of the subjunctive

[3] The reader may note that, compared to earlier quantitative studies using the LIP corpus, the number of verbs analyzed in the current study differs. This is due both to the fact that previous studies imposed *a priori* lexical choices, whereas the present analysis considered all governors that triggered the subjunctive in context (i.e., determined corpus-internally), and to the restriction of the current study to a sub-section of LIP containing only spontaneous, unscripted conversations. For a detailed discussion of stylistic and social variation in contemporary speech, see Digesto (2019, 2021).

as semantically driven, positing an alternative account where syntactic environments and lexical collocations underpin its usage. The insights gleaned from examining speech data from a several Romance languages and a wide temporal range have been pivotal in characterizing the subjunctive not as a feature constrained by semantic properties but as one deeply embedded within the lexicon of certain governing verbs. Poplack and her associates (1992; 2013) notably illustrate this phenomenon in French, where the verbs *falloir* 'it is necessary', *vouloir* 'to want', and *aimer* 'to like' significantly trigger the selection of the subjunctive in speech. This phenomenon, found also in a study of subjunctive usage in Parisian French (Kastronic 2016), indicates a systematic entrenchment of the subjunctive within particular lexical environments cross-varieties. Likewise, contrary to prevailing assumptions in the literature regarding the semantic/pragmatic motivation for selecting the subjunctive mood in embedded completive clauses in Italian, Digesto (2019; 2021) showed that the use of subjunctive is mainly restricted to a handful of main clause verbs (e.g., *bisognare* 'it is necessary'*, credere* 'to believe'*, pensare* 'to think'*, sembrare* 'it seems') and a single embedded verb (suppletive forms of *essere* 'to be'). Similar trend, i.e. a less productive use of the subjunctive in speech that is mainly triggered by a handful of matrix verbs and embedded verbs, particularly favoured with irregular/suppletive morphology, is observed in other Romance languages examined through the lens of variationist method, such as Portuguese and Spanish (Poplack et al. 2018; also Berlinck 2019). This pattern towards lexicalization, identified by some scholars as a case of subjunctive attrition (Bybee et al. 1994), would mark a departure from a semantically-based usage of the subjunctive through desemanticization or 'semantic bleaching'. In other words, the loss of a semantic contribution, through a process of *ritualization* "brought about trough routine repetition" (Haiman 1994: 3; see also Bybee 2003), results in *obligatorification* (Lehmann 1995: 12), or what is referred to elsewhere, in variationist research, as *lexical routinization* (Poplack et al. 2018: 238).

Lexical routinization implies that semantics contribute minimally to none to the choice of subjunctive, favouring structural and lexical constraints over semantic underpinnings (Haiman 1994; Bybee 2003; Lehmann 1995). While variationist research on contemporary usage of the subjunctive mood, based on recordings of actual speech data highlighted common patterns in Romance pointing mostly to a lexicalized use in discourse, variationist sociohistorical research showed that similar trends have been operative in Italian since at least the 16th century (Digesto 2019) and since the 19th century in Canadian French (Poplack et al. 2013) and in Brazilian Portuguese (Berlinck 2019), at our knowledge, the oldest historical benchmarks investigated by following

variationist methodology. Results showed that, contrary any expectation of a semantically-based use of the subjunctive in previous stages of these languages, the grammar of usage has remained remarkably stable, with no semantic contributions evidenced in previous stages of the targeted varieties but rather a persistence of a lexicalized pattern as outlined above.

We are still left with the question: Can the lexical conditioning of the subjunctive be considered an innovative pattern that emerged in the vernacular daughter languages, assuming that in the ancestral language, Latin, the subjunctive mood was triggered by semantic factors? Or is it a case of inheritance and transmission from Latin, suggesting that the lack of semantic contribution *already* characterized the use of the subjunctive at its roots? Building on previous variationist research, this contribution explores the use of the subjunctive within a corpus of Latin data, in a variationist framework. This approach ensures comparability with results observed in Italian, as well as more broadly across other Romance languages, particularly French and Portuguese.

2. The evolution of the subjunctive from Latin to Romance

It is generally observed that a distinction between *meaningful* and *meaningless* morphology was found in Latin, which stems from the observation that the subjunctive mood was used in various syntactic environments, in both main and various subordinate clauses. While its application in main clauses is considerably more extensive in Latin than in modern Romance languages (Harris 1978), the subjunctive is predominantly associated with *subordination* (Magni 2009: 244; Noonan 2007; Jespersen 1924; Palmer 2001). Traditionally, the opposition between indicative and subjunctive moods in Latin is seen as carrier of semantic differences (Harris 1974: 169). However, during the Vulgar Latin period, the subjunctive would have begun to function more as a conventional or stylistic marker in certain contexts, indicating a shift towards a predominantly formal and mechanical attraction of the embedded mood (Handford 1947: 149). It is commonly accepted that the shift from paratactic to hypotactic syntax prompted a reorganization of independent clauses into more complex, dependent structures (Haudry 1973; Harris 1978; Murphy 2008). This reorganization is said to mirror the subjunctive's original optative value in earlier independent structures, rendering its morphology *meaningful*, especially when influenced by verbs expressing volition or order, as well as by verbs of hindering or preventing (Magni 2009: 245; Bybee et al. 1994).

PARATAXIS	>	HYPOTAXIS
volo; veniat$_{SBJV}$		*volo ut veniat*$_{SBJV}$
'I wish it; let him come'		'I want him to come'

Figure 2: Shift from parataxis (with subjunctive used in the main clause) to hypotaxis (with subjunctive used in the subordinate clause) in Latin. Adapted from Harris (1978: 168).

The shift to hypotactic structures led to the gradual permeation of the subjunctive in syntactic contexts traditionally associated with the indicative, such as indirect questions and adverbial sentences, which are syntactically considered *non-harmonic* with the original main clause meaning (Magni 2009: 247-250). This broadened use of the subjunctive denotes a departure from its original meaning and rather becoming a marker of subordination (Magni 2009: 260; Harris 1978).

Another significant syntactic development affecting subjunctive use in subordinate clauses, and its inheritance in Romance languages is the typological shift from OV (Object-Verb) to VO (Verb-Object) word order (Murphy 2008):

- *Left-branching subordination* included a) infinitival constructions, b) participial constructions, and c) subjunctive mood with no additional marker of subordination;
- *Right-branching subordination* included overt pre-posed subordinating conjunctions, i.e. subordination is marked pre-verbally.[4]

This shift led to an increase in explicit subordinating conjunctions and reinforced the subjunctive's syntactic subordinating function, moving away from a semantically-driven selection process. It suggests a preference for finite complementation over non-finite forms such as the *Accusativus cum Infinitivo*, promoting a widespread use of the subjunctive in right-branching structures. Despite some claims of SVO being the colloquial norm in earlier Latin (Pinkster 1990: 188), this shift likely furthered the use of the subjunctive as a standard marker of subordination in Latin, weakening any hypothesis of a semantically-based usage.

Similarities and discrepancies related to the type of matrix verbs that select the subjunctive are well-documented between Latin and Italian (and Romance in general): where in Latin verbs of thinking and saying typically select the indicative,

[4] Herman (1989) similarly noted, providing quantitative evidence, that in Late Latin, there was a syntactic competition between finite complements and non-finite complements (*Accusativus cum Infinitivo*). He pointed out that "before the verb, AcI is practically the only possibility, whereas after the verb, the binary choice between AcI and *quod/quia* clauses becomes possible" (*ibid.* 133).

supposedly categorically, Italian sees permeation of the subjunctive into these contexts. On the other hand, necessity, emotive and volitive verbs do not contrast in terms of mood selection between the ancestor and the daughter language—all supposedly triggering subjunctive categorically (Harris & Vincent 1997: 67; Disterheft & Viti 2010: 248; Clackson 2011: 137; Magni 2009; Palmer 2001). However, the main discrepancy lies in the type of constructions these main-clause verbs govern in the targeted languages: verbs of saying and thinking, as well as volitive verbs are generally considered quasi-categorical contexts for selecting *non-finite* constructions in Latin, as opposed to Italian, and to some extent in other Romance languages.

The expected outcome in relation to the type of main predicate can be schematized in the following way:

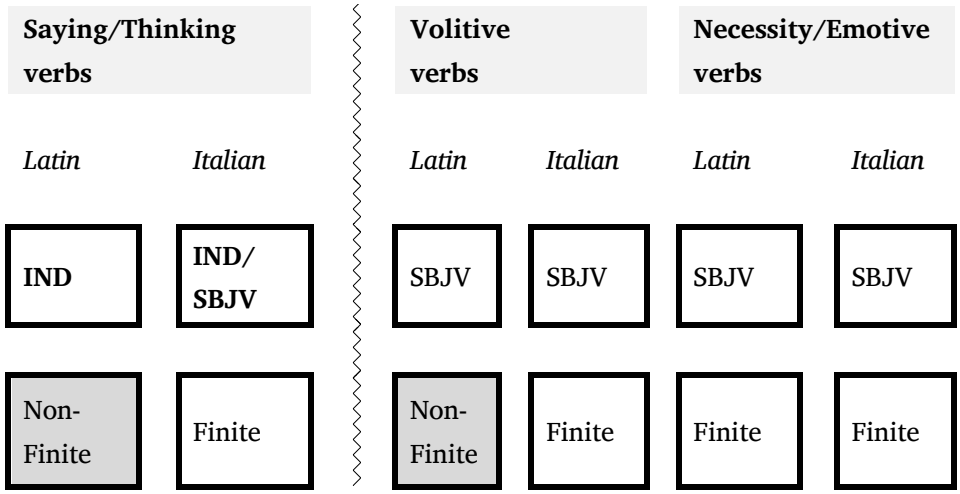


Figure 3: Prediction of mood selection according to the type of main predicate, in Latin and Italian, based on what is generally reported in the literature.

3. A variationist approach on subjunctive selection

3.1. Methodology

The approach to understanding subjunctive selection through the lens of variationist sociolinguistics offers a methodologically robust framework, placing empirical analysis at its core. This perspective posits that language variation is not random but is systematically structured (*ordered heterogeneity*; Labov 2001). It pivots on the fundamental notion that speakers, embedded within their communities, make use of alternate linguistic forms to express identical meanings or functions. This is the

abstract construct of the *linguistic variable* (Labov 1972; Labov 1984). Within this framework, the subjunctive mood is analyzed as a variable element, a variant form that alternates in discourse with other forms (such as the indicative) to express the same meaning or function in discourse (Poplack 2011: 212; Poplack & Levey 2010: 398). However, for many accounts of mood selection, the subjunctive and the indicative are ascribed distinct functions, which posit a methodological caveat when considering these two moods as variants that alternate to express equivalent meanings or functions in discourse. One way to circumvent this issue could be via the identification of those contexts where the subjunctive is *supposed* to occur and subsequently ascertain whether it actually does occur there in the data. However, previous research has highlighted the lack of agreement on which contexts or meanings *exactly* should trigger the subjunctive and therefore the rules governing subjunctive usage remain somewhat generally elusive (see Poplack et al. 2013, for French; Digesto 2019, for Italian). Following Poplack (1992, 2013, and Poplack at al. 2018), we identify the contexts of use *corpus-internally*, by locating all the contexts where the subjunctive is *actually used* to determine where it *could* be used. We identified all the unambiguous subjunctive forms and exhaustively extracted them from the dataset. This method enables the identification of which main clause verbs, called *governors*, were accompanied by the subjunctive, yielding the list of all the governors that selected the subjunctive in the dataset. Once the lexical identities of the governors are identified, the next step is to go back through the data and to extract all of the variants that compete with the subjunctive. This process is known as *circumscribing the variable context*. It allows to fully account for the variation in discourse by including "every case where the variable element occurs in the relevant environments as we have defined them" (Labov 1972: 72), but also where it could have occurred but it *did not* (Poplack & Tagliamonte 2001: 89). Following Poplack et al. (2019: 229), the variable context is here defined as every tensed embedded clause governed by a matrix element—whether verbal, nominal, or adjectival—that triggered the subjunctive at least once.

During the extraction phrase of the data, homophony between the subjunctive and the indicative was considered, in order to select unambiguous subjunctive morphology. Homophony is found in Latin for a wide variety of forms, mainly belonging to the active voice and overlapping between future indicative and perfect subjunctive. For instance, 1st conjugation verbs (e.g., *amo* 'I love') between future perfect indicative and perfect subjunctive: 2SG (*amaveris*), 3SG (*amaverit*), 1PL

(*amaverimus*); 2PL (*amaveritis*), 3PL (*amaverint*); 2nd conjugation verbs (e.g., *moneo* 'I warn'): 2PL (*monueritis*) and 3PL (*monuerint*); 3rd (e.g., *dico* 'I say') and 4th conjugation verbs (e.g., *audio* 'I hear'): 1SG future indicative and present subjunctive (*ducam*) at the active voice as well as at the passive voice (*ducar*); a few forms of the future perfect indicative and perfect subjunctive such as 2SG (*duxeris*), 3SG (*duxerit*), 1PL (*duxerimus*), 2PL (*duxeritis*), 3PL (*duxerint*); as well as 2SG, 3SG, 1PL, 2PL, 3PL between perfect future indicative and perfect subjunctive of almost all irregular verbs such as *sum*, *possum*, *eo*, *volo*, *nolo*, *malo*, *fero* and *do*.[5] On the other hand, homophony in Italian is limited to 2nd person singular of the present indicative and the present subjunctive with first group verbs (*-are*), 1st person plural of the present indicative and the present subjunctive of all three conjugation groups, and finally, 2nd plural of the simple past indicative and the imperfect subjunctive of all three conjugation groups.

The method applied in the current study allows for direct comparability between findings from synchronic and diachronic research on Italian and other Romance languages, and those derived from Latin data. We implemented the *Comparative Variationist Method* (Poplack & Tagliamonte 2001): rather than modeling change as proceeding directly from stage A to B, thereby overlooking the transition period from point A to point B in time and consequently portraying change as very abrupt, characterising the comparative method (Campbell 2013), we assume that a given language inherits the variants of a particular function from its ancestor, along with the linguistic conditioning of their variability. Thus, the genetic relationship is determined on the basis of the *conditioning* on variability, as revealed by which factors trigger the use of the subjunctive in discourse, rather than relying solely on the forms themselves.

3.2. Data

The contemporary data analyzed in this study come from two well-established corpora of spoken Italian: the *Lessico di frequenza dell'italiano parlato* (LIP; De Mauro et al. 1993) and the *C-ORAL-ROM Corpus of spoken Romance languages* (Cresti & Moneglia 2005). Both corpora include recordings of conversations, dialogues, and monologues, primarily collected in major urban centres such as Milan, Florence,

[5] 'I am', 'I can', 'I go', 'I want/I wish', 'I do not want/I do not wish', 'I prefer', 'I carry/I bring', 'I give'.

Rome, and Naples. For the purposes of this analysis, only those recordings capturing naturally occurring, unscripted speech in informal contexts were retained, while any data from institutional or pre-planned settings (e.g., televised interviews or press briefings) were systematically excluded. While invaluable for empirical research, these corpora are not without limitations. As previously noted (Digesto 2019: 44; Digesto 2021), there are sampling biases: Florentine data dominate the informal speech in C-ORAL, and LIP lacks socio-demographic metadata, making it difficult to verify whether speakers are truly representative of their local communities; where speaker information is available, the data tend to skew toward individuals with high levels of education. Nonetheless, these corpora offer a rare opportunity to conduct quantitative analysis of subjunctive usage in *spontaneous speech*. All data were manually concordanced, and all unambiguous instances of subjunctive morphology were retained for analysis. Given the prestige associated with the subjunctive, its frequency may be somewhat inflated in these sources. However, as shown in previous studies (e.g., Poplack & Tagliamonte 2001; Poplack & Levey 2010), while frequency may vary with social factors, the structural conditioning of variation is typically stable across genres and speaker groups.

3.3. Selecting the Latin benchmark

Variationist research mainly focuses on spoken usage, and particularly *vernacular*, i.e. "the style in which the minimum attention is given to the monitoring of speech" (Labov 1972: 208), since this is not only characterized by inherent variability but it also appears to be "the most systematic data for linguistic analysis" (Labov 1984: 29). This might be particularly challenging for analyzing early language stages or ancestral languages, as significant changes often emerge in spoken rather than written form. Without recordings or direct evidence of these stages, older texts cannot be reliably considered accurate representations of vernacular speech (Ayres-Bennett 2000; Nevalainen & Raumolin-Brunberg 2017: 26; Romaine 1982: 14). Written texts often fall short of capturing non-standard features; even texts intended to reflect spontaneous speech are shaped by prescriptive norms or linguistic ideology. The inclusion of non-standard linguistic features or sociolinguistic stereotypes in texts might also reflect the author's intent to depict lower-class/uneducated speech rather than an authentic portrayal of vernacular norms (Labov 1994: 11). To address this,

researchers have turned to speech-like surrogates, notably popular theatrical plays that aim to replicate the daily speech of different social classes, including the lower, working, and upper/aristocratic classes (Palmer 2001; Haudry 1973; Murphy 2008).

Since this research aims to explore the continuity and innovation of subjunctive use from Latin to Romance, with a specific emphasis on Italian, to assess how and to what extent contemporary usage aligns with historical practices in the ancestor language, it is essential to 1) select the specific variety of Latin for investigation and 2) choose texts that resemble speech. First, this emphasizes the need to focus on Vulgar Latin (VL), the spoken precursor to the Romance languages, rather than Classical Latin, primarily associated with written texts. Second, the choice of the data to examine fell onto an example of material that reflects features of VL and non-standard linguistic behaviour: we selected the *Satyricon*, a comic romance attributed to Petronius and dated to the 1st century (Durante 1981: 29). Specifically, the chapters of the *Cena Trimalchionis* ('Trimalchio's dinner party') constitute the VL corpus adopted in this study. Despite the limitation of the sample, with 12,576 words, it offers a unique glimpse into the informal, speech-like Latin that deviates from the normative Classical tradition.

This distinctive source allows a deep dive into the vernacular Latin of a different social stratum, showcasing a vivid portrayal of 'popular' speech through the dialogue of Trimalchio and his rude dining companions. This representation is celebrated as "the happiest achievement of ancient realism" (Scarsi 1996: xix), filled with elements like stories, proverbs, stereotypes, gossip and beliefs and providing more naturalistic forms of expression that are valued in variationist research. Though not a modern theatrical work, the *Satyricon* was intended for live performance to an audience of average citizens (Scarsi 1996), distinct from the elite circles often depicted in classical texts. Through Petronius's work, we gain insights into the *sermo plebeius*, the common speech of the first century A.D. (Palmer 1988: 153), offering a valuable perspective for understanding the use of the subjunctive in Vulgar Latin and its evolution in Romance. The chapters of the *Cena* not only contain a vocabulary that is "forceful, coarse, often indecent" (Palmer 1988: 152-153), but they also lay bare some explicit linguistic choices diverging from the standard of Classical Latin. For instance, errors of declension (7), vernacular verb forms (8), as well as the non-standard use of indicative instead of the normative subjunctive (9).

(7) *Lacte gallinaceum si quaesieris invenies.*
[*lacte* instead of the normative accusative form *lac*] (XXXVIII.249)[6]
'You can have cock's milk if you desire.'

(8) *Itaque statim urceatim plovebat*$_{[IMPRF.IND.3SG/ACT]}$.
[instead of *pluebat*$_{[IMPRF.IND.3SG/ACT]}$] (XXXVIII.249)
'And immediately the rain came down in bucketfuls.'

(9) *Cum interim nemo curat quid annona mordet*$_{[PRES.IND.3SG/ACT]}$.
[instead of *mordeat*$_{[PRES.SBJV.3SG/ACT]}$] (XXXVIII.249)
'Meanwhile, nobody cares how high the market price is.'

Despite the limitations of the historical dataset, we can use this data to uncover the internal conditioning of the subjunctive, which serves as major component of comparative analysis. Through systematic comparison of VL and Italian, and subsequently Romance, we will be in a position to ascertain whether and to what extent the subjunctive usage differs from the ancestor language to a daughter language.

3.4. Operationalizing hypotheses

Building on previous variationist studies (Poplack et al. 2013; Poplack et al. 2018), this study investigates semantic and lexical conditioning of subjunctive selection in Italian and Latin, focusing on the following parameters:

i. The semantic category of the governor
ii. The lexical identity of the governor
iii. The structure of the matrix clause and polarity
iv. The presence of other indicators of non-factual modality
v. The type of complementation

As outlined above, a recurrent claim in the literature is that specific semantic properties of governing verbs drive subjunctive selection. This study tests such claim by coding each occurrence according to parameters (i), (iii), and (iv), as detailed in

[6] The codes identify the chapter of the Satyricon and the line number at which the utterance occurred.

the ensuing section. Additionally, every token was coded based on the lexical identity of the governor. This analysis focused on three aspects: 1) the frequency of subjunctive usage for each governor, 2) the proportion that each governor represents within the entire governor pool, and 3) the extent to which each governor accounts for subjunctive morphology. A high score on these measures could indicate limited productivity and suggest a trend toward lexicalization of the subjunctive. If the lexical identity of the governor contributes to subjunctive selection *independently* of the intended meaning, this could signify that the productivity of the subjunctive is declining, and its use is confined to specific lexical governors.

Going beyond the envelope of variation, we also explore the type of complementation preferred by each governor, aiming to discern the extent to which the subjunctive use in Latin was influenced by the type of the complement (finite versus non-finite). This analysis could reveal the degree of syntactic contrast and to what extent this affects mood selection, particularly subjunctive, in Latin as opposed to the daughter languages.

In variationist research, a quantitative approach to modeling variation and change is paramount:

> The effect of a given factor is inferred by comparing its individual rate of subjunctive to the overall rate for the pooled data […] If a factor shows a rate of subjunctive selection higher than the overall rate, its effect is deemed favouring. The greater the difference, the stronger the effect. Likewise, the same method is applied when deeming the effect disfavouring (rate of subjunctive selection of a given factor lower than the overall rate). A neutral effect is deemed when there is no substantial difference between the rate of subjunctive of a given factor and the overall rate. (Digesto 2021: 13).

Employing the same analytical method and quantitative framework as earlier variationist research enables cross-linguistic comparison of findings. By examining the aforementioned parameters, both in isolation and in combination, this study seeks to examine whether the subjunctive was a productive and semantically-driven mood in Latin, and whether its productivity and semantic contribution were subsequently weakened in Romance languages. This could be evidenced by the diminishing role of semantic factors in the selection of the subjunctive and its increasing entrenchment in specific lexical environments (Lehmann 1995; Bybee 2003; Haiman 1994), as

demonstrated by the lexical routinization of the subjunctive in Romance (Poplack et al. 2018).

4. Results: contemporary spoken Italian

Overall results are based on the extraction and coding of the dependent variable, i.e., the grammatical mood (subjunctive or indicative). A first observation is that mood selection under governors that triggered the subjunctive at least once in the dataset shows robust variability: the subjunctive is the majority variant, occurring in 66% of cases (N = 404/616), followed by the indicative at 34% (N = 212/616), confirming that alternation remains frequent and that subjunctive mood is apparently quite productive in discourse. This also establishes our benchmark for interpreting any favouring effects of independent factors, particularly semantics.

4.1. The semantics of the governing verb

Among the core claims in the literature on subjunctive selection is that mood choice is semantically conditioned, determined in part by the meaning of the matrix predicate. To empirically test this claim, this study follows the variationist methodology developed by Poplack (1992; 2013), which evaluates such hypotheses against actual usage patterns in discourse rather than analyst assumptions. Operationalizing semantic claims is a crucial but often difficult task for the analyst. This is due on the one hand to the lack of consensus across studies on exactly which semantic categories, meanings, and governors should be considered for a study of the subjunctive in speech, and on the other hand, to the fact that many of the meanings proposed in previous research stem from analyst introspection into the supposed beliefs, feelings, or desires of the speaker. This is problematic because, as Poplack (2013: 162) points out, there is often no objective way to reconstruct speaker intent or underlying meaning from discourse, let alone to operationalize and test it. In the present study, I follow the method implemented by Poplack (1992, 2013), which circumvents this issue to some extent by relying on observable patterns in naturalistic data to test the contribution of semantic factors to subjunctive selection.

Rather than reproducing previous classifications that often rely on traditional notions of *volition*, *emotion*, or *opinion*, this study adopts a three-way typology of

predicate classes based on Wandruszka's (1998) descriptive framework: *volitive* (10a), *dubitative* (10b), and *factive* (10c).

(10) a. VOLITIVE

Ma po' mi comprava i' pesce. Voleva che lo ***cucinassi***$_{[3SG.IMPRF.SUBJ]}$. (C.001.398)

'But then s/he would buy me fish. S/he wanted me to cook it.'

b. DUBITATIVE

Mi sembra che ***sia***$_{[3SG.PRES.IND]}$ *una felce.* (L.120.81)

'It seems to be that it is a fern.'

c. FACTIVE

Non *mi dispiace che tu* ***faccia***$_{[3SG.PRES.SUBJ]}$ *con me questo compito.* (L.10.21)

'I don't mind you doing this homework with me.'

If semantic conditioning underlies subjunctive selection, we expect consistent patterns across members (i.e., governors) of each class: the subjunctive should appear categorically with volitive predicates, variably with dubitative ones, and be disfavoured with thematic-factives. Crucially, no single lexical item should determine the behaviour of an entire category; rather, the consistency of effect across predicates within each class is the relevant diagnostic. This approach allows for a more systematic test of semantic hypotheses and avoids the pitfalls of relying on a handful of a priori selected governors, which has been a common limitation in previous research.

The results for this factor group, as presented in Table 1, show clear differences in subjunctive rates across the three categories.

Semantic Class	% Subj	N
Factive	95%	21/22
Volitive	82%	108/131
Dubitative	60%	272/450
Communicative	23%	3/13
Total	**66%**	**404/616**

Table 1: Subjunctive selection rates by semantic class of the governing verb.

A first important result is that all semantic classes exhibit variation in mood selection, albeit to differing degrees, including the volitive class (82%), "widely touted as the

wellspring of subjunctive morphology" (Poplack et al. 2018: 234). Contrary to expectations, the strongest favouring effect emerges in the factive class (95%), followed by volitive predicates (82%), while dubitative governors slightly disfavour the subjunctive (60%). A small number of tokens headed by *dire* 'to say' pattern rather with *verba dicendi*, aligning more with assertive than epistemic usage, and show a markedly low subjunctive rate (23%), as in (11).

(11) a. *Io, in terza media, c'avevo una che diceva che Leopardi* ***era***$_{[3SG.IMPRF.IND]}$ *immortale.* (C.087.55)
'Me, when I was in middle school, I had a teacher who used to say that Leopardi was immortal.'

b. *Non si può dire che nel terzo mondo* ***siano***$_{[3SG.PRES.SUBJ]}$ *tutti cattivi. (*L.412.140)
'One cannot say that in the Third World they are all bad.'

It is worth noting that volitive and dubitative classes account for the bulk of the dataset (94%, n = 581/616) and of the subjunctive forms produced in discourse (94%, n = 380/404). On the surface, one could claim that this hierarchy between volitive and dubitative governors supports Wandruszka's as well as other accounts of subjunctive selection (Loengarov 2006; Giorgi & Pianesi 1997; Schmitt Jensen 1970, among others) placing variability primarily within the epistemic domain of opinion or dubitative predicates. However, this presupposes internal uniformity: that every member of a given semantic class would pattern alike. Closer examination reveals that this is not the case. Variation persists within each class, undermining the assumption that semantic category alone can predict mood selection consistently.

Volitive Governor	Translation	% Subj	N	% Data	% Subj Morph
bisognare	to be necessary	93%	41/44	34%	38%
(non) volere	(not) to want	84%	16/19	15%	15%
sperare	to hope	93%	14/15	11%	13%
bastare	to be enough / suffice	50%	6/12	9%	6%
aspettare	to wait / to expect	50%	4/8	6%	4%
è importante	it is important	100%	4/4	3%	4%
fare sì	to ensure / to make sure	75%	3/4	3%	3%
esigere	to demand	100%	3/3	2%	3%

Volitive Governor	*Translation*	*% Subj*	*N*	*% Data*	*% Subj Morph*
avere paura	to be afraid	75%	3/4	3%	3%
è inutile	it is useless	50%	2/4	3%	2%
richiedere	to require / to demand	67%	2/3	2%	2%
non avere senso	to make no sense	100%	1/1	1%	1%
evitare	to avoid	100%	1/1	1%	1%
non preoccuparsi	to not worry	100%	1/1	1%	1%
preferire	to prefer	100%	1/1	1%	1%
sentire il timore	to feel fear	100%	1/1	1%	1%
l'importante è	the important thing is	100%	1/1	1%	1%
servire	to be needed / to be of use	100%	1/1	1%	1%
lasciare	to let / to allow	100%	1/1	1%	1%
fare in maniera	to ensure	50%	1/2	2%	1%
fare in modo	to ensure / to make sure	100%	1/1	1%	1%
Total		**82%**	**108/131**		

Table 2: Subjunctive selection across volitive predicates.

As shown in Table 2, while the overall subjunctive rate across volitive governors is 82%, closer inspection reveals considerable variability among individual verbs. Some governors strongly favour the subjunctive, such as *bisognare* 'it is necessary' and *sperare* 'to hope' (93%), while others favour it to a lesser degree, as with *volere* 'to want' (84%). A few select the subjunctive categorically (*è importante* 'it is important', *evitare* 'to avoid', *preferire* 'to prefer', etc., all at 100%), while others disfavour it or yield markedly lower rates, including *fare sì* 'to ensure' (75%) and *bastare* 'to be enough' (50%). Strikingly, just three verbs, *bisognare* 'it is necessary', *volere* 'to want', and *sperare* 'to hope', account for 60% of all volitive tokens (78/131) and 66% of the subjunctive morphology observed (71/108), with *bisognare* 'it is necessary' alone representing 34% of the data and 28% of the subjunctive forms. These patterns mirror previous variationist results (Poplack et al. 2018; Digesto 2021), underscoring the extent to which a small number of predicates can drive the overall favouring effect.

A similar trend is observed within the class of dubitative governors, which comprises the bulk of the dataset: 73% of all tokens and 67% of the total subjunctive morphology.

Dubitative Governor	Translation	% Subj	N	% Data	% Subj Morph
(non) pensare	to (not) think	73%	64/88	20%	24%
(non) sembrare	to (not) seem	75%	48/64	14%	18%
(non) credere	to (not) believe	79%	45/57	13%	17%
non è	to not be	33%	37/112	25%	14%
(non) parere	to (not) appear	58%	15/26	6%	6%
può darsi	to be possible / it might be	58%	11/19	4%	4%
(non) essere sicuro	to (not) be sure	63%	5/8	2%	2%
(non) dire	to (not) say	21%	3/14	3%	1%
avere la sensazione	to have the feeling	100%	3/3	1%	1%
non è detto	to not be certain	100%	3/3	1%	1%
ritenere	to consider	100%	3/3	1%	1%
supporre	to suppose	100%	3/3	1%	1%
avere l'impressione	to have the impression	50%	2/4	1%	1%
ci sta	to be possible	100%	2/2	0%	1%
è impossibile	to be impossible	100%	2/2	0%	1%
mettere	to assume	100%	2/2	0%	1%
non sapere	to not know	33%	2/6	1%	1%
non succedere	to not happen	100%	2/2	0%	1%
presupporre	to presuppose	67%	2/3	1%	1%
può essere	to be possible	100%	2/2	0%	1%
(non) è possibile	to be (not) possible	33%	1/3	1%	0%
assicurarsi	to make sure	100%	1/1	0%	0%
avere il dubbio	to have doubt	100%	1/1	0%	0%
calcolare	to calculate	50%	1/2	0%	0%
controllare	to check	100%	1/1	0%	0%
dedurre	to deduce	100%	1/1	0%	0%
dubitare	to doubt	50%	1/2	0%	0%
è ovvio	to be obvious	33%	1/3	1%	0%
immaginare	to imagine	25%	1/4	1%	0%
non avere il dubbio	to have no doubt	100%	1/1	0%	0%
non è vero	to not be true	50%	1/2	0%	0%
presumere	to presume	100%	1/1	0%	0%
reputare	to deem / to consider	100%	1/1	0%	0%
rischiare	to risk	100%	1/1	0%	0%
trovare	to find	50%	1/2	0%	0%
verificare	to verify	100%	1/1	0%	0%
Total		**60%**	**272/450**		

Table 3: Subjunctive selection across dubitative predicates.

Despite the overall rate of subjunctive selection being 60% (lower than the corpus-wide average of 66%), not all verbs in this category conform to a disfavouring pattern. Some governors, such as *avere la sensazione* 'to have the feeling', *supporre* 'to suppose', and *ritenere* 'to consider', show categorical selection of the subjunctive (100%), while others, such as *pensare* 'to think' (73%), *sembrare* 'it seems' (75%), and *credere* 'to believe' (79%), exhibit a clear favouring effect. Verbs like *(non) essere sicuro* '(not) to be sure' (63%) hover close to the class average, suggesting a slightly favouring effect. By contrast, others such as *parere* 'it seems' (58%) and *può darsi* 'it might be' (58%) show slight disfavour, and *non è* 'it is not' (33%) and *dire* 'to say' (21%) markedly disfavour the subjunctive. Notably, three verbs, *pensare* 'to think', *sembrare* 'it seems', and *credere* 'to believe', account for nearly half of all dubitative tokens and together represent 59% of all the subjunctive morphology attested in this class, reinforcing the influence of the lexical identity of the governor rather than a genuinely semantic effect, and suggesting that the subjunctive may be triggered by these specific verbs themselves rather than by their purported underlying meaning in discourse.

4.2. The structure of the matrix clause

Among the structural features of the matrix clause hypothesized to condition subjunctive selection is sentence type. If the subjunctive signals a weaker speaker commitment to the truth of the embedded proposition, as generally claimed (Carlier et al. 2012; Quer 2001; Wandruszka 1998; Loengarov 2006; Manzini 2000; Veland 1991), it should be favoured in non-declarative contexts, such as interrogatives and conditionals, which have been variously characterized as non-assertive or less assertive than their affirmative counterparts (Haspelmath 2003: 220; Givón 2018: 139). Negation, too, has long been associated with reduced assertiveness and is often thought to favour subjunctive morphology (Thompson 1998; Acquaviva 1996; Manzini 2000; Costantini 2011; Giannakidou 1995; Giannakidou & Mari 2015; Giorgi & Pianesi 1997; Loengarov 2006; Veland 1991). To test this hypothesis, each token in the dataset was coded according to the structure of the matrix clause, affirmative declarative, interrogative, or conditional. In addition, the data was analyzed according to polarity of the clause, affirmative versus negative. Following Poplack (2013, 2018), these factor groups were intended as independent approximations of the assertiveness of the matrix predication, separate from the semantic class of the governor.

Despite these predictions, the results do not lend support for the hypotheses. If anything, subjunctive is disfavoured in non-declarative contexts.

	% Subj	**N**
declarative	66%	390/592
non-declarative	58%	14/24
Total	**66%**	**404/616**

Table 4: Subjunctive rate by sentence type: declarative vs. non-declarative matrix clauses.

Now turning to the testing of affirmative versus negative contexts. Likewise, results do not lend support for the hypothesis that less-assertive contexts favour the subjunctive; if anything, the opposite holds, with affirmative contexts favouring it (72%) and negative ones highly disfavouring it (49%).

	% Subj	**N**
affirmative	72%	320/443
negative	49%	84/173
Total	**66%**	**404/616**

Table 5: Subjunctive rate by polarity of the matrix clause.

Not all governors appeared in both affirmative and negative contexts. If we isolate those that did, we find 20 such governors, listed in the table below with their rates of subjunctive selection.

Governor	**Translation**	**Neg. %Subj**	**Aff. %Subj**	**p-value**	**Sign.**
credere	to believe	73% (11/15)	64% (14/22)	0.7235	n.s.
dire	to say	100% (2/2)	13% (3/24)	-	-
pensare	to think	67% (6/9)	73% (58/79)	0.6999	n.s.
sembrare	to seem	83% (5/6)	74% (43/58)	1.0000	n.s.
volere	to want	75% (3/4)	87% (13/15)	-	-
parere	to appear	100% (1/1)	56% (14/25)	-	-
è possibile	it is possible	100% (1/1)	0% (0/2)	-	-
essere sicuro	to be sure	100% (1/1)	40% (2/5)	-	-
avere il dubbio	to have doubt	100% (1/1)	50% (1/2)	-	-
Sub-total		**83% (35/42)**	**64% (148/232)**		

Governor	*Translation*	*Neg. %Subj*	*Aff. %Subj*	*p-value*	*Sign.*
Only Neg Gov		***%Subj***			
non è	it is not	33% (37/112)			
non sapere	not to know	33% (2/6)			
non è detto	it is not certain	100% (3/3)			
non è vero	it is not true	50% (1/2)			
non dispiacersi	not to mind	100% (1/1)			
non è giusto	it is not fair	100% (1/1)			
non avere senso	not to make sense	100 % (1/1)			
non è da dire	it is not to be said	100% (1/1)			
non è umano	it is not humane	100% (1/1)			
non preoccuparsi	not to worry	100% (1/1)			
Sub-total		**61% (80/131)**			

Table 6: Subjunctive rate by polarity (affirmative vs. negative) for frequent governors. Fisher's exact test conducted only for governors with N ≥ 5 in both polarity contexts.

Eleven of these governors occurred only in negative form, all with very low token counts or as singletons. Only *non sapere* 'not to know' (12a below) appears with slightly more tokens, though still few, and disfavours the subjunctive, which is consistent with expectations that this verb is not typically associated with subjunctive use (Wandruszka 1998: 442; Schmitt Jensen 1970: 234). A substantial portion of the negative data, however, comes from a single governor, *non è* 'it is not' (12b), which is not only frequent but also strongly disfavours the subjunctive (33%).

(12) a. SAPERE 'TO KNOW'

Io non l'ho studiato perché <u>non sapevo</u> ***era*** *[3SG.IMPRF.IND] da fare.* (C.010.23)

'I didn't study it because I didn't know it had to be done.'

b. NON È 'IT IS NOT'

Ah <u>non è</u> che lo ***devo***[3SG.PRES.IND] *pagare io.* (L.120.81)

'Ah, it's not that I have to pay for it.'

As for the remaining negative contexts that appear to favour the subjunctive, the apparent effect is inflated by the presence of only one to three occurrences and is not interpretable in any meaningful sense.

If we examine the sub-group that showed variability in both affirmative and negative

governors, while subjunctive selection appears to be higher in negative clauses (83%) than in affirmative ones (64%) for these items, low token counts prevent firm conclusions. For those with ≥5 tokens in both contexts, we conducted Fisher's exact tests[7] to assess whether the observed differences were statistically significant. None reached significance at the $p < 0.01$ level (n.s.), suggesting that the presence of negation alone does not independently determine mood selection.

Overall, neither interrogative and conditional sentence types nor polarity support a semantically-driven account of mood selection. Results either run counter to expectations or implicate the same governors, *credere* 'to believe (13a), *pensare* 'to think' (13b), *sembrare* 'to seem' (13c), *volere* 'to want' (13d), suggesting that subjunctive selection is more closely tied to lexical identity than to the semantics of non-assertiveness.

(13) a. CREDERE 'TO BELIEVE'

Io credo che ***sia*** [3SG.PRES.SUBJ] *stata la meglio ditta di tutta l'Italia.* (C.322.151)

'I believe that it has been the best firm of all Italy.'

b. PENSARE 'TO THINK'

*Eh invece penso ora tu c'****abbia*** [3SG.PRES.SUBJ] *parecchia esperienza.* (C.5.110)

'Eh but I think now you have lots of experience.'

c. SEMBRARE 'TO SEEM'

*Ma veramente, sembra c'****abbia*** [3SG.PRES.SUBJ] *quattordici, quindic'anni.* (C.072.227)

'But honestly, it seems he is fourteen, fifteen years old.'

d. VOLERE 'TO WANT'

Senti, vuoi che ***compri*** [3SG.PRES.SUBJ] *i popcorn e la Coca-Cola?* (L.412.96)

'Listen, do you want me to buy popcorn and coke?'

Notably, earlier quantitative studies (e.g., Schneider 1999; Veland 1991), motivated by the assertion hypothesis, excluded negative and interrogative clauses on the grounds that these contexts categorically trigger the subjunctive. The present findings refute this assumption and highlight that such exclusions not only compromise the principle of accountability, but also preclude the possibility of empirically and systematically testing the very assumptions they seek to uphold.

[7] Fisher's exact test is preferable when working with small sample sizes or when expected frequencies in contingency table cells are very low (Gorman & Johnson 2013: 2019).

4.3. Other indicators of non-factual modality

A final test of semantics follows the principle of redundancy (Poplack et al. 2018), which posits that "any meaning expressed by the subjunctive would be echoed elsewhere in the discourse" (Poplack et al. 2018: 233). In the case of the subjunctive, often described as marking non-factual or uncertain propositions, one would expect its use to be reinforced by other contextual indicators of non-factuality, whether lexical, modal, or structural. These include adverbs such as *forse* 'maybe', *probabilmente* 'probably', valuative expressions like *magari* 'perhaps' or *speriamo che* 'hopefully', epistemic modals such as *può darsi* 'it might be', or metadiscursive constructions like *se non mi sbaglio* 'if I'm not wrong'. Other indicators comprise the use of matrix moods such as the conditional, future, imperative, or subjunctive, forms argued to mark speaker detachment from the assertion or to promote so-called "modal attraction" (Gatta 2002; Hooper 1975; Klein 1975; Loengarov 2006; Manzini 2000; Santulli 2009; Veland 1991; Wandruszka 1991).

According to these accounts, such elements are sometimes described as 'subjunctivizing' factors or triggers of the subjunctive, and have occasionally been treated as categorical contexts for its selection. From a variationist standpoint, however, such assumptions are tested empirically. The current study therefore codes for the presence of explicit indicators of non-factual modality and assesses whether they in fact favour subjunctive selection.

Following Poplack (2013; 2018), these indicators were operationalized as a factor group distinct from sentence type and from the lexical identity of the embedded predicate. In keeping with the principle of orthogonality (Guy 1988: 126), a context already captured in another factor group, such as conditional sentences, was excluded here to avoid redundancy and confounding effects. That is, although conditional clauses may structurally promote a non-factual interpretation and should theoretically favour the subjunctive (Manzini 2000: 243), they were already analyzed in the "sentence type" factor group (Section 4.2), and are therefore not repeated here. This approach isolates the effect of non-factual indicators not otherwise captured, allowing for a more reliable test of the semantic conditioning hypothesis.

(14) a. MOOD OF THE MATRIX CLAUSE

Ecco, a me piacerebbe che 'un ci ***fosse***$_{[3SG.IMPRF.SUBJ]}$ *questa-questo comportamento schizofrenico.* (C.430.241)

'So, I would like it not to be this-this schizophrenic behaviour.'

b. AUXILIARY USED MODALLY

Uno potrebbe pensare che architettura e struttura sono$_{[1SG.PRES.IND]}$ *la stessa cosa.* (C.451.25)

'One could think that architecture and structure are the same concept.'

c. LEXICAL

Magari ci poteva stare che io andassi$_{[1SG.IMPRF.SUBJ]}$ *a cambiarli.* (C.224.61)

'Maybe it made sense that I could go to exchange them.'

	% Subj	**N**
Presence	72%	55/76
Absence	65%	349/540
Total	**66%**	**404/616**

Table 7: Subjunctive rate according to the presence of other indicators of non-factual modality.

A first observation is that indicators of non-factual modality are relatively rare in running discourse, with only 12% of tokens enabling us to test whether their presence affects mood selection. According to our hypothesis, one would expect that any meaning conveyed by the subjunctive would be echoed by other contextual cues contributing to a non-factual reading. Indeed, results show a higher rate of subjunctive in the presence of such indicators (72%) compared to their absence (65%). However, this difference is not statistically significant at $p < 0.01$ (χ^2, $N = 616$) $= 1.44$, $p = 0.23$), suggesting that these elements do not reliably condition the selection of the subjunctive. Moreover, if we break down the category according to the specific type of indicator coded for, results show that not all indicators behave uniformly: modal auxiliaries categorically trigger the subjunctive; among the two categories that account for most of the data of indicators of non-factual modality, tense and mood of the matrix clause favour the subjunctive (77%), while the presence of lexical indicators has no effect (66%).

Type of Indicator	**% Subj**	**n Subj**
Auxiliary used modally	100%	7/7
Combination of factors	100%	2/2
Mood of the matrix clause	77%	17/22
Lexical	66%	31/47
Absence	64%	347/538
Total	**66%**	**404/616**

Table 8: Subjunctive rate according to various indicators of non-factual modality.

Summarizing, factors designed to capture a non-factual or doubtful reading through explicit cues in discourse do not consistently influence subjunctive selection, contra the hypothesis of semantic contribution. If anything, the only context where the effect is consistent is with modal auxiliaries, though low token counts preclude firm conclusions. These findings corroborate those for the other semantics-based factor groups: if semantics were genuinely explanatory of subjunctive variability, we should not observe such inconsistency both across and within factor groups. Of the three groups tested, none accounted for a semantic contribution. Instead, results suggest that a handful of governors account for 47% of the data and 56% of all subjunctive morphology: *pensare* 'to think' (73%), *sembrare* 'it seems' (75%), *credere* 'to believe' (79%), *bisognare* 'it is necessary' (93), *volere* 'to want' (84%), and *sperare* 'to hope' (93). Each favours the subjunctive, and moreover, these are the only verbs that show a higher proportion of subjunctive morphology than their share of the data, suggesting that subjunctive selection is more closely tied to lexical identity than to the semantic context (see Appendix).

This trend was observed also among dubitative governors (*pensare* 'to think', *sembrare* 'it seems', and *credere* 'to believe') and volitive predicates (*volere* 'to want', *bisognare* 'it is necessary', and *sperare* 'to hope') when we tested the semantic category of the main predicate. While the semantic classification of these predicates on the surface predicts the expected outcomes, with volitive favouring more than other classes the selection of the subjunctive, albeit not categorically so in our dataset, the high variability in terms of rates of subjunctive usage across governors and the dominant presence of just a few verbs within each semantic class suggest that the effect is less a function of semantic class per se and more attributable to these specific verbs.

5. Diachronic perspective on subjunctive selection

5.1. A glimpse into subjunctive selection in Italian: from the 16th century to contemporary speech

Previous variationist research examined the use of the Italian subjunctive across five centuries (Digesto 2019), particularly scrutinizing completive clauses to assess whether its use was productive as well as semantically motivated in earlier stages of the language. Digesto (2019) utilized written sources to examine historical stages of Italian, alongside recordings of contemporary spontaneous speech[8]. The historical benchmarks were established based on a corpus of theatrical scripts, predominantly comedies. These texts, while written, serve as proxies for spoken Italian across time, embodying features indicative of oral communication (Stefanelli 2006: 57), alongside documented vernacular features such as mood alternation under identical matrix verbs (15), exemplified by *credere* 'to believe'; the use of the adverb of negation *mica*, "characteristic of colloquial, informal, registers" (16) (Maiden & Robustelli 2007: 405); the indirect object pronoun *gli*$_{3SG.M}$ (masculine) instead of *le*$_{3SG.F}$ (feminine) with 3rd person singular feminine indirect objects (17).

(15) *Possono loro credere sul serio che la mia figliuola* ***è***$_{[3SD.PRES.IND]}$ *morta? Che io* ***sia***$_{[3SD.PRES.SUBJ]}$ *pazza? Che questa che ha con sé* ***è***$_{[3SD.PRES.IND]}$ *una seconda moglie?* (COHI.20C.098.754)
'Can they really believe that my daughter is dead? That I am crazy? That the one who is with him is his second wife?'

(16) *Non ho mica voluto io che s' arrivasse a una cosa così grave.* (COHI.20C.072.2253)
'I didn't really want to come to such a dangerous point.'

(17) *Gli*$_{[3SG.M]}$ *ho dato speranza di condurla, ancor oggi, alle voglie sue.* (COHI.18C.033.280)
'I gave her hope that I will persuade her again today.'

[8] Contemporary speech data comes from LIP (De Mauro et al. 1993) and C-ORAL-ROM (Cresti & Moneglia 2005), while the theatrical plays selected as a benchmark for the historical data were compiled as part of the COHI, *Corpus of Historical Italian* (Digesto 2019). We refer the reader to Digesto (2019; 2021) for detailed description of the corpora.

The findings (2019) indicate a relative stability in the patterns of usage of the subjunctive in Italian over time. Despite its seeming productivity, as it remains the dominant variant in completive clauses both in contemporary (66%) and written historical Italian (79% in the 16th century; 81% in the 18th century), its frequent usage is attributed to a limited number of lexical governors.

	COHI			SPEECH
	16th Century	**18th Century**	**20th Century**	**20th/21st Century**
%Subj	79%	81%	74%	66%
n Subj	299/379	408/505	207/279	404/616
n words	19,621	37,264	40,565	387,825
n governors	81	88	60	71

Table 9: Overall rate of subjunctive selection (%Subj), number of subjunctives (*n* Subj), number of governors that occurred with a subjunctive (*n* governors), and total number of words (*n* words) for each time period.

These trends hold despite the variances in the nature of the data, timespan, conversational topics, and the social backgrounds of both speakers in spontaneous interactions and characters in the comedies. It has been noted that a small subset of verbs such as *credere* 'to believe', *pensare* 'to think', as well as *sembrare* 'it seems' in contemporary speech accounts for the majority of the variation observed in discourse; that the subjunctive is likely to occur with suppletive forms of *essere* 'to be'; and that dedicated tests of semantic contribution did not substantiate the theory of a semantically-driven subjunctive, neither historically nor in present-day use. This evidence supports the premise that a lexicalization pattern has been operative since at least the 16th century.

In examining specific semantic categories (2019: 188), such as *necessity*, *volition*, and *emotion*, which are theoretically categorical triggers of subjunctive mood due to their strong modal determination, results revealed variability with indicative mood yet remaining strong predictors for the selection of the subjunctive over time. Although the effect of necessity verbs appeared as an epiphenomenon of a lexical effect due to *bisognare* 'it is necessary', *emotive* and *volitive* verbs demonstrated a more consistent effect. Despite this consistent selection, often at high rates, occasionally categorically, suggesting semantic motivation at least for these contexts, the rarity of

these governor types within the variable context and the disproportionate frequency of certain governors within the same semantic category, especially among volitive verbs, complicate the picture. Notably, within the *volitive* semantic class, the verb *volere* 'to want' singularly represented a significant proportion of usage, constituting 40% in the 16th century, 57% in the 18th century, and 73% in the 20th century, casting doubt on a purely semantic interpretation and suggesting a more lexically driven selection of the subjunctive.

5.2. Tracing complementation patterns from Vulgar Latin to Italian

Considering the aforementioned distinctions between Latin and its Romance descendants, particularly Italian, regarding 1) the preference for finite versus non-finite constructions according to the main clause predicate and 2) the shift from left- to right-branching structures, and considering the way that the variable context for subjunctive has been defined, syntactic complementation differences are likely to influence our findings. Initially, we identified subjunctive completive clauses within the Latin corpus using the methodology outlined above. To determine the continuity of subjunctive governors across languages and through time, we examined each governor identified in VL within the Italian dataset, noting the absence or presence of these governors in the Italian context, and vice versa. This approach helped to ascertain the extent of governor transmission from Latin to Italian, their ongoing association with the subjunctive mood, and the nature of transmitted conditioning.

However, the potential impact of Latin's distinct syntactic strategies on our analysis remains a concern. Specifically, we may encounter two scenarios: a) certain governors might not be identified within the variable context if they did not select a finite complement clause, hence no subjunctive usage; b) some governors in our dataset may exhibit a preference for non-finite complementation, leading to infrequent subjunctive selection (Figure 2). To better understand if a tendency for non-finite complementation in VL influenced the selection of subjunctive or indicative mood in completive clauses, we expanded the variable context for VL. This included examining every governor mentioned in Petronius's *Cena* and those found in Italian but not in VL, independently of their frequency of selecting non-finite complements. Such an approach nonetheless enables more reliable results on the use of subjunctive in VL's completive clauses and a deeper comprehension of governor persistence across languages and through time.

5.3. The overall trajectory of the subjunctive

In the VL data, 20 distinct governors were identified with an overall near-categorical rate of subjunctive selection: 97% (71/73).

Governor	Translation	%Subj	N
rogo	to ask	100%	22/22
curo	to ensure	89%	8/9
suadeo	to recommend	100%	6/6
timeo	to fear	100%	5/5
oportet	it is necessary	100%	5/5
caveo	to beware	100%	3/3
licet	it is permitted	100%	3/3
persuadeo	to persuade	100%	3/3
nolo	not to want	100%	2/2
non est	it is not	100%	2/2
spero	to hope	50%	1/2
dubito	to doubt	100%	2/2
video	to understand	100%	2/2
miror	to be astonished	100%	1/1
experior	to find out	100%	1/1
exspecto	to expect	100%	1/1
volo	to want	100%	1/1
veto	to forbid	100%	1/1
indignatus est	to be resented	100%	1/1
efficio	to make	100%	1/1
Total		**97%**	**71/73**

Table 10: Subjunctive selection with verbal governors in Vulgar Latin ('Cena Trimalchionis' from the *Satyricon* by Petronius). Governors are sorted by total number of tokens.

The indicative only appears in two tokens, one with the governor *spero* 'I hope' (18) and one with *curo* 'I care' in its negative form (19).

(18) ***Spero*** *tamen iam veterem pudorem sibi imponit*$_{[2SG.PRES.IND]}$. (VLCo.XLVII.492).
'I hope at least that my stomach imposes some decencies on itself.'

(19) *Nam quod strabonus est*$_{[3SG.PRES.IND]}$ ***non curo.*** (VLCo.XLVII.492)
'For instance, that he be crossed-eyed, I don't care.'

The initial findings highlight a notable lack of variability in the use of governors in VL. Moreover, the low number of governors contrasts with what has been found diachronically in Italian: by the 16th century, the earliest period examined in variationist research, Italian already exhibited 81 governors, highlighting a dynamic expansion. Despite this growth in Italian, the role of governors in VL regarding subjunctive selection mirrors earlier observations, emphasizing their influence over the quantity of available data. In other words, more data does not entail more subjunctives. The *Cena Trimalchionis* contributes a corpus of 12,576 words, comparable in size to the 16th-century Italian corpus (19,621 words; 81 governors), yet VL demonstrates a much smaller pool of governors (N = 20) and a limited number of occurrences for nearly all (totaling only 73). This suggests that complement clauses governed by a verb were not a highly productive context for the subjunctive in VL, unlike the trend seen in Italian.

Exploring the continuity of governors over time, the persistence of governors was examined, which enables us to establish stability and change with respect to the governors when these are coupled with their relative rates of occurrence and frequency within the variable context. Results reveal initial insights into lexical transmission. The near-total overlap of governors between VL and Italian (19 out of 20) preliminarily indicates lexical continuity. However, this figure pales in comparison to the 239 lexical types identified in Italian across the centuries (Digesto 2019: 214). Investigating these Italian governors within the Latin dataset, only one, *credo* ('I believe'), was found not to select the subjunctive in VL, indicating a potential reason for the limited number of governors in VL: they might not have constituted subjunctive-selecting contexts historically. Except for *indignatus est* 'to be resented', all other 19 governors identified in the VL data were transmitted to Italian, with a third of them (N = 6/20) persisting through all five examined periods. These include *exspecto* 'I expect', *dubito* 'I doubt', *volo* 'I want', *nolo* 'I do not want', *oportet* 'it is necessary', and *credo* 'I believe', though *credo* uniquely favoured the indicative in VL, pointing to nuanced shifts in mood selection over time.

The disparity in the number of governors and the scant occurrence of tokens in VL as compared to Italian necessitates further exploration, particularly in light of the previously discussed potential syntactic divergences in complementation. This discrepancy prompts an inquiry into whether the distinct syntax between the languages, specifically the prevalent use of infinitival clauses in VL, including those with an accusative subject (as noted by Bolkestein 1989 and Bodelot 2003), could account for the limited presence of finite complement clauses and, consequently, the sparse subjunctive data in VL. To

determine if the linguistic scenario observed with verbal governors in Latin primarily stems from profound syntactic differences, an analysis was conducted on each Latin governor's complementation type.

	Vulgar Latin		**16th Century**		**18th Century**		**20th Century**		**Actual Speech**	
	%	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**
Sharedness = 5										
nolo (≃non volere)	100	2/2	100	5/5	100	26/26	100	3/3	75	3/4
volo (≃volere)	100	1/1	100	22/22	100	47/47	100	49/49	87	13/15
oportet (≃bisognare)	100	5/5	100	1/1	100	29/29	100	9/9	92	35/38
dubito (≃dubitare)	100	1/1	100	1/1	75	3/4	67	2/3	50	1/2
exspecto (≃aspettare)	100	1/1	100	3/3	88	7/8	100	1/1	50	4/8
credo (≃credere)	0	0/6	83	39/47	98	42/43	69	22/32	79	45/57
Sharedness = 3										
rogo (≃pregare)	100	16/16	100	8/8	100	3/3				
video (≃guardare)	100	1/1	100	3/3	100	3/3				
suadeo (≃consigliare)	100	6/6	100	1/1	100	1/1				
Sharedness = 2										
exspecto (≃attendere)	100	1/1	100	1/1						
Sharedness = 3										
non est (≃non è)	100	2/2	100	1/1					33	37/112
esperior (≃scoprire)	100	1/1			100	1/1				
caveo (≃badare)	100	3/3			80	4/5	0	0/1		
rogo (≃chiedere)	100	16/16			100	1/4			66	2/3
Sharedness = 4										
miror (≃meravigliarsi)	100	1/1	100	1/1	100	4/4				
timeo (≃consigliare)	100	5/5	75	3/4	100	2/2				
curo (≃assicurare)	100	7/7			33	2/2	0	0/6		
timeo (≃avere paura)	100	5/5			100	7/7	100	1/1	75	3/4
licet (≃permettere)	100	3/3			100	6/6	100	2/2		
spero (≃sperare)	50	1/2			38	3/8	50	1/2	93	14/15
efficio (≃fare)	100	1/1	100	16/16	100	7/7			71	5/7
Sharedness = 2										
exspecto (≃attendere)	100	1/1	100	1/1						
veto (≃vietare)	100	1/1			100	1/1				

Table 11: Persistence of governors across time periods (from 2 to 5) and rate of subjunctive selection. Governors are paired according to roughly the same meaning. The columns regarding 16th, 18th and 20th century data report the results of the corpus of Italian comedies over time.

As indicated earlier (Figure 3), there is a suggestion that Latin favoured non-finite complements more so than its Romance successors (as per Magni 2009, among others). It is reasonable to hypothesise that ongoing syntactic shift from OV to VO structures in Latin, alongside a continued preference for infinitival constructions, results would manifest markedly different strategic choices compared to Italian, which is predominantly an SVO language. Should a preference for non-finite complementation obstruct the emergence of the subjunctive in VL, this would elucidate the observed low productivity in the ancestral language and provide concrete evidence of a fundamental syntactic disparity between VL and Italian.

TYPE OF COMPLEMENT:		FINITE						NON-FINITE		TOTAL	
		SUBJS		IND		AMBIGUOUS					
GOVERNOR		**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**
rogo	to ask	100%	22	0%	0	0%	0	0%	0	0	22
curo	to ensure	89%	8	11%	1	0%	0	0%	0	0	9
suadeo	to recommend	100%	6	0%	0	0%	0	0%	0	0	6
timeo	to fear	100%	5	0%	0	0%	0	0%	4	0	5
oportet	it is necessary	56%	5	0%	0	0%	0	44%	0	15	9
caveo	to beware	75%	3	0%	0	25%	1	0%	0	0	4
licet	it is permitted	43%	3	0%	0	57%	4	0%	0	0	7
persuadeo	to persuade	43%	3	0%	0	57%	4	0%	4	0	7
nolo	not to want	33%	2	33%	1	0%	0	67%	0	15	6
non est	it is not	100%	2	0%	0	0%	0	0%	0	0	2
dubito	to doubt	100%	2	0%	0	0%	0	0%	6	0	2
video	to understand	25%	2	0%	0	0%	0	75%	1	23	8
spero	to hope	33%	1	0%	0	0%	0	33%	0	4	3
miror	to be astonished	100%	1	0%	0	0%	0	0%	0	0	1
efficio	to make	100%	1	0%	0	0%	0	0%	0	0	1
indignatus est	to be resented	100%	1	0%	0	0%	0	0%	0	0	1
experior	to find out	100%	1	0%	0	0%	0	0%	0	0%	1
exspecto	to expect	100%	1	0%	0	0%	0	0%	11	0	1
volo	to want	8%	1	0%	0	0%	0	92%	0	42	12
veto	to forbid	100%	1	0%	0	0%	0	0%	0	0	1
credo	to believe	0%	0	86%	6	14%	1	0%	0	0	7
Total		**62%**	**71**	**7%**	**8**	**9%**	**10**	**23%**	**26**		**115**

Table 12: Number of tokens (N) and proportion of finite vs. non-finite complementation (%) for each verbal governor in Vulgar Latin data. Governors are sorted by total number of tokens with subjunctive morphology.

Only five out of twenty-one governors opted for infinitival clauses over finite subordinate ones, with *oportet* 'it is necessary', *nolo* 'I do not want', *volo* 'I want', *spero* 'I hope', and *video* 'to understand' showing a preference for non-finite constructions. A significant portion of these non-finite contexts is dominated by *volo* 'I want' (42%) and further increased by *nolo* 'I do not want', together accounting for 58% of cases eschewing subjunctive complements. Apart from these, the larger part of governors (16/21) exclusively chose finite complements. The difficulty in differentiating subjunctive from indicative moods due to morphological similarities in Latin might have influenced our results, suggesting an underestimation of subjunctive productivity.

Given the high degree of morphological homophony between the subjunctive and the indicative in Latin, our results had a relatively greater likelihood to be affected by such cases, preventing us from distinguishing the embedded mood and generating the apparent non-productivity observed here. However, ambiguous morphology was not a significant factor in limiting the diversity of subjunctive governors, as only three governors out of the 20 subjunctive-selecting contexts showed such ambiguity: *caveo* 'to beware', *licet* 'it is permitted', and *persuadeo* 'I persuade'. Contrary to expectations of widespread infinitival usage, our findings suggest that the choice between finite and non-finite complements does not explain the differences between VL and Italian, particularly regarding the expansion and variability of subjunctive selecting governors in Italian. The predominance of finite complements (89 tokens) over non-finite ones (26 tokens) among our identified governors suggests a preference for *finite* complementation in VL.

6. The grammaticalization path of the subjunctive

6.1. Subjunctive use from Latin to Italian

Although VL presented a limited dataset, notable patterns emerged concerning the use of the subjunctive in verb-governed completive clauses across VL and Italian. First of all, there is superficial evidence of linguistic change; specifically, the overall rate of subjunctive selection decreases from a quasi-categorical 97% in VL to 66% in contemporary Italian speech, albeit reaching 81% in the 18th century (81%).

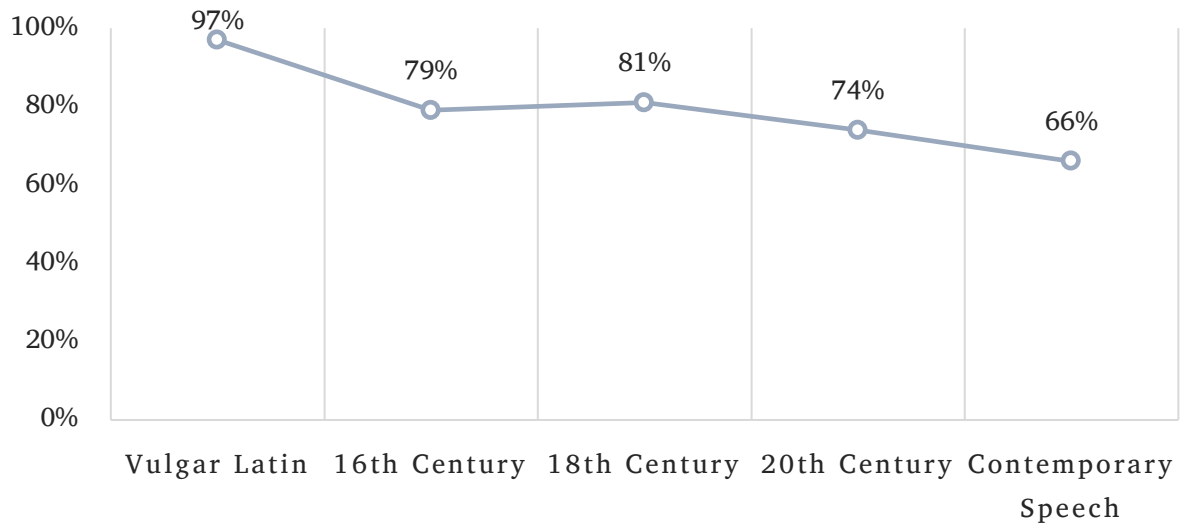


Figure 4: The trajectory of subjunctive selection across time from Vulgar Latin to 20th/21st century Italian.

The comparison between VL and Italian highlights a significant discrepancy in the number of governors, showcasing VL's relative unproductivity in this aspect, especially when juxtaposed with the more expansive governor pool in contemporary Italian. Despite analyzing a VL corpus of similar size to that of 16th-century Italian, VL exhibits a much smaller set of governors, totaling only 20 verbs, indicating a stark contrast across the languages. This observation raises questions about the dynamics of linguistic change, particularly why there is such a variance in governor numbers between VL and Italian. A deeper analysis into the application of subjunctive morphology across various completive clauses offers additional insight. The systematic coding and examination of subjunctive use within both VL and Italian contexts reveal a distinct split: in Italian, subjunctive morphology increasingly appears with verbal governors, whereas in VL, it is predominantly associated with non-verbal governors, especially conjunctions like *ut* 'so that' in VL (and its Italian counterpart *in modo che* 'so that').

(20) *Ego gloriosus volo efferri* ***ut*** *totus mihi populus bene imprecetur*$_{SBJV}$.
(VLCo.LXXVIII.1251).
'I want to be carried out in splendour, so that the whole crowd calls down blessings on me.'

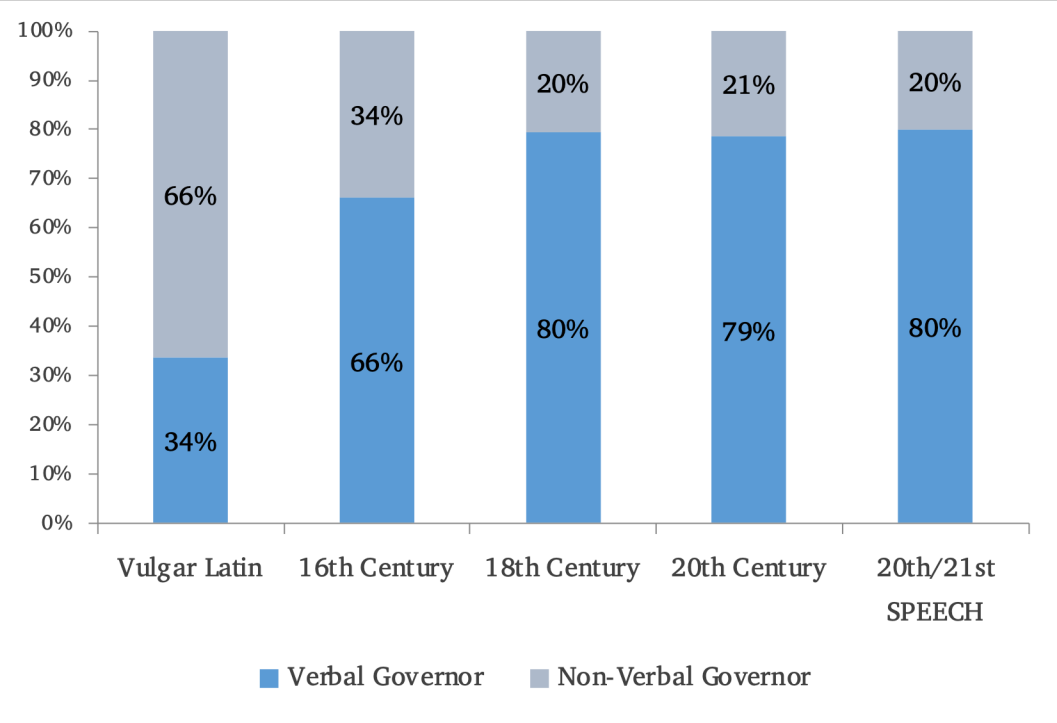
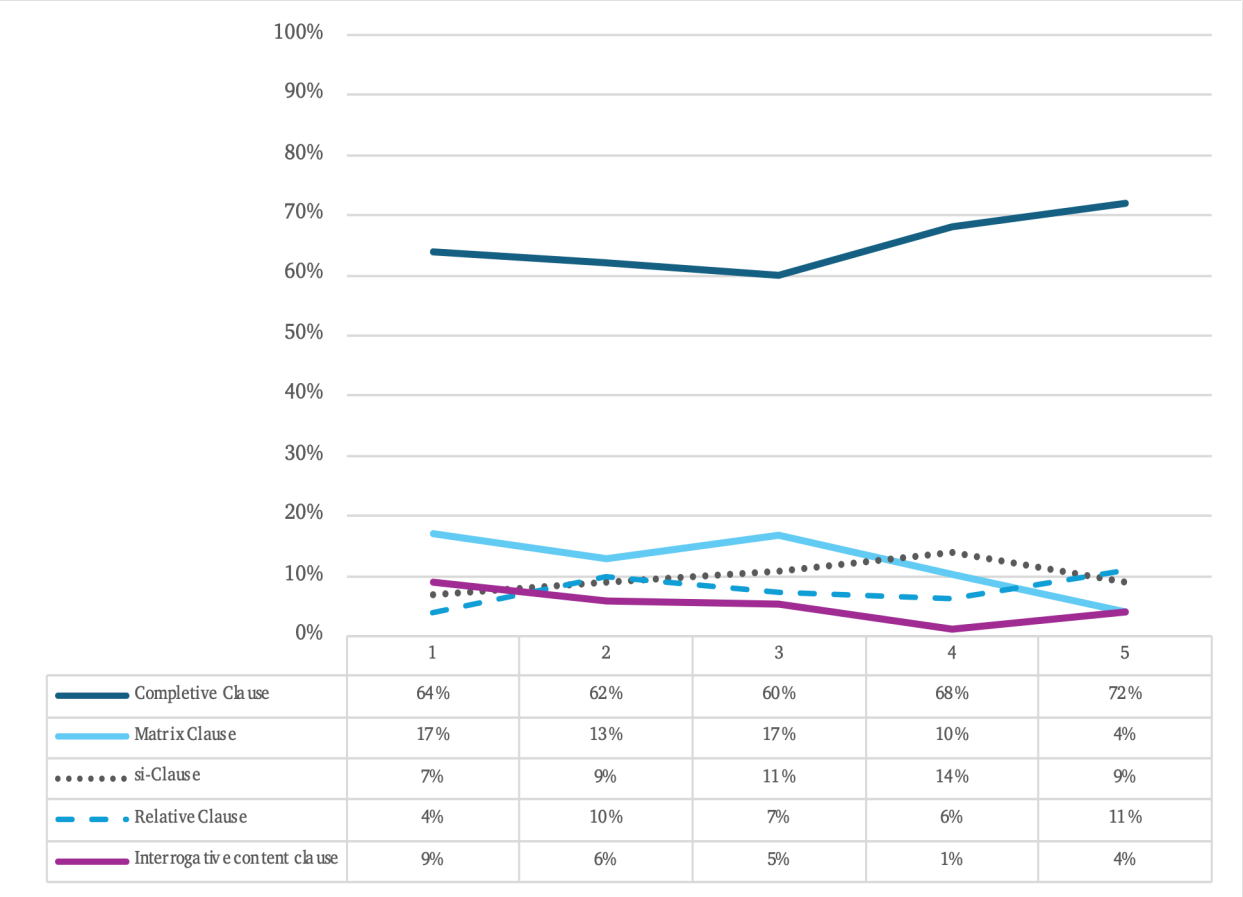


Figure 5: Distribution of subjunctive morphology in the context of completive clauses across time (from Vulgar Latin to contemporary Italian speech) when governed by either a verbal governor (object of study of the current research) or by a non-verbal governor.

This divergence underscores a fundamental shift in the usage of the subjunctive, suggesting that while in Latin, the subjunctive's association with conjunctions might have been reinforced by the syntactic transition from OV to VO, facilitating primarily finite complementation, Italian displays the opposite trend. This shift becomes evident in the 16th century, with a significant percentage (66%) of subjunctives governed by verbs, a figure that solidifies in the 18th century at 80%. Analyzing the distribution of subjunctive morphology across different contexts over time illuminates the evolving patterns of its usage.

Consistent with earlier research (Harris 1978; Noonan 2007), subjunctive morphology in VL *primarily* marked subordination, particularly within completive clauses, indicating its established role as a subordination marker. The distribution analysis reveals its limited use in independent clauses, showing a decline over time from 17% in VL to 4% in contemporary Italian. An overall decreasing trend is observed in the context of interrogative content clauses, from 9% in VL to 4% in contemporary Italian. Subjunctive use within *if*-clauses has remained relatively consistent, while relative clauses see an upward trajectory. The distribution pattern indicates a stronger association of subjunctive morphology with completive clauses in both Latin and Italian, suggesting this syntactic context as the primary domain for

subjunctive usage, a trend that has also strengthened from 64% in VL to 72% in modern Italian.



	1	2	3	4	5
Completive Clause	64%	62%	60%	68%	72%
Matrix Clause	17%	13%	17%	10%	4%
si-Clause	7%	9%	11%	14%	9%
Relative Clause	4%	10%	7%	6%	11%
Interrogative content clause	9%	6%	5%	1%	4%

Figure 6: Distribution of subjunctive morphology across syntactic environments, from Vulgar Latin to contemporary Italian speech.

Figure 5 clearly shows that the subjunctive had *already* spread out in VL to other subordinating contexts, such as relative clauses, indirect question clauses, etc. (as pointed out by Harris 1974, 1978; Magni 2009; Haudry 1973), which supports the notion that a *generalization* had already occurred by that time. Its extended use in subordinate syntactic position can be taken as evidence of the subjunctive being interpreted as a *subordinate marker* and a shift from a stronger to a weaker hypotactic link between the clauses (Magni 2009). This extended subordinating use in Latin, and the subsequent transmission of its subordinating function to Italian, along with an expanded governor pool, highlights the continuity and evolution of the subjunctive's role across languages.

6.2. Semantic factors over time

The only result landing support to a supposed semantic conditioning in Italian was related to quasi-categorical selection of subjunctive mood with verbs expressing volition, which remains stable over time. When we examine the semantic class of the VL subjunctive governors, the semantic class with the richest lexical inventory is indeed that of verbs marking volition (85% of the data).

Semantic Class	Governor	Translation	%	N
Volitive	*rogo*	to ask	100%	22/22
	curo	to ensure	89%	8/9
	suadeo	to recommend	100%	6/6
	timeo	to fear	100%	5/5
	oportet	it is necessary	100%	5/5
	caveo	to beware	100%	3/3
	licet	it is permitted	100%	3/3
	persuadeo	to persuade	100%	3/3
	nolo	not to want	100%	2/2
	spero	to hope	50%	1/2
	exspecto	to expect	100%	1/1
	volo	to want	100%	1/1
	veto	to forbid	100%	1/1
	efficio	to make	100%	1/1
	Sub-Total		**97%**	**62/64**
Dubitative	*video*	to understand	100%	2/2
	non est	it is not	100%	2/2
	dubito	to doubt	100%	2/2
	experior	to find out	100%	1/1
	Sub-Total		**100%**	**7/7**
Factive	*miror*	to be astonished	100%	1/1
	indignatus est	to be resented	100%	1/1
	Sub-Total		**100%**	**2/2**
	TOTAL		**97%**	**71/73**

Table 13: Governors and their rates of subjunctive selection according to semantic class in VL data.

Subjunctive morphology in Latin was predominantly associated with volitive contexts, confirming the notion that these verbs primarily govern the use of the

embedded subjunctive, aligning with its original paratactic function. This observation supports the idea that the syntactic patterns of harmonic hypotactic contraction observed in Latin have been preserved and transmitted to Italian, demonstrating a continuity and inheritance within the linguistic history of the languages (Labov 1989). If we assume that the contemporary Italian subjunctive kept the semantic motivation in this context, we nonetheless need to acknowledge the fact that it is *also* found in other non-harmonic contexts such as those with other types of governors, e.g., *experior* 'to find out', albeit to a lesser extent, and is also used in other embedded complements, such as interrogative-content clauses.

Thus, the subjunctive's role in Latin and its evolution in Italian can be viewed through both a semantic lens, retaining meanings from paratactic origins, and a syntactic and lexical perspective, highlighting a continuing lexical pattern but also adaptability to new lexical governors over time.

Despite the predominance of volitive governors as triggers, the subjunctive also fulfills a broader subordinating function, a dual characteristic that has been carried over into Italian. Although it had already generalized as a marker of subordination and spread out to a variety of embedded contexts, the Latin subjunctive was mainly used with volitive verbs within *that*-clauses (Figure 7).

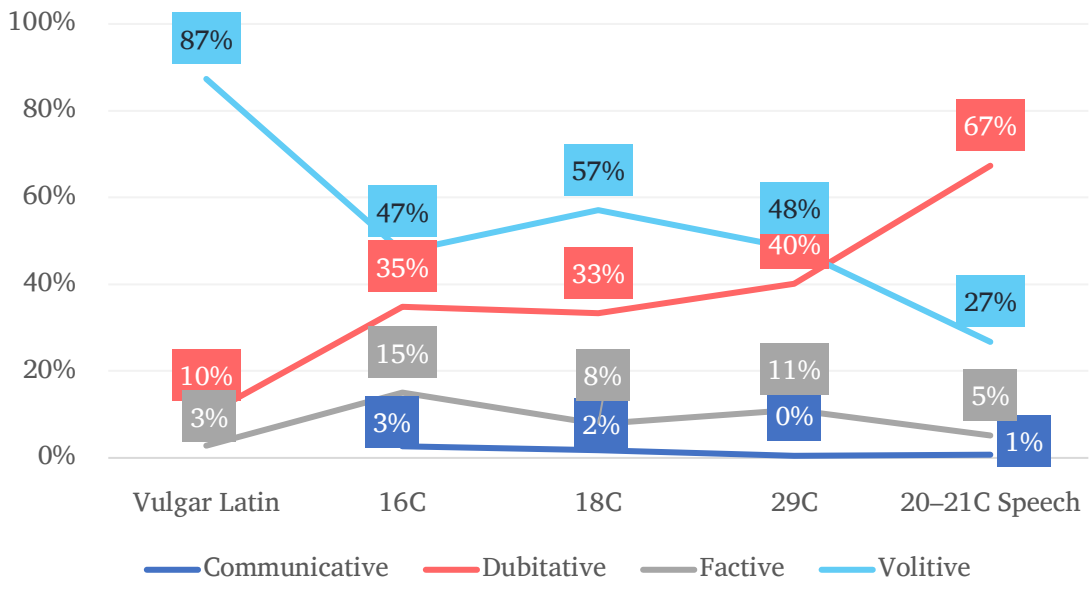


Figure 7: Distribution of subjunctive morphology across semantic classes in Vulgar Latin, historical, and contemporary Italian data.

Regarding the distinction in triggering contexts between VL and Italian, communicative and opinion verbs in VL typically lead to non-finite complement clauses, contrasting with the expanded governor pool in Italian. This would explain the important discrepancy observed with respect to the Italian situation, i.e. increased number of governors[9]. On the other hand, the strong association of the subjunctive with volitive verbs can be considered a pattern transferred from Latin into Italian, and one which is remarkably stable across centuries. We also observe an innovative trend: subjunctive morphology is gaining ground in other contexts that are non-harmonic with the original Latin's (e.g., *dubitative verbs* in the graph), providing additional support for the assumption that subjunctive use has a "relative stability so often found in Romance […] but ha[s] ceased to have any semantic import" (Harris 1974: 175).

When we examine other factor groups designed to gauge the contribution of semantics, we find that highly stable results across time with respect to the presence of other indicators of non-factual modality, which consistently favour subjunctive selection. However, as previously shown (see Digesto 2019: 197), multivariate analysis revealed that, with the exception of the 16th-century data, this factor group does not reach statistical significance, a result corroborated by the chi-square tests presented in Table 14.

Period	% Subj *presence*	% Subj *absence*	Total	χ^2 (df = 1)	p-value
16C	91% (126/138)	72% (173/241)	299/379	20.079	< 0.001 ***
18C	83% (154/186)	80% (254/319)	408/505	0.762	0.38 (n.s.)
20C	76% (51/67)	74% (156/212)	207/279	0.171	0.68 (n.s.)
20-21C Speech	72% (55/76)	65% (349/540)	404/616	1.768	0.18 (n.s.)

Table 14: Subjunctive rates according to presence of other non-factual indicators in Italian across time.

By contrast, the only finding lending support to the semantic hypothesis concerns the structure of the matrix clause: non-declarative sentences favour the subjunctive across the diachronic data and show a statistically significant effect (Digesto 2019: 197). Yet, since no such effect is observed in contemporary speech, this could suggest the erosion of semantic conditioning in present-day Italian. Still, this is an isolated result, and the broader pattern, including robust lexical effects, militates more strongly against a semantically-based analysis.

[9] It is worth noting that only a few governors account for more and more of the variation in discourse in Italian across time, namely *credere* ('to believe'), *pensare* ('to think') and *sembrare* ('to seem').

Period	% Subj *non-declarative*	% Subj *declarative*	Total	χ^2 (df = 1)	p-value
16C	96% (72/75)	75% (227/304)	227/304	16.43	5.0e-05 ***
18C	95% (138/146)	75% (270/359)	270/359	24.94	5.9e-07 ***
20C	81% (79/98)	71% (128/181)	128/181	3.25	0.071 (n.s.)
20-21C Speech	58% (14/24)	66% (390/592)	390/592	0.58	0.446 (n.s.)

Table 15: Subjunctive rates according to declarative and non-declarative structure of the matrix clause in Italian across time.

Likewise, the effect of polarity on subjunctive selection is consistent over time, with negative polarity favouring the selection of the subjunctive in diachronic data. The hierarchy is reversed in contemporary data.

Period	% Subj Aff.	% Subj Neg.	Total	χ^2 (df = 1)	p-value
16C	77% (260/339)	98% (39/40)	299/379	9.30	< 0.01 **
18C	78% (339/432)	95% (69/73)	408/505	10.36	< 0.01 **
20C	74% (183/246)	73% (24/33)	207/279	0.04	0.84 (n.s.)
20-21C Speech	72% (320/443)	77% (47/61)	367/504	0.63	0.43 (n.s.)

Table 16: Subjunctive rates according to affirmative and negative polarity across time. *For the 20–21C Speech data, the rate for negative polarity excludes tokens with the lexical governor *non è* 'it is not', which categorically disfavours the subjunctive and skews the overall distribution.

However, as shown in section 4.2 above, *non è* 'it is not' alone was responsible for lowering the overall rate of negative polarity to 49%; if we exclude the 37/112 tokens of *non è* 'it is not', the rate rises to 77% (47/61), patterning exactly like previous stages of the language. Nonetheless, recall that close inspection showed that where data was available for analysis, no significant difference was observed. In earlier periods (16C, 18C), negative contexts clearly favour the subjunctive, and the difference is statistically significant ($p < 0.01$). In the 20C dataset, however, the effect disappears entirely, and in contemporary speech (20–21C), the apparent disfavouring effect is driven almost exclusively by a single lexical item (*non è* 'it is not'), which categorically disfavours the subjunctive. Once these tokens are removed, no significant effect of polarity remains at $p < 0.01$, as shown in table 16 above. This bolsters previous observations that polarity is not an independent conditioning factor in mood selection, but rather reflects the influence of specific lexical governors.

6.3. Unveiling the path to lexicalization: insights across Romance languages

The variationist method has been employed to study subjunctive selection in completive clauses across Romance languages, including French, Italian, and Portuguese. This approach, building on the methodology developed by Poplack for studying the use of French subjunctive (1992) and leveraged in this study, has been applied to contemporary spoken Romance language corpora, facilitating a synchronic comparison (Poplack et al. 2018). Drawing from the findings of Poplack and her associates (2018), we note that subjunctive-selecting governors in the VL benchmark predominantly use the subjunctive, contrasting with the robust variability observed in all the descendant languages. Moreover, some governors select the subjunctive frequently, while others do so less often : "just a few of them make up a disproportionately large part of the entire governor pool, and the rest are very rare." (Poplack et al. 2018: 239). Although a wide range of governors can trigger the subjunctive mood, only a select few contribute *significantly* to the observed variation in discourse.

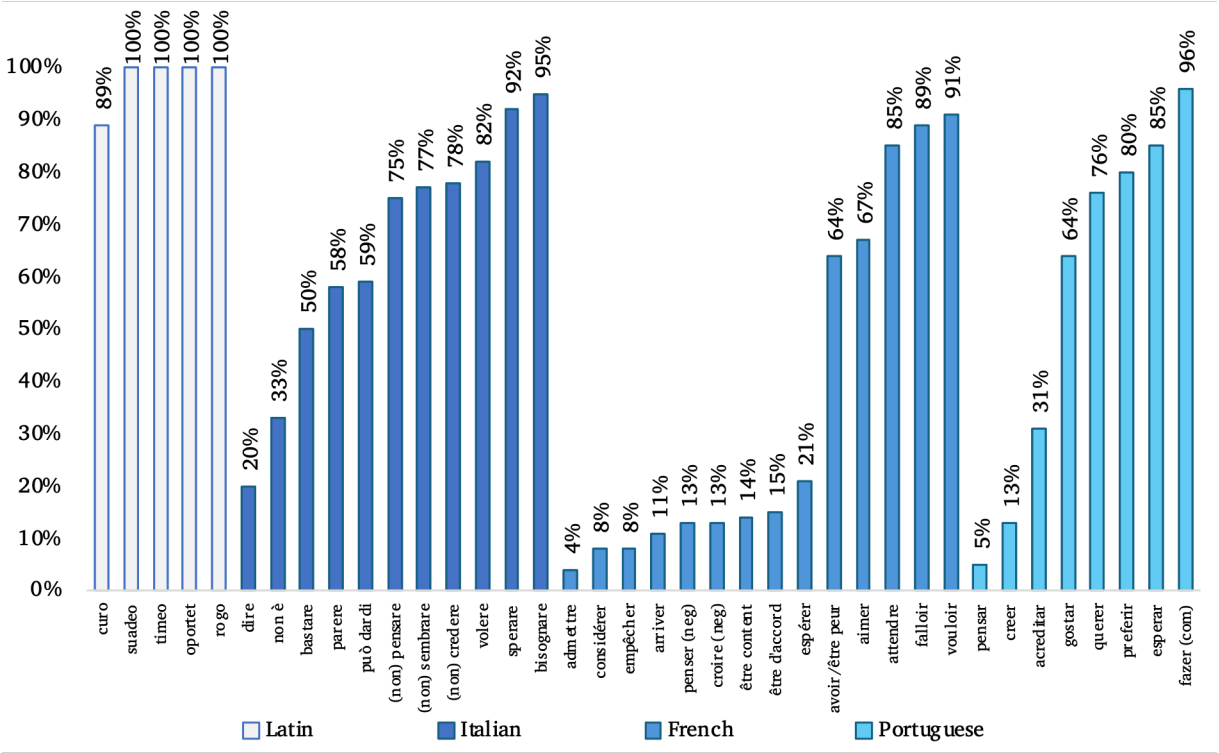


Figure 8: Subjunctive rate across frequent governors by language, adapted from Poplack et al. (2018: 239).

In a study of a 2 million-word spoken French corpus from Ottawa-Hull, Canada (Poplack 1989), *falloir* 'it is necessary' and *vouloir* 'to want' were found to represent 89% of all subjunctive morphology in complement clauses. In Brazilian Portuguese, *querer* 'to want' accounts for approximately a quarter of the governors and about 30% of the subjunctive morphology. Italian exhibits greater dispersion across a broader array of governors, with *pensare* 'to think', and *credere* 'to believe', *sembrare* 'it seems', *bisognare* 'it is necessary' and *volere* 'to want' accounting for 44% of the governor pool and more than half of the subjunctive morphology in discourse.

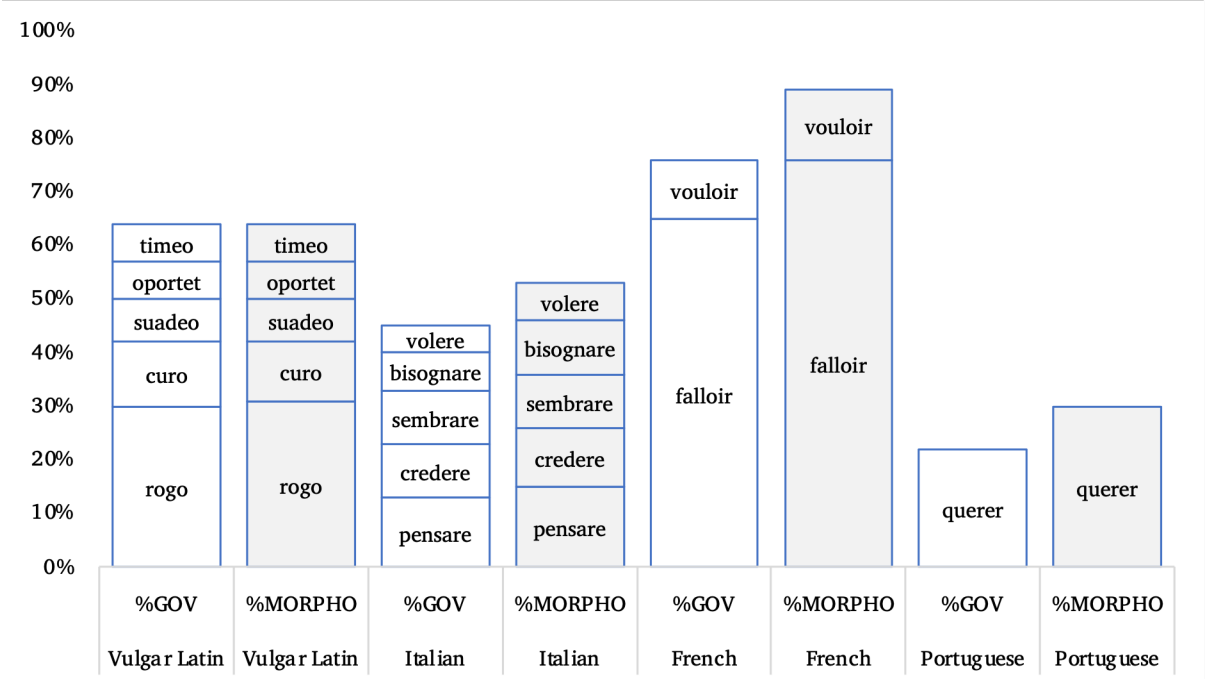


Figure 9: Dispersion of subjunctive morphology across governors by language, as measured by proportion frequent governors represent of the governor pool (%GOV) and proportion of subjunctive morphology they account for (%MORPHO), adapted from Poplack et al. (2018: 241).

Considering the semantic classes of governors and the identities of these in VL, there is evident transmission to the daughter languages, especially regarding volitive governors. This is exemplified by *falloir* 'it is necessary' (rate of subjunctive selection, 89%) and *vouloir* 'to want' (91%) in French, *querer* 'to want' (76%) in Portuguese, *bisognare* 'it is necessary' (100%) and *volere* 'to want' (92%) in Italian, indicating a preservation of these contexts within the Romance languages.

Additionally, Poplack and her associates (2018) highlighted that despite the eligibility of virtually all verbs for subjunctive morphology, only a limited number of

verbs actually carry it in discourse. This selective usage, responsible for a significant portion of subjunctive morphology in Portuguese, more than a third in both languages, and even more so in French (almost 70%) and Italian (where irregular forms of *essere* 'to be', mainly *sia*$_{[be.1\text{-}2\text{-}3SG.PRES.SUBJ]}$ contribute significantly to the selection of the subjunctive; Digesto 2019: 142), suggests a lexically-driven and non-productive use of the subjunctive in these languages.

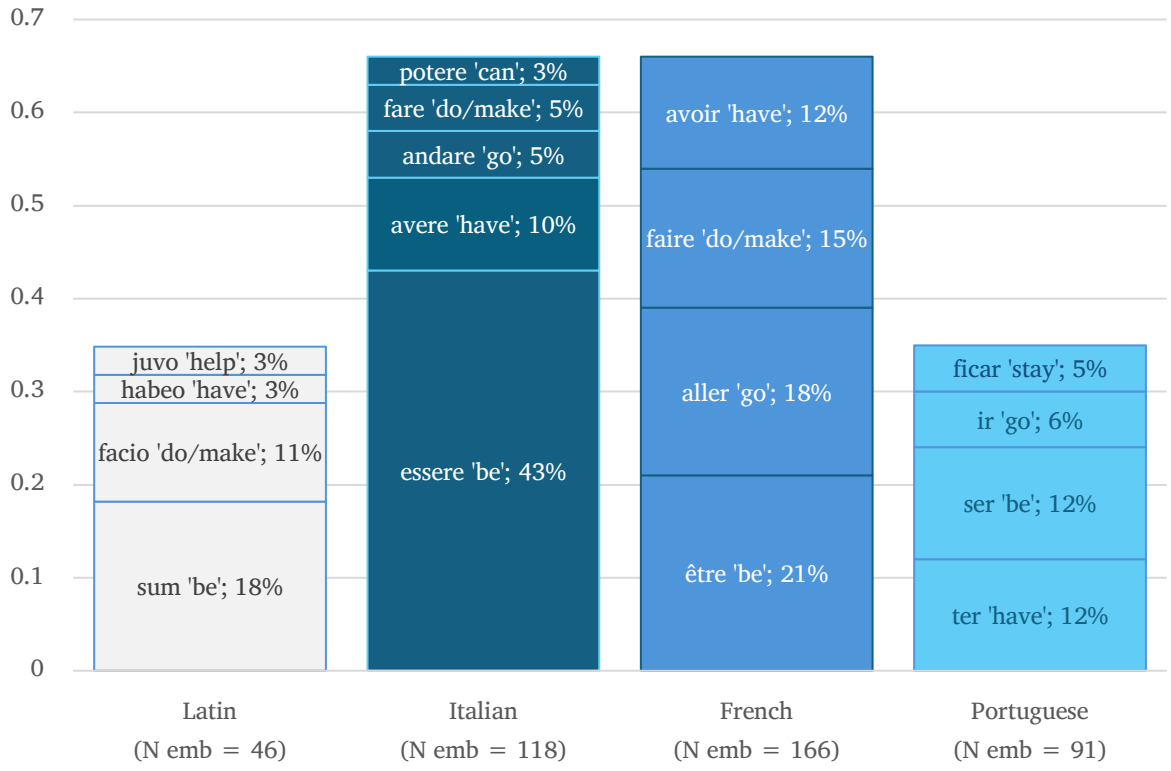


Figure 10: Distribution of subjunctive morphology across embedded verbs, adapted from Poplack et al. (2018: 243).

7. Discussion

This study contributes to our understanding of mood selection in completive clauses across Romance, offering an insight into the application of the variation method to examine typological relevant linguistic issues. We examined whether the lexicalized pattern of the subjunctive in contemporary Romance reflects a loss of semantic conditioning or the persistence of an inherited structure from Vulgar Latin. By analyzing VL data using established empirical methods, we traced these patterns to their origins and enabled a cross-linguistic comparison of subjunctive use in contexts

where these patterns remain operative. Despite the limitations of the data and the challenge of representing historical speech accurately, our findings reveal remarkable parallels between VL and Italian, extending to the broader Romance language domain, emphasizing a high "degrees of Romanceness" (Portner 1998: 38).

Our examination of the subjunctive in VL, particularly within Trimalchio's dinner party from the Satyricon, demonstrated a near-categorical selection of this mood in complement clauses, limited to a few governors. This observed pattern wasn't due to the VL corpus's size; a comparable 16th-century Italian corpus offered a greater diversity of governors and subjunctive governors. Furthermore, previous variationist research provided similar insights: the Canadian French corpus counts more than 2 million words; nevertheless, the corpus yielded an even smaller governor pool (n = 37!). This is related to the fact that the subjunctive is context-related, primarily influenced by the governor's lexical identity. However, almost all the VL governors belong to the semantic classes of volitive verbs. Surprisingly, the rate of subjunctive selection according to these semantic classes appeared fairly stable across time; governors within them highly favour or categorically favour the use of the subjunctive. While rates of subjunctive selection although decrease overtime (see *volere* 'to want' in both its affirmative and negative counterparts in VL, *volo* 'I want' and *nolo* 'I do not want'), these remain strong predictors of subjunctive selection, nonetheless accounting for less and less variation in discourse.

The prototypical use of the subjunctive in completive clauses could be considered 'harmonic' with the pre-existing paratactic constructions of Latin. This raises a critical question: is the previous paratactic (independent clause) meaning of the subjunctive still retained in the new hypotactic constructions, or is it just a syntactic mechanism devoid of any meaning and therefore reflects the grammaticalization of the Latin subjunctive as a subordinating marker? One perspective posits that the subjunctive was semantically motivated in VL, and given the transmission of its contexts of use to Romance languages (volitive verbs), we could consequently assume that these contexts have retained their original semantic contribution. Conversely, another interpretation challenges this semantic assumption, suggesting that the Latin subjunctive was already a morphosyntactic rather than a morphosemantic device, and that these morphosyntactic characteristics had been transferred into the daughter languages. The evidence at our disposal suggests a lack of lexical productivity in these so-called "semantic" contexts and a handful of governors responsible for the overall apparent semantic effect of these classes. Moreover, the surprising continuity of the

Latin pattern in Italian over the centuries, as well as French and Portuguese, weakens the semantic hypothesis due to the fact that these categories are both sparsely populated with governors, and infrequent in the data overall (see also Poplack et al. 2018). Evidence collectively points to a lexicalized nature of the use of the subjunctive in discourse. Furthermore, we can assume that the use of the subjunctive in VL was already "automatic", as previously suggested by Harris (1974), as evidenced by the categorical association of the subjunctive with a very small set of verbal governors, such as *rogo* 'to ask/to want', *oportet* 'it is necessary', among others.

In support of the choice of ruling out the question of the semantic role of volitive governors, there is also evidence that the overwhelming majority of the variation observed in Italian–both historically and in contemporary data–can be accounted for by the clear lexicalization of the subjunctive in discourse. As reported above, this stems from the fact that most of the variation in discourse is accounted by a handful of governors and by suppletive forms of *essere* 'to be'.

Based on the overall rates of subjunctive selection, the trajectory from VL to contemporary Romance languages shows a decline in the overall rate.[10] Despite the considerable temporal gap difference between the VL and the earlier Romance benchmark (16th century Italian), as well as the evolution within the Romance family, the analysis uncovered a widespread, consistent pattern across these languages. The subjunctive, though it has adapted to the lexical and syntactic characteristics of each language, maintains a core set of governing principles derived from VL. These play a crucial role in linguistic variation and account for most of the subjunctive morphology used in the targeted Romance varieties, notably the effect of *to want* across all three Romance languages, and *it is necessary* in Italian and more particularly in French. Nevertheless, an innovative trend has emerged: Italian, diverging from its Latin roots, has increasingly employed verbs of thinking such as *credere* 'to believe' and *pensare* 'to think' as subjunctive triggers. This development, where subjunctive forms are found in contexts not aligned with original Latin usage, is mirrored in other Romance languages, although they are quite rare in the variable context and, particularly in French Canadian speech, they highly disfavour the use of the subjunctive. Previous research into the dynamics of the lexical contribution in contemporary Italian subjunctive has identified a significant number of governors within what is classified

[10] In addition to the overall rate of subjunctive in contemporary Italian discourse reported in Figure 4 (66%), the use of the subjunctive, while quite robust in some Romance languages compared to others, shows a decline in French to 76% and more significantly in Portuguese to 55% (Poplack et al. 2018: 229).

as the semantic class of opinion verbs. Yet, the majority of variation and subjunctive use attributed to opinion verbs is concentrated amongst only a few governors, such as *credere* 'to believe', *pensare* 'to think', *sembrare* 'to seem'. Additional analyses investigating the supposed semantic contribution, examining the distinction between affirmative and negative constructions of verbs such as *credere* 'to believe' and *pensare* 'to think', along with the identification of contextual cues suggesting uncertainty or doubt, have not demonstrated a significant divergence in the choice between subjunctive and indicative forms under opinion verbs.

We observed a diminishing effect and a significant reduction in the distribution of the subjunctive with verbs of volition, which are the principal contexts of its use in VL data. Despite this, the daughter languages appear to still exhibit a perceived "semantic" impact of volitive verbs on subjunctive selection, though this effect appears to be actually an epiphenomenon resulting from the lexical identities characterizing these so-called semantic classes (see *querer* 'to want', *falloir* 'to be necessary', *vouloir* 'to want', *volere* 'to want', *bisognare* 'to be necessary'). The findings related to the contribution of a limited number of embedded verbs, which are predominantly responsible for subjunctive selection in Romance languages, further challenge the notion of a semantic contribution. Instead, they point to a lexicalized pattern of subjunctive use.

These findings lend support to the view that the idea of the subjunctive as semantically meaningful is a relatively modern development. For much of its documented history, the subjunctive was regarded primarily as a formal marker of subordination (*subiunctivus*, as its name underscores its hypotactic status) rather than as a carrier of inherent meaning. Metalinguistic analyses of normative grammars across centuries, both in Italian (Digesto 2019) and in French (Poplack & Dion 2013), similarly indicate that concerns about the meaning of the subjunctive emerge particularly in the modern period, coinciding with the rise of contemporary linguistic inquiry. Since the Latin era, the subjunctive was seen as morphosyntactically functional but semantically empty: *per se non exprimat sensum* ('by itself it does not express meaning,' Diomede, Latin grammarian; Keil ed. [1857: 340]).

8. Conclusion: bridging the past and future

The quantitative approach helped in elucidating grammatical (dis)similarities between languages beyond merely identifying the presence or absence of a form,

especially when syntactic patterns are expected to favour one form over another. Regarding the expected dissimilar syntactic outcomes between the ancestor language and Italian in particular, the subjunctive mood has broadened its subordinating role, a function that was already well-established by the 16th century in Italian, as exemplified by an increasing number of governors over time, even though a select few account for the majority of the variation in discourse.[11]

Furthermore, the shift from OV to VO structures in VL seems not to have impacted the lexical governors that trigger the subjunctive in the ancestor language. In fact, the findings indicate that the predominant strategy among governors is finite complementation and that the subjunctive mood has consistently served as the mood of subordination, primarily within completive clauses, a role that has remained relatively unchanged from VL to Italian. However, it's noteworthy that complement clauses featuring the subjunctive were initially triggered mainly by non-verbal governors, especially with *ut*, a pattern that has evolved in the descendant languages (see Figure 4: we observe increasing use of the subjunctive with verbal governors as opposed to non-verbal ones).

These results became apparent from a systematic and empirical examination of the subjunctive across languages and over time, adopting a corpus-based perspective to shed light on the processes governing its selection. Focusing on actual speech or close facsimiles of speech (such as dialogues written for stage plays) for historical benchmarks also allows for an exploration of inherent variability and its structured heterogeneity, and helps mitigate the potential influence of normative injunctions.

The systematic and exhaustive analysis of all instances where the subjunctive appears in the corpus, factoring in lexical, semantic, and syntactic factors, both in isolation and in combination, allows us to advance beyond mere overall occurrence rates. Relying solely on these rates, even within a defined semantic class, can obscure underlying patterns that aren't immediately apparent. Conversely, by

[11] It is worth noting that Digesto (2019; 2022) showed that the subjunctive is increasingly associated with suppletive forms of *essere* 'to be' and the complementizer *che* 'that' over time. Furthermore, in contemporary speech data, it has gained social prestige, evolving into a sociolinguistic marker: speakers with higher levels of education predominantly use the subjunctive within the variable context, employing a diverse array of governors alongside 'that' and suppletive forms of *essere*. However, no significant differences are observed in the use of the subjunctive with a core set of governors, such as *credere* 'to believe', *pensare* 'to think', *sembrare* 'to seem', *bisognare* 'to be necessary', and *volere* 'to want', across varying levels of speaker's education or speech styles.

clearly delineating a variable context, as well as by empirically operationalizing hypotheses about the contributions of linguistic (as well as extralinguistic) factors, we can discern cross-linguistic differences and similarities through quantifiable patterns of linguistic variation.

Abbreviations

1 = first person	F = feminine	PL = plural
2 = second person	IMPRF = imperfect tense	PRES = present tense
3 = third person	IND = indicative mood	SG = singular
ACT = active voice	M = masculine	SBJV = subjunctive mood

References

Acquaviva, Paolo. 1996. The Logical Form of Negative Concord. *Working Papers in Linguistics*, 1–29. Università Ca' Foscari di Venezia.

Ayres-Bennett, Wendy. 2000. Voices from the past. Sources of Seventeenth-Century Spoken French. *Romanische Forschungen,* 112(3). 323–348.

Berlinck, Rosane. 2019. Subjuntivo vs indicativo em orações completivas: percurso diacrônico no português brasileiro. In Ernestina Carrilho, Ana Maria Martins, Sandra Pereira & João Paulo Silvestre (eds.), E*studos linguísticos e filológicos oferecidos a Ivo Castro*, 217–245. 1st ed. Lisboa: Centro de Linguística da Universidade de Lisboa.

Binazzi, Neri. 2015. La frequente rinuncia al "che" nel parlato fiorentino: caratteristiche del fenomeno e spunti di riflessione per la lingua comune. *Studi di grammatica italiana*, vol. 33, 255–294. Le Lettere. Firenze: Accademia della Crusca.

Bodelot, Colette. 2003. *Grammaire fondamentale du latin: Les propositions complétives en latin.* Tome X (Bibliothèque d'études Classiques). Louvain: Peeters.

Bolkestein, A. Machtelt. 1989. Parameters in the expression of embedded predications in Latin. In Gualtiero Calboli (ed.), *Subordination and other Topics in Latin: Proceedings of the Third Colloquium on Latin Linguistics*, Bologna, 1-5 April 1985, 3–36. Amsterdam/Philadelphia: John Benjamins Publishing.

Bonomi, Ilaria. 1993. I giornali e l'italiano dell'uso medio. *Studi di grammatica italiana,* 15. 181-201.

Bybee, Joan L. 2003. Mechanisms of change in grammaticization: The role of frequency. In Brian Joseph & Richard Janda (eds.), *The Handbook of Historical Linguistics*, 602–623. London: Blackwell.

Bybee, Joan L., Revere Perkins & William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World.* Chicago: University of Chicago Press.

Campbell, Lyle. 2013. *Historical Linguistics: An Introduction*. 3rd ed. Cambridge/Massachusetts: The MIT Press.

Carlier, Anne, Walter De Mulder & Béatrice Lamiroy. 2012. Introduction: The pace of grammaticalization in a typological perspective. *Folia Linguistica,* 46(2). 287–302. (doi:10.1515/flin.2012.010)

Chinellato, Paolo. 2001. L'interpretazione morfosemantica del modo congiuntivo in italiano e in tedesco. *Rivista di grammatica generativa* 26. 3–20.

Clackson, James. 2011. *A Companion to the Latin Language.* Chichester, UK: John Wiley & Sons.

Costantini, Francesco. 2011. Subjunctive obviation in nonargument clauses. *Working Papers in Linguistics,* 39–61. Università Ca'Foscari di Venezia.

Cressot, Marcel. 1947. *Le style et ses techniques.* Paris: Presses Universitaires de France.

Cresti, Emanuela & Massimo Moneglia. 2005. *C-ORAL-ROM integrated reference corpora for spoken Romance languages.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

(http://public.eblib.com/choice/publicfullrecord.aspx?p=622285) (Accessed May 30, 2016.)

De Mauro, Tullio, Federico Mancini, Massimo Vedovelli & Miriam Voghera. 1993. *Lessico di frequenza dell'italiano parlato.* Milano: Etas.

Digesto, Salvio. 2019. Verum a fontibus haurire. *A Variationist Analysis of Subjunctive Variability Across Space and Time: from Contemporary Italian back to Latin.* Ottawa, Canada: University of Ottawa.

Digesto, Salvio. 2021. Lexicalization and Social Meaning of the Italian Subjunctive. *Cadernos de Linguística*, 2(3). e609. (doi:10.25189/2675-4916.2021.V2.N3.ID609)

Disterheft, Dorothy & Carlotta Viti. 2010. Subordination. In Silvia Luraghi & Vit Bubenik (eds.), *Continuum companion to historical linguistics* (Bloomsbury Companions), 230–249. London: Continuum.

Dorigo, Carlo Alberto. 1951. Problemi di metodo e prospettive nuove negli studi della sintassi italiana. *Aevum. Rassegna di scienze storiche, linguistiche e filologiche*, XXV(4). 355–361.

Durante, Marcello. 1981. *Dal latino all'italiano moderno: saggio di storia linguistica e culturale.* Bologna: Zanichelli.

Elsig, Martin & Shana Poplack. 2006. Transplanted dialects and language change: question formation in Québec. *Penn Working Papers in Linguistics*, 12(2), Selected Papers from NWAV 34. 77–90.

Farkas, Donka F. 1992. On the semantics of subjunctive complements. In Paul Hirschbühler & Ernst Friedrich K. Koerner (Eds.), *Romance Languages and Modern Linguistic Theory* (Current Issues in Linguistic Theory, vol. 91), 69–104. Amsterdam: John Benjamins.

Gatta, Francesca. 2002. Osservazioni sul congiuntivo in margine ad una giornata televisiva. In Leo Schena, Michele Prandi & Marco Mazzoleni (eds.), *Intorno al congiuntivo, Atti del Convegno Nazionale "Attorno al congiuntivo. Storia, tipologia, traduzione" (Forlì, 2 – 3 marzo 2000)*, 83–92. CLUEB. Bologna: CLUEB.

Giannakidou, Anastasia. 1995. Subjunctive, habituality and negative polarity. *Semantics and Linguistic Theory* (SALT), 5. 132–150.

Giannakidou, Anastasia & Alda Mari. 2015. Mixed (Non)Veridicality and Mood Choice in Complement Clauses. *CLS* 51. Chicago, France.

Giorgi, Alessandra & Fabio Pianesi. 1997. *Tense and Aspect: from Semantics to Morphosyntax (Oxford Studies in Comparative Syntax)*. New York: Oxford University Press.

Givón, Talmy. 2018. *Negation in language: Between semantics and pragmatics.* On Understanding Grammar: Revised Edition. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Gorman, Kyle & Daniel Ezra Johnson. 2013. 214 Quantitative Analysis. In Robert Bayley, Richard Cameron, & Ceil Lucas (eds.), *The Oxford Handbook of Sociolinguistics,* 214–240. Oxford: Oxford University Press.

Guy, Gregory R. 1988. Advanced VARBRUL analysis. In Kathleen Ferrara, Becky Brown, Keith Walters & John Baugh (eds.), Linguistic Contact and Change: Proceedings of the Sixteenth Annual Conference on New Ways of Analyzing Variation, 124–136. Austin, TX: University of Texas Department of Linguistics.

Haiman, John. 1994. Ritualization and the development of language. In William Pagliuca (ed.), *Perspective on Grammaticalization*, 3–28. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Handford, Stanley A. 1947. *The Latin Subjunctive. Its Usage And Development From Plautus To Tacitus.* London: Methuen.

Harris, Martin. 1974. The subjunctive mood as a changing category in Romance. In John M. Anderson & Charles Jones (eds.), *Historical Linguistics II*, 169–188. Amsterdam: North-Holland Publishing Company.

Harris, Martin. 1978. *The evolution of French syntax: a comparative approach* (Longman Linguistics Library 22). London/New York: Longman.

Harris, Martin & Nigel Vincent. 1997. *The Romance Languages*. London: Routledge.

Haspelmath, Martin. 2003. The geometry of grammatical meaning: semantic maps and cross-linguistic comparison. In Michael Tomasello (ed.), *The new psychology of language: cognitive and functional approaches to language structure,* vol. 2, 211–242. Mahwah, New Jersey: Lawrence Erlbaum.

Haudry, Jean. 1973. Parataxe, hypotaxe et correlation dans la phrase latine. *Bulletin de la Société de Linguistique de Paris*, 68. 147–186.

Herman, József. 1989. Accusativus cum Infinitivo et subordonnée à quod, quia en latin tardif: Nouvelles remarques sur un vieux problème. In Gualtiero Calboli (ed.), *Subordination and Other Topics in Latin: Proceedings of the Third Colloquium on Latin Linguistics*, Bologna, 1-5 April 1985, 133–152. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Hooper, Joan. 1975. On assertive predicates. In John Kimball (ed.), *Syntax and semantics*, vol. 4, 91–124. New York: Academic Press.

Jespersen, Otto. 1924. *The philosophy of grammar.* London: Allen & Unwin.

Kastronic, Laura. 2016. *A Comparative Variationist Approach to Morphosyntactic Variation in Contemporary Hexagonal and Quebec French.* Ottawa, Canada: University of Ottawa.

Keil, Heinrich. 1857. *Flavii Sospatri Charisii Artis grammaticae libri v. Diomedis Artis grammaticae libri III. Ex Charisii Arte grammatica excerpta*. Lipsia: In aedibvs B.G. Teubneri.

Klein, Flora. 1975. Pragmatic constraints on distribution: the Spanish subjunctive. *Papers from the regional meeting of the Chicago Linguistic Society* XI. 353–365.

Labov, William. 1994. *Principles of Linguistic Change, Internal Factors. Language in Society*. Vol. 1: Internal Factors. Cambridge and Oxford: Blackwell Publishers.

Labov, William. 1972. *Sociolinguistic Patterns.* Philadelphia: University of Pennsylvania Press.

Labov, William. 1984. Field Methods of the Project on Linguistic Change and Variation. In John Baugh & Joel Sherzer (eds.), *Language in Use: Readings in Sociolinguistics*, 28–54. Englewood Cliffs, NJ: Pretence-Hall. (Accessed July 17, 2017.)

Labov, William. 1989. The child as linguistic historian. *Language Variation and Change*, 1(1). 85–97.

Labov, William. 2001. *Principles of Linguistic Change, Social Factors. Language in Society*. Vol. 2: Social Factors. Malden and Oxford: Blackwell Publishers.

Lehmann, Christian. 1995. *Thoughts on Grammaticalization*. 2nd revised. Munich: Lincom.

Leone, Alfonso. 1949. *L'uso del congiuntivo in Latino. Principi di sintassi ragionata.* Catania: tip. Coniglione & Giuffrida.

Lindschouw, Jan. 2011. *Étude des modes dans le système concessif en français du 16e au 20e siècle et en espagnol moderne : évolution, assertion et grammaticalisation.* Copenhagen: Museum Tusculanum Press.

Loengarov, Alexander. 2006. L'alternance indicatif/subjonctif dans les langues romanes. Motivation sémantico-pragmatique et grammaticalisation. Doctoral dissertation, University of Leuven (KUL).

Lombardi Vallauri, Edoardo. 2003. Pragmaticizzazione dell'incompletezza sintattica nell'italiano parlato: le ipotetiche sospese. In *Il parlato italiano*.

Magni, Elisabetta. 2009. Mood and Modality. In Philip Baldi & Pierluigi Cuzzolin (eds.), *New Perspectives on Historical Latin Syntax. Constituent Syntax: Adverbial Phrases, Adverbs, Mood, Tense*, Vol. 2, 193–275. Berlin/New York: Mouton de Gruyter.

Maiden, Martin & Cecilia Robustelli. 2007. *A Reference Grammar of Modern Italian* (Routledge Reference Grammars). 2nd edn. London/New York: Routledge.

Manzini, Maria Rita. 2000. Sentential Complementation: The subjunctive. In Peter Coopmans, Martin Everaert & Jane Grimshaw (eds.), *Lexical Specification and Insertion,* 241–267. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Mellet, Sylvie. 1994. Le subjonctif. In Sylvie Mellet, Marie-Dominique Joffre & Guy Serbat (eds.), *Grammaire fondamentale du latin. Le signifié du verbe*, 173–209. Louvain: Peeters.

Murphy, Melissa Dae. 2008. The role of typological drift in the development of the Romance subjunctive: a study in word-order change, grammaticalization and synthesis. Doctoral dissertation: The University of Texas, Austin.

Nevalainen, Terttu & Helena Raumolin-Brunberg. 2017. *Historical Sociolinguistics: Language Change in Tudor and Stuart England.* 2nd Edition. London: Routledge.

Noonan, Michael. 2007. Complementation. In Timothy Shopen (ed.), *Language Typology and Syntactic Description,* Vol. 2, 52–150. Complex Constructions. Cambridge: Cambridge University Press.

Palmer, Leonard R. 1988. *The Latin Language (The Great Languages)*. Norman: University of Oklahoma Press.

Palmer, Frank R. 2001. *Mood and Modality*. Cambridge University Press.

Perotti, Pier Angelo. 1996. Sulle interrogative indirette in latino. *Latomus: revue d'études latines,* 55(3–4). 329–338.

Pinkster, Harm. 1990. *Latin syntax and semantics.* London: Routledge.

Poletto, Cecilia. 2000. *The Higher Functional Field: Evidence from the Northern Italian Dialects.* New York: Oxford University Press.

Poplack, Shana. 1989. The care and handling of a mega-corpus: The Ottawa-Hull French project. In Ralph W. Fasold & Deborah Schiffrin (eds.), *Language Change and Variation* (Current Issues in Linguistic Theory 52), 411–451. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Poplack, Shana. 1990. Prescription, intuition et usage: Le Subjonctif français et la variabilité inhérente. *Langage et Societe,* 54. 5–33. (doi:10.3406/lsoc.1990.2499)

Poplack, Shana. 1992. The inherent variability of the French subjunctive. In Christiane Laeufer & Terrell A. Morgan (eds.), *Theoretical analyses in Romance linguistics: selected papers from the nineteenth Linguistic Symposium on Romance Languages* (LSRL XIX), The Ohio State University, April 21-23, 1989, 235–263. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Poplack, Shana. 2011. Grammaticalization and linguistic Variation. In Bernd Heine & Heiko Narrog (eds.), *The Oxford Handbook of Grammaticalization,* 209–224. Oxford: Oxford University Press.

Poplack, Shana, Allison Lealess & Nathalie Dion. 2013. The evolving grammar of the French subjunctive. *International Journal of Latin and Romance Linguistics*, 25(1). 139–195. (doi:10.1515/probus-2013-0005)

Poplack, Shana & Stephen Levey. 2010. Contact-induced grammatical change: A cautionary tale. In Peter Auer & Jürgen Erich Schmidt (eds.), L*anguage and Space - An international handbook of linguistic variation*, Vol. 1: Theories and methods, 391–419. Berlin: Mouton de Gruyter.

Poplack, Shana & Sali A Tagliamonte. 2001. *African American English in the Diaspora.* Oxford: Basil Blackwell.

Poplack, Shana, Rena Torres Cacoullos, Nathalie Dion, Rosane e Andrade Berlinck, Salvio 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.), *Manuals in Linguistics: Romance Sociolinguistics*, 217-252. Berlin: Mouton de Gruyter.

Portner, Paul. 1998. The semantics of mood. *GLOT International*, 4(1). 3-8

Portner, Paul & Aynat Rubinstein. 2013. Mood and contextual commitment. In *Proceedings of SALT, 22*. 461–487. CLC Publications.

Quer, Josep. 2001. Interpreting mood. Probus. *Walter de Gruyter,* 13(1). 81–111. (doi:10.1515/prbs.13.1.81)

Romaine, Suzanne. 1982. *Socio-historical linguistics: its status and methodology*. Cambridge: Cambridge University Press.

Santulli, Francesca. 2009. Il congiuntivo italiano: morte o rinascita? In *Rivista Italiana di Linguistica e di Dialettologia*. 167–195.

Scarsi, Mariangela. 1996. *Satyricon / Petronius* (Classici Giunti). Firenze: Giunti.

Schachter, Paul & Timothy Shopen. 2007. Parts-of-speech systems. In Timothy Shopen (ed.), *Language Typology and Syntactic Description*. Vol. 1: Clause Structure, 1–60. 2nd edn. Cambridge: Cambridge University Press. (doi:10.1017/CBO9780511619427.001)

Schmitt Jensen, Jørgen. 1970. *Subjonctif et hypotaxe en italien: une esquisse de la syntaxe du subjonctif dans les propositions subordonnées en italien contemporain.* Odense: Odense University Press.

Schneider, Stefan. 1999. *Il congiuntivo tra modalità e subordinazione: uno studio sull'italiano parlato*. Roma: Carocci.

Shlonsky, Ur. 2006. Projection étendue et cartographie de SC. *Nouveaux cahiers de linguistique française,* (27). 83–93.

Simone, Raffaele. 1993. Stabilità e instabilità nei caratteri originali dell'italiano. In Alberto Sobrero & Paola Benincà & Gaetano Berruto (eds.), *Introduzione all'italiano contemporaneo*. Vol. I: Le strutture. 41–100. Roma - Bari: Laterza.

Soliman, Luciana T. 2002. *Modalità e implicazioni aspettuali: analisi contrastiva italiano/francese del congiuntivo nell'ottica psicomeccanica.* In Atti del Convegno Attorno al congiuntivo. Storia, tipologia, traduzione. 291–306.

Stefanelli, Stefania. 2006. *Va in scena l'italiano: la lingua del teatro tra Ottocento e Novecento.* Firenze: Cesati.

Terrell, Tracy & Joan Hooper. 1974. A semantically based analysis of mood in Spanish. *Hispania,* 57. 484–494.

Thompson, Sandra A. 1998. A discourse explanation for the cross-linguistic differences in the grammar of incorporation and negation. In Anna Siewierska & Ja Jung Song (eds.), *Case, Typology and Grammar*. Vol. 38, 309–342. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Veland, Reidar. 1991. L'influence modale des verbes d'opinion en italien contemporain: Étude d'un corpus. *Studia Neophilologica: A Journal of Germanic and Romance Languages and Literature*, 63(2). 209–220.

Voghera, Miriam. 1993. La grammatica del LIP. In Tullio De Mauro, Federico Mancini, Massimo Vedovelli & Miriam Voghera (eds.), *Lessico di frequenza dell'italiano parlato*, 86–111. Milano: Etas.

Voghera, Miriam. 2001. Teorie linguistiche e dati di parlato. In Francesco Albano Leoni, Lucia Cresti & Giorgio Fiorentino (eds), *Dati empirici e teorie linguistiche*, 75-95. Roma: Bulzoni.

Wandruszka, Ulrich. 1998. Frasi subordinate al congiuntivo. In Lorenzo Renzi & Giampaolo Salvi (eds.), *Grande grammatica italiana di consultazione*, vol. II, 415–481. Bologna: Il Mulino.

Whaley, Lindsay. 1997. *Introduction to Typology: The Unity and Diversity of Language.* Thousand Oaks, CA: SAGE Publications. (doi:10.4135/9781452233437)

Appendix

This appendix lists all matrix governors that selected the subjunctive at least once in the contemporary Italian dataset. For each, it provides an English translation, the subjunctive rate, token count, proportion of overall data (% Data), proportion of total subjunctive forms (% Morpho), and the increment, a metric indicating whether the

governor contributes disproportionately more subjunctive morphology than expected given its frequency in the data.

Governor	Translation	% Subj	N	% Data	% Morpho	Increment
non è	it is not	33%	37/112	18%	9%	-9%
(non) pensare	to (not) think	73%	64/88	14%	16%	2%
(non) sembrare	to (not) seem	75%	48/64	10%	12%	1%
(non) credere	to (not) believe	79%	45/57	9%	11%	2%
bisognare	to be necessary	93%	41/44	7%	10%	3%
(non) parere	to (not) appear	58%	15/26	4%	4%	-1%
(non) dire	to (not) say	19%	5/26	4%	1%	-3%
(non) volere	to (not) want	84%	16/19	3%	4%	1%
può darsi	it might be	58%	11/19	3%	3%	0%
sperare	to hope	93%	14/15	2%	3%	1%
bastare	to be enough / suffice	50%	6/12	2%	1%	0%
(non) essere sicuro	to (not) be sure	63%	5/8	1%	1%	0%
aspettare	to wait / to expect	50%	4/8	1%	1%	0%
non sapere	to not know	33%	2/6	1%	0%	0%
è facile	it is easy	100%	4/4	1%	1%	0%
è importante	it is important	100%	4/4	1%	1%	0%
avere paura	to be afraid	75%	3/4	1%	1%	0%
è meglio	it is better	75%	3/4	1%	1%	0%
fare sì	to ensure / to make sure	75%	3/4	1%	1%	0%
avere l'impressione	to have the impression	50%	2/4	1%	0%	0%
è inutile	it is useless	50%	2/4	1%	0%	0%
immaginare	to imagine	25%	1/4	1%	0%	0%
(non) è bene	to (not) be good	100%	3/3	0%	1%	0%
avere la sensazione	to have the feeling	100%	3/3	0%	1%	0%
esigere	to demand	100%	3/3	0%	1%	0%
non è detto	to not be certain	100%	3/3	0%	1%	0%
ritenere	to consider	100%	3/3	0%	1%	0%
supporre	to suppose	100%	3/3	0%	1%	0%
presupporre	to presuppose	67%	2/3	0%	0%	0%
richiedere	to require / to demand	67%	2/3	0%	0%	0%
(non) è possibile	to be (not) possible	33%	1/3	0%	0%	0%
è ovvio	it is obvious	33%	1/3	0%	0%	0%
ci sta	it is possible	100%	2/2	0%	0%	0%
è difficile	it is difficult	100%	2/2	0%	0%	0%
è impossibile	it is impossible	100%	2/2	0%	0%	0%
mettere	to assume	100%	2/2	0%	0%	0%
non succedere	to not happen	100%	2/2	0%	0%	0%

Governor	*Translation*	*% Subj*	*N*	*% Data*	*% Morpho*	*Increment*
può essere	it may be	100%	2/2	0%	0%	0%
calcolare	to calculate	50%	1/2	0%	0%	0%
dubitare	to doubt	50%	1/2	0%	0%	0%
fare in maniera	to ensure	50%	1/2	0%	0%	0%
non è vero	to not be true	50%	1/2	0%	0%	0%
trovare	to find	50%	1/2	0%	0%	0%
(non) dispiacersi	to (not) mind	100%	1/1	0%	0%	0%
(non) è giusto	to (not) be fair	100%	1/1	0%	0%	0%
assicurarsi	to make sure	100%	1/1	0%	0%	0%
avere il dubbio	to have doubt	100%	1/1	0%	0%	0%
controllare	to check	100%	1/1	0%	0%	0%
dedurre	to deduce	100%	1/1	0%	0%	0%
è bello	it is nice	100%	1/1	0%	0%	0%
è inevitabile	it is inevitable	100%	1/1	0%	0%	0%
è strano	it is strange	100%	1/1	0%	0%	0%
essere orgoglioso	to be proud	100%	1/1	0%	0%	0%
evitare	to avoid	100%	1/1	0%	0%	0%
fare in modo	to ensure / to make sure	100%	1/1	0%	0%	0%
fare piacere	to please	100%	1/1	0%	0%	0%
l'importante è	the important thing is	100%	1/1	0%	0%	0%
lasciare	to let / to allow	100%	1/1	0%	0%	0%
non avere il dubbio	to have no doubt	100%	1/1	0%	0%	0%
non avere senso	to make no sense	100%	1/1	0%	0%	0%
non è da dire	it is not to be said	100%	1/1	0%	0%	0%
non è umano	it is not humane	100%	1/1	0%	0%	0%
non preoccuparsi	to not worry	100%	1/1	0%	0%	0%
preferire	to prefer	100%	1/1	0%	0%	0%
presumere	to presume	100%	1/1	0%	0%	0%
reputare	to deem / to consider	100%	1/1	0%	0%	0%
rischiare	to risk	100%	1/1	0%	0%	0%
sentire il timore	to feel fear	100%	1/1	0%	0%	0%
servire	to be needed	100%	1/1	0%	0%	0%
sorprendere	to surprise	100%	1/1	0%	0%	0%
verificare	to verify	100%	1/1	0%	0%	0%
Total		**66%**	**404/616**			

CONTACT

salvio.digesto@carleton.ca