ZERL
Zeitschrift für Europäische Rechtslinguistik

Das eJournal der Europäischen Rechtslinguistik (ERL)
Universität zu Köln

Linguistic variation in primary EU legal texts: Evidence from (null) subjects in the Italian Constitution and the Italian version of the Treaty of the European Union

Katrin Schmitz

30. Dezember 2025
https://doi.org/10.18716/ojs/zerl.2025/12842
www.zerl.uni-koeln.de

Abstract English

Many recent studies in legal linguistics investigated the variation within and across legal varieties of various EU Treaty languages focusing on the question of whether the multilingualism of the EU has led to the development of so-called eurolects with properties different from those of the national legal variety. But a crucial property of many Treaty languages, namely null subjects, has not been analysed systematically. A proper assessment of the presumed influence of English as the main drafting language requires detailed analyses of how syntax and pragmatics interact in this grammatical domain together with the specific requirements of legal language in the legal varieties of null subject Treaty languages. The present study aims to contribute to this overarching goal by providing a quantitative and qualitative analysis of two primary Italian legal texts, namely the complete text of the Italian version of the Treaty of the European Union (TEU) and its functionally equivalent parts of the Italian Constitution (CONST). The central finding is that the domain of subject expression in the two studied legal documents, although far apart in time, follows the individual language properties and processing properties, contradicting cross-linguistically motivated assumptions related to an absence of null subjects. Observed effects of the influence from English concern a reduced number of postverbal subjects and the increased use of overt subject pronouns (*egli*/*esso* series) in the TEU.

Abstract Deutsch

Viele neuere rechtslinguistische Arbeiten haben die Variation innerhalb und zwischen Varietäten verschiedener EU-Vertragssprachen im Hinblick auf die Frage untersucht, ob die Mehrsprachigkeit der EU zur Entwicklung sogenannter Eurolekte geführt hat, die von denjenigen der jeweiligen nationalen Rechtssprachvarietät abweichende Eigenschaften aufweisen. Jedoch wurde eine zentrale Eigenschaft vieler EU-Vertragssprachen nicht hinreichend berücksichtigt, nämlich die Nullsubjekteigenschaft. Für eine angemessene Erfassung des angenommenen Einflusses des Englischen als wichtigste Entwurfssprache muss das komplexe Zusammenspiel von Syntax und Pragmatik dieses grammatischen Bereichs in Kombination mit den spezifischen Anforderungen von Rechtssprache in den Rechtssprachvarietäten von Nullsubjekt-Vertragssprachen untersucht werden. Die vorliegende Untersuchung versteht sich als ein Beitrag zu diesem übergreifenden Ziel, indem sie eine quantitative und qualitative Analyse von zwei primären italienischen Rechtstexten, nämlich die italienische Fassung des Unionsvertrags (TEU) und der funktional entsprechenden Teile der italienischen Verfassung (CONST) vorlegt. Als zentrales Ergebnis lässt sich festhalten, dass die beiden untersuchten italienischen Rechtsdokumente trotz ihres zeitlichen Abstands den individuellen sprachlichen Eigenschaften und Verarbeitungs-präferenzen folgen; sprachübergreifende Behauptungen über die Struktur von Rechtstexten, die einen Ausschluss von Nullsubjekten beinhalten, werden widerlegt. Effekte des angenommenen Spracheneinflusses wirken sich im TEU in einer reduzierten Anzahl postverbaler Subjekte sowie einer Zunahme der Verwendung von overten Subjektpronomina der *egli*/*esso*-Serie aus.

Abbreviations

AdvP	Adverbial Phrase
AP	Adjective Phrase
DP	Determiner-Phrase
CONST	(Italian) Constitution
EU	European Union
NP	Nominal Phrase
PP	Prepositional Phrase
TEU	Treaty of the European Union

Contents

1. Introduction 5

1.1. Linguistic variation and (null) subjects in the legal context 5

1.2. Aim of this study 8

2. Subject expression in (Standard and Legal) Italian 10

2.1. Null and overt subjects in the grammatical system 10

2.1.1. The null subject property: Syntactic licensing and pragmatic identification of null subjects 10

2.1.2. Subject pronoun types 12

2.1.3. Subject position 14

2.2. The distribution of null and overt subjects in written legal Italian 15

2.3. Research questions and hypotheses of this study 17

3. Methodology 18

3.1 The Data 18

3.2. The Analysis 18

4. Results 20

4.1 Quantitative Analysis 20

4.1.1 Distribution of null vs. overt subjects 20

4.1.2 Types and positions of overt subjects 21

4.2 Qualitative Analysis 23

4.2.1. Null subjects and their referents 23

4.2.2 *Esso/egli* subjects and their referents 25

4.2.3. Postverbal subjects in the TEU 27

5. Discussion 29

6. Concluding remarks 30

7. References 31

1 Introduction

1.1 Linguistic variation and (null) subjects in the legal context

Recent years have seen a notable increase of studies of legal varieties of various EU languages with respect to linguistic variation. They focused on the question of whether the multilingualism of the EU has led to the development of so-called eurolects with properties different from those of the national legal variety (cf. in particular the *Eurolect Observatory Project*[1], MORI 2018a, and the critical reviews of the volume by MATTILA 2019 and HEINEMANN 2019). Eurolects are linguistic systems and a contact variety of the national legal variety; they enlarge the sociolinguistic space of the Treaty languages (MORI & SZMRECSANYI 2020: 2). The results obtained by the *Eurolect Observatory Project* as well as earlier work (cf. VENTURI 2012) for phenomena of morphosyntactic nature confirm the existence of the eurolects.

Although many EU Treaty languages are typologically null subject languages, among them most Romance Treaty languages, subject expression remains a largely understudied domain in legal language research to date[2]. These languages permit their subjects to be phonologically silent under topic continuity (cf. section 2 for details). MATTISSEN (2022: 470) observes that the use of subject types differs in the TEU versions of those Romance Treaty languages which allow null subjects. Although chiefly concerned with reference tracking of overt (subject and object pronouns), MATTISSEN (2022: 470) discusses the following two example cases, namely Article 17, para. 1 (example 1) and Article 3, para. 3 TEU (example 2) in the Spanish, Italian, Portuguese and Romanian versions (cf. MATTISSEN 2022: 470, examples (4-5)). Null subjects identified by personal marking are highlighted by underlining the verb whereas overt pronouns are marked in bold[3]:

(1)

[ES] 1. **La Comisión** promoverá […]. Velará [...]. Supervisará [...]. Ejecutará [...]. Ejercecerá asimismo [...], asumirá la represensatición exterior de la Unión. Adoptará [...].	[IT]. **La Commissione** promuove [...]. Vigila […]. Vigila […]. Dà esecuzione […]. Esercita […]. Assicura la representanza esterna dell'Unione, […]. Avvia […]
[PT] 1. **A Comissão** promove […]. **A Comissão** vela […]. Controla […]. **A Comissão** executa […]. Exerce […]. […], **a Comissão** assegura a representacão externa da União. Toma […].	[RO] (1) **Comisia** promovează […]. **Aceasta** asigură [...]. **Comisia** supraveghează [...]. **Aceasta** execută [...]. **Comisia** exercită [...]. [...], **aceasta** asigură reprezentarea externă a Uniunii. **Comisia** adoptă [...].

[1] See https://www.unint.eu/en/research/research-projects/33-page/490-eurolect-observatory-project.html [accessed April 10th, 2020].

[2] The detailed analysis for Italian (MORI 2018b) does not include a proper analysis of the domain of null subjects, neither does the referred previous work by MORI (2003). In contrast, recent work by CANAVESI and MORI (2024) and TOMBESI (2024) include null subjects (cf. section 2.2 for details).

[3] It is unclear why MATTISSEN (2022: 470) highlights the adverb *asimismo* 'likewise' in the Spanish version in ex. (2), given here in its original version, as it is not a subject.

(2)

[ES] 3. **La Unión** establecerá un mercado interior. Obrará [...]. **Asimismo**, promoverá el progreso cierntífico y técnico.	[PT] 3. A União estabelece um mercado interno. Empenha-se […]. A União fomento a progresso cientifico e tecnológico.
La Uniòn [...]	A União […].
La Uniòn [...]	A União […].
La Uniòn [...]	A União […].
[RO] (3) Uniunea instituie o piaţă internă. Aceasta acţionează […]. **Aceasta** promovează progresul ştiinţific şi tehnic.	[IT] 3. **L'Union**e instaura un mercato interno. Si adopera […]. **Essa** promuove il progresso scientifico e technologico.
Uniunea […].	L'Unione [...].
Aceasta [...]	**Essa** [...].
Uniunea […].	**Essa** [...].

In example (1), the Spanish and Italian versions do not use subject pronouns but null subjects in the non-first sentences of the paragraph while the Portuguese version reduplicates the lexical subject in all sentences and the Romanian version varies between reduplication of the lexical subject and the use of the demonstrative pronoun *aceasta*. In example (2), both the Spanish and Portuguese versions reduplicate the subject in all non-first sentences, while the Romanian version again alternates reduplication and use of *aceasta* and Italian uses the personal pronoun *essa* in the third and fourth sentence. Obviously, the pro-drop property shared by these legal varieties combines with language-specific pronoun types and processing preferences leading to systematic microvariation. Linguistic variation has, however, to be strongly controlled to ensure coherence and legal certainty in the multilingual EU setting. To guarantee the unity of the European law in legislation and jurisdiction, every EU document must be available in the 24 Treaty languages and its versions are equally authentic and authoritative. Given the different linguistic properties of the Treaty languages, the linguistic means used in each Treaty language to achieve the unity may differ. Among various types of possible divergences between language versions of EU legal acts (cf. BURR 2013), language-specific syntactic and pragmatic differences, including the present case of null subjects, are considered as instances of the "linguistic-pragmatic convention" (cf. BURR 2013: 1515). As such, they should not threaten the unity of EU law. An interesting question concerns the interplay of various factors shaping the variation reproduced in the accepted linguistic-pragmatic conventions. These factors include intra-linguistic properties (syntactic and pragmatic features) of null and overt subjects and language-specific and processing preferences of the speech community. The latter include, among others, the distance between anaphoric subject pronouns and their antecedent (cf. e.g. CONTEMORI & DI DOMENICO 2021 and section 2.1). But extra-linguistic factors also play a role, namely the degree of systematic language contact in the individual speaker/writer and the legal culture reflected in the requirements of the legal language variety. Both will be briefly explained in the following lines.

In psycholinguistics, language contact in the bi-/multilingual speakers is a crucial issue, particularly regarding its effects across different linguistic domains and developmental

conditions. The interplay of syntactic licensing of the subject omission and pragmatic identification of the referent as well as its discourse status is one of the most studied phenomena. Psycholinguistic research of the last decades has highlighted the vulnerability of this interplay in bilingual speakers raised from early age in environments where Italian or Spanish are in constant contact with English as dominant language. Evidence includes temporary influence of the dominant language or even convergence processes and thus new varieties where null subjects tend to be replaced by overt subject pronouns and preverbal overt subjects are preferred over postverbal ones (cf. e.g. MONTRUL 2016, SCHMITZ et al. 2016, SCHMITZ & SCHERGER 2025). This constellation also holds for many European professionals (lawyers, translators, lawyer-linguists) with multilingual professional training and possibly also a bi-/multilingual background.

Obviously, principles proper to legal language and legal genres represent a crucial factor shaping legal texts, which also have an impact on overt subject expression. In any legal context, the most important principle requires the avoidance of ambiguity and precision of interpretation. Compliance with this principle consequently leads to a high level of redundancy, generally due to the pleonastic use of lexical items, which in turn involves a violation of the principle of conciseness, a distinctive feature of specialized discourse, since the number of lexemes employed is far higher than necessary (cf. GOTTI 2012: 52). Anaphoric reference is one of the most common devices deployed to increase textual cohesion. In conjunction with other referential phenomena, such as ellipsis, substitution and lexical cohesion, it forms the textual framework which – combined with suitable cohesive devices – accounts for the constituent features of a text (cf. GOTTI 2012: 55f). Finally, also the type of legal genre plays a role. GOTTI (2012: 61) highlights the fact that one of the phenomena that most distinguishes legal discourse is compliance with the norms governing the construction of its different text genres, i.e. there is a close link between the type of legal text and its structure.

A more specific issue concerns the role of the sentence level in legal texts for anaphoric reference. Based on English national law documents, HILTUNEN (2012: 41ff.) describes several fundamental properties of syntactic structures in legal texts, claiming in particular that a) the sentence constitutes the basic syntactic unit and is traditionally constructed as a self-contained, context-free entity; (b) sentences in statutory texts are independent; (c) there is avoidance of grammatical ties across sentence boundaries, including pronominal anaphoric references; (d) the most common sentence type is a complex sentence (one main clause with one or more subordinated clauses). Most importantly for the use of (null) subjects, HILTUNEN (2012: 48f.) claims the absence of pronoun references, not only within the sentence but also across sentence boundaries. The preferred strategy is the repetition of full noun and verb phrases, even at the expense of readability. MATTISSEN (2022: 464) confirms the importance of the traceability of discourse referents for the clarity of legal texts in the multilingual legislation context of the EU, both within and across texts over a series of sentences which involves a syntactic analysis of EU legal texts. Nevertheless, the sentence

level (and a related syntactic analysis) remains important, as neuronal machine translation tools still operate primarily at the sentence level (MATTISSEN 2022: 464).

Since EU legal acts are equally mandatory in all 24 official languages of the EU and form together an original (cf. Art. 55 TEU), the translation of the negotiated drafts into the official languages and the subsequent text editing by the lawyer-linguists face a high-quality standard. The resulting texts have to be semantically convergent and possess a similar text structure (subdivided in articles, paragraphs, sentences etc., cf. MATTISSEN 2022: 464). To draft legislation at the supranational EU level, several resources have been developed and translated into all 24 Treaty languages, namely the *Joint Handbook for the presentation and drafting of acts subject to the ordinary legislative procedure* by the European Parliament, the Council of the EU and the European Commission (2020) (henceforth *Joint Handbook*) and the *Joint Practical Guide of the European Parliament, the Council and the Commission for persons involved in the drafting of European Union legislation* (2015, henceforth *Joint Practical Guide*). They mainly provide formal requirements regarding the division of legal acts, standardized formulations of specific text parts and other standardisation. There are, however, also linguistically relevant remarks concerning the clarity and unambiguity of formulations and syntactic constructions (cf. *Joint Handbook* D. 4.4.2, *Joint Practical Guide* 1.1, 4, 4.5.2, 5.2-5.23; cf. MATTISSEN 2022: 466 and references therein). A rule of special interest for the present study refers to the non-substitution of an introduced referent at the beginning of each new paragraph in an article (cf. *Joint Handbook* D. 4.3, *Joint Practical Guide* 1.4.1, 6-6.4); this rule allows for a citation of each paragraph in isolation (cf. MATTISSEN 2022: 467 for the German implementation). There is no specification, however, for paragraphs consisting of several sentences, allowing thus also null subjects as will be shown in this study.

1.2 Aim of this study

The preceding section has introduced three groups of factors which affect the use of (null) subjects in the legal varieties and potentially increase the observed variation, namely (a) language-specific properties and processing preferences of null and overt subject pronouns (as witnessed in the examples in 2 and psycholinguistic research), (b) speaker-/writer-specific factors, i.e. individual bi-/multilingual background and training of professionals in the multilingual EU context, and (c) properties of legal language in general and of particular legal genres and legal systems. The present study aims to contribute to a better understanding of the systematicity of the use and distribution of null and overt subjects in the legal varieties of the null subject Treaty language Italian. To this end, the study provides a detailed quantitative and qualitative analysis of two documents of primary law, namely the Italian version of the EU Treaty (henceforth TEU) and the functionally comparable parts of the Italian Constitution (henceforth CONST), expanding largely the preliminary analysis in SCHMITZ (2023).

Several reasons motivated the choice of documents of primary law. First, the *Observing Eurolects* project bases its work on a large data base including a huge number of directives

(secondary law involving several technolects) and a joint analytic template (see MORI 2018a). The absence of an analysis of null subjects in the Italian report (MORI 2018b) might, at least partially, be explained by the conditions set in the Italian version of the *Joint Handbook* (2020: 68, section D.4.3), requiring the repetition of lexical subjects and objects instead of the use of subject and object pronouns of an article or paragraph in another article or paragraph, giving the following example:

'Anziché scrivere:

"1. La Commissione adotta tutte le misure necessarie …

2. **Essa** adotta **tali** misure secondo la procedura …",

si scriva:

"1. La Commissione adotta tutte le misure necessarie …

2. **La Commissione** adotta **le misure di cui al paragrafo 1** secondo la procedura …". '

Second, these conditions were made for EU legal acts (i.e. secondary law) whereas, the fulfilment in primary EU law texts has not yet been systematically evaluated. To the best of the author's knowledge, EU primary law has so far not been systematically analysed in terms of null subjects. One might criticise the comparison of CONST and TEU since they are far apart in time and could be argued to be hardly comparable[4] if the observed social and linguistic changes leading to a "Neo-Standard" of Italian (cf. BERRUTO 2012) are considered: CONST would then be written in the traditional standard and the more recent Italian versions of TEU in the "Neo-Standard". Such an assumption, however, faces several problems: (1) There is no break but a co-existence of the traditional standard norm and the neo-standard norm and there is no evidence that the traditional standard is losing its official prestige (CERRUTI et al. 2017: 16ff.). (2) According to AUER (2017: 369ff.), the neo-standard is situated below the traditional standard and linked to one or more of the following characteristics: orality, informality, subjectivity/personalization and modernity. None of these features applies to legal language which is a highly controlled and formal written variety associated with formality, impersonality/neutrality and archaisms (cf. e.g. MATTILA 2016). These general arguments already suggest a continuity of the traditional norm in legal Italian allowing a comparison across time. (3) The continuity in the application of the traditional standard in legal Italian can also be proven empirically: among the morphosyntactic properties associated with the neo-standard, is the use of the oblique forms *lui, lei* and *loro* as subject personal pronouns, replacing *egli/ella, essi/esso* (BERRUTO 2017: 45). The detailed analysis presented in this paper shows that such a substitution did not happen in the Italian version of the TEU. Overall, it will show a strong degree of continuity and similarity in the use of null and overt subjects, especially traditional pronominal forms, which underline the systematicity and robustness of the traditional standard in legal Italian.

[4] I thank an anonymous reviewer for raising the issue of restandardization of the Italian language (cf. BERRUTO 2012).

Finally, and most importantly, in accordance with the genre criteria in GOTTI (2012), CONST and TEU are statutory texts with important and ideal properties for the aim of this study. Like other legal texts, they are written and formal but they are closer to general language than specific EU legal acts, which often combine several specialized languages and thus do not provide a clear picture of the specific impact of legal language on subject expression. At the same time, it is highly probable that the used consolidated TEU version includes typical properties of the Italian eurolect.

The paper has the following structure: Section 2 presents the subject domain in Italian, including also available studies on the distribution of null and overt subjects in different written legal varieties of Italian and the working hypotheses. Section 3 presents the methodology of the present paper. The results are presented in section 4 and discussed in Section 5. Section 6 concludes the paper with some implications of the results for further legilinguistic and psycholinguistic research in the subject domain in legal varieties of null subject treaty languages.

2 Subject expression in (Standard) Italian

2.1 Null and overt subjects in the grammatical system

2.1.1 The null subject property: Syntactic licensing and pragmatic identification of null subjects

As detailed in many works and resumed, among others, in SCHMITZ et al. (2016) and SCHMITZ (2023), Italian is a null-subject language that allows for null subjects in all grammatical persons, tenses and sentence types (main and subordinate clauses). It does not require an overt phonological realization of subjects since person and number agreement morphology assures their recovery. Pragmatics, in particular topic (dis)continuity, regulates the choice of null or phonologically overt subjects: unless a new referent is introduced, giving rise to *topic discontinuity* whereas the maintenance of the same referent *is topic continuity*. Establishing a contrast between referents in the discourse also requires overt subjects (cf. MAYOL 2010 for different types of contrasts). The examples from the Italian version of the Treaty of the European Union, henceforth TEU, in (3) and (4) illustrate, in turn, *topic continuity* (see ◎ of the second subject in example 4, Art. 3, para. 3, cl. 4 TEU) and *topic discontinuity* (from the Italian and English version of TEU, Art. 5, para. 1, cl. 1 and 2):

(3) **L'Unione** combatte l'esclusione sociale e le discriminazioni e Ø promuove la giustizia e la protezione sociale, la parità tra donne e uomini, la solidarietà tra le generazioni e la tutela dei diritti del minore.
'It shall combat social exclusion and discrimination, and shall promote social justice and protection, equality between women and men, solidarity between generations and protection of the rights of the child.'

(4) **La delimitazione delle competenze dell'Unione** si fonda sul principio di attribuzione. **L'esercizio delle competenze dell'Unione** si fonda sui principi di sussidiarità e proporzionalità.
'The limits of Union competences are governed by the principle of conferral. The use of Union competences is governed by the principles of subsidiarity and proportionality.'

The interpretation of null subjects and overt subject pronouns depends on the possibility to identify their referents. This process is called anaphora resolution, particularly in psycholinguistic research, or reference tracking (cf. MATTISSEN 2022). The interpretation may differ among null subject languages and explain some of the observed preferences that we may find in both non-legal and legal language. For Italian, the influential study by CARMINATI (2002) introduced the Position of Antecedent Hypothesis (PAH) and claimed that in this language, null-subject pronouns co-refer with antecedents in the specifier of the inflectional phrase (IP) in generative terms, i.e. the structural position occupied by the preverbal subject of a tensed verb, while overt pronouns refer to antecedents in a lower position. This psycholinguistic experimental study investigated intra-sentential anaphora resolution based on items with an initial subordinate clause containing the referents and a following main clause including the null or overt subject pronoun – a rather rare order for legal texts (cf. section 3.2 for details on the analysed primary law texts). Other studies followed this design and investigated the grammatical function of the referent and its structural position to check the validity of the PAH in further null subject languages (cf. summaries in CONTEMORI and DI DOMENICO 2021, LOZANO and QUESADA 2023). In their recent work, CONTEMORI and Di DOMENICO (2021) provided further experimental comparative work on comprehension and production of Spanish and Italian speakers. Overall, these authors found microvariation in the two null-subject languages, with Spanish observing the PAH, but to a lesser degree than Italian. Most recently, LOZANO and QUESADA (2023) pointed to the fact that current research does not agree on the cross-linguistic validity of the PAH. Based on the results of their corpus study with monolingual Spanish and bilingual L1-English/L2-Spanish speakers, LOZANO and QUESADA (2023: 20) argue in favour of a larger flexibility and of the integration of further factors, e.g. the distance between anaphora and their referents. This latter factor might be pertinent also for the analysis of legal written language. Furthermore, legal texts often contain reference chains, i.e. two or more expressions in a complex sentence or across sentences within a legal text or even across related legal texts referring to the same discourse referent (MATTISSEN 2022: 464). The resulting requirement of traceability of the discourse referents both intra- and intertextually across several sentences is designed in European legal acts in the way that, upon maintenance of idiomatic expression, a verification of single sentences is possible (MATTISSEN 2022: 464). Within the scope of the present study, the referents of null and overt subject pronouns, as well as the distance between them will be examined at the article level to detect possible differences in sentence structure between the monolingual Italian CONST and the Italian version of the TEU, which may in turn affect the distribution of null subjects in these legal documents (cf. sections 3.2 and 4.2).

2.1.2 Subject pronoun types

Overt subjects in Italian can be lexical (proper names, noun phrases, determiner phrases) or pronominal (cf. MATTISSEN 2022: 468 for a list of factors influencing the substitution of lexical subjects by pronouns). Looking more closely at the different types of overt subject pronouns (and pronouns able to function as a subject, among other grammatical functions), we find demonstrative[5], relative, interrogative and personal pronouns which, according to CARDINALETTI & STARKE (1999), can be strong, weak or clitic. As a consistent null subject language, Italian does not have subject clitics (in contrast to the non-null subject language French). The only exception is *si impersonale* which refers to an unspecified group of humans, which usually includes the speaker (cf. CHIERCHIA 1995), being thus also pertinent to legal texts which favour impersonal formulations. According to both CHIERCHIA (1995; 107) and D'ALESSANDRO (2002: 39), *si impersonale* is a clitic pronoun. D'ALESSANDRO (2002: 39) further assumes that Italian impersonal *si* is referential. All other subject personal pronouns are tonic pronouns falling in two categories: strong pronouns like 3rd person *lui/lei* and weak pronouns where we find *egli/ella* and *esso/essa.* Their classification and description in both a synchronic and a diachronic perspective continue to intrigue researchers as the recent contributions by CAPPELLARO (2017), CARDINALETTI (2021) and CORDIN (2021) and work cited therein witness. Focusing on different aspects, each one contributes to a better understanding of their distribution in current written Italian texts, including legal Italian, as the following lines will show.

CAPPELLARO (2017) focuses on the semantic specialization of Italian subject pronouns. She argues that within the *egli/esso* series, *esso* (unlike its feminine counterpart *essa* and both plural forms *essi/esse* as well as *egli/ella*) has developed obtaining the semantic specialization [- human] only. In Modern standard language – which CAPPELLARO (2017: 114) does not relate to BERRUTO's (2012) definition of Neo-standard Italian - *lui, egli* and *esso* are in complementary distribution according to (in)formal register (non-)human reference. CAPPELLARO (2017: 114) defines *esso* as [+ formal, - human], *lui* as [- formal, + human] and *egli* as [+ formal, + human]. Based on a detailed quantitative and qualitative analysis of 48 texts from the 13th century to the first decade of the 20th century, including treatises, CAPPELLARO (2017) shows the overall distribution and development of *esso*. Importantly, in the past centuries, *esso* could refer to both human and non-human referents and both *esso* and *egli* were near synonyms; until today, this holds for *ella* and *essa*.

CARDINALETTI (2021) compares the distribution and syntactic description of Italian weak pronouns in Old and Modern Italian (without reference to the debate on standardization). Weak pronouns in Modern Italian include the subject pronouns *tu* 'you' and *egli/esso* 'he/she' as well as the object dative form *loro* 'to them' found in high registers (cf. CARDINALETTI 2021: 131). The weak character of *egli/esso* is motivated syntactically: They cannot be dislocated to the left and appear in postverbal position (unlike strong pronouns), but, unlike

[5] Cf. MATTISSEN (2023) for a comparative analysis of the use of demonstrative pronouns and their referents in different EU legal acts and Treaty languages. The analysis in the present study will focus on the use of personal pronouns.

subject clitics, they can be separated from the from the verb by parenthetic elements or adverbs (cf. CARDINALETTI 2021: 135). For the distribution of *egli*/*esso* in Modern Italian, CARDINALETTI (2021: 136-137) observes two situations of use (based on literary examples not reproduced here): a) when the antecedent is not prominent, and b) when the antecedent is perceived as too "distant" by the speaker/writer. Finally, CARDINALETTI (2021: 141) shows that the strong pronouns *lui*/*lei* are used "weakly" in the substitution of null pronouns in contexts of topic maintenance. For the textual conditions in which *egli*/*esso* are preferred over the null weak form, CARDINALETTI (2021: 151) presumes an ongoing language change leading to the substitution of the weak forms *egli*/*esso* by the innovative forms *lui*/*lei* which, in turn, are becoming weak in this process (cf. CARDINALETTI 2021: 151-152).

CORDIN (2021) follows CARDINALETTI (2021) in the overall classification and description of the Italian weak pronouns as well as in the ongoing grammaticalization but adds pertinent details on both their diachronic development and distribution in contemporary Italian including written corpus data. *Egli* and *esso* are characterized as residuals from a grammaticalization process which started in the 15th century and was interrupted, resulting in the absence of subject clitics (cf. CORDIN 2021: 159). In contemporary Italian, the residual nominative pronouns *egli* and *esso* (and their inflected forms) are restricted to the third person and to formal register, mainly in written texts (CORDIN 2021: 160). Based on the KiParla corpus with semi-sponteanous speech collected in university contexts, CORDIN 2021: 162ff.) discusses utterances and elaborates the following differences between *egli* and *esso* (with its inflected forms): a) the use of the weak subject pronouns is very rare and limited to situations which require a very controlled speech (e.g. lectures and exams), b) the forms with the root *ess-* are less rare and more frequent in postverbal position, often strengthened by an adjective or adverb, d) *essi* always refers to human subjects. The focus of CORDIN's paper – and of special interest for this paper – is the presentation and discussion of criteria for the selection between *pro* and the weak subject pronouns *egli* and *esso*, emphasizing the central role of the text (cf. CORDIN 2021: 168ff.). Based on SALVI and VANELLI (2004: 193), CORDIN (2021: 170ff.) presents an important notion related to the communicative structure: the expected referent, in turn related to the theme and the thematic continuity. Both *pro* and *egli* can only introduce an expected referent while the pronoun *lui* refers to an unexpected referent. In examples with expected referents, *egli* is preferred to the null pronoun when the antecedent and the pronoun are "relatively distant" in the discourse as the following example (5) from SALVI and VANELLI (2004: 193, (21c), in CORDIN (2021: 170, (28)) illustrates:

(5) Il solitario Abu Mazen$_i$ sta dimostrando, a mio avviso, in queste ore un grande coraggio. Da una posizione ingrata, impopolare, insostenible, egli$_i$ (*? pro$_i$) difende la ragione. (*Repubblica*, 22 agosto 2003).

CORDIN (2021: 171) further discusses the notions of expected referent and distance. While the expected referent is interpreted as introduced in a preceding utterance to the position of the pronoun (often, but not necessarily with the role of theme), the distance between antecedent and pronoun normally refers to the number of phrases separating the elements.

Since there are also examples with a short distance between antecedent and pronoun, CORDIN (2021: 171) critically discusses the conception of distance in terms of length of the phrasal chain and proposes to replace it by the concept of syntactic independence of the utterances where the antecedent and the pronoun are placed. The explicit expression of the features number, person, gender, animacy and case of the pronoun facilitate the interpretation of the anaphoric relationship established between elements in two different utterances. Overall, CORDIN (2021: 172) characterizes written Italian as an "apparently hypertrophic" system with three types of subject pronouns (strong, weak, null) which are not competing since the anaphoric relation of the contexts in which they appear differs: the strong pronoun is related to an unexpected referent, the null pronoun to an expected referent within the complex sentence and the weak pronoun to an expected referent which is often found outside the complex sentence.

Summarizing and combining the characterizations proposed by CAPPELLARO (2017), CARDINALETTI (2021) and CORDIN (2021), the class of weak Italian subject pronouns in standard written Italian includes the following forms, each with different syntactic, semantic and textual features (with differences among the analyses highlighted in bold): (a) *pro* [± formal, ± human, + sentence-internal expected referent], (b) *esso* [+ formal, **± human**, ± sentence-external expected referent], (c) *essa/esse/ella* [+ formal, ± human, ± sentence-external expected referent], *essi* [+ formal, **+ human,** ± sentence-external expected referent], (d) *egli* [+ formal, **± human,** ± sentence-external expected referent].

2.1.3 Subject position

The studies on subject pronouns presented so far provide important criteria for determining the choice and position of subject pronouns based on the textual, semantic and syntactical features of the sentence level and beyond in (written) Italian. Another important dimension for the position of overt subjects is related to the lexical properties of the involved verbs and their interaction with focus. English and Italian differ here, especially regarding the position of lexical subjects: unlike English, Italian allows postverbal subjects and their availability in the TEU may thus give a further hint regarding the limits of the presumed influence of English. The subject position interacts with verb type and focus type, as shown by CALOI et al. (2018: 2): in wide focus, e.g. in answers to general questions of the type "Che è successo?" ‚What happened?', subjects precede inergative verbs and follow ergative ones. In contrast, subjects in narrow focus, e.g. in answers to questions which focalize the subject like „Chi ha fatto X?" ‚Who did X?', always follow the verb irrespective of its type. The following examples from SCHMITZ & SCHERGER (2025: 16, (7-10)), illustrate these cases:

(6) Wide focus:

(a) Unergatives:	Che è successo? Gianni ha urlato.
	‚What happened? Gianni cried.'
(b) Ergatives:	Che è successo? È arrivato Gianni.
	‚What happened? Gianni arrived.'

(7) Narrow focus:

(a) Unergatives:	Chi ha urlato? Ha urlato Gianni.
	‚Who cried? Gianni cried.'
(b) Ergatives:	Chi è arrivato? È arrivato Gianni.
	‚Who arrived? Gianni arrived.'

The properties of the Italian subject domain will be integrated in the qualitative analysis of the results (cf. section 4.2). The following section briefly presents the state of research in the domain of null and overt subjects in written legal Italian.

2.2 The distribution of null and overt subjects in written legal Italian

As mentioned in the introduction and further detailed in SCHMITZ (2023: 136-137), the analysis of Italian eurolect by MORI (2018b: 227) does not present a complete analysis of null subjects in the project corpus, stating simply a "clear-cut Non-Pro Drop preference in directives". MORI (2018b: 227-228) further claims that in Italian eurolect, subjects mainly occur in sentence beginning position and are generally not omitted, because of the syntactic skeleton of EU master texts written-negotiated-revised in English and French (both Non-Pro Drop languages). For national legislative Italian, MORI (2018b: 228) claims, however, that the inversion V + S (in passive sentences) is allowed when a contextual reason requires it: namely the presence of multiple subjects and the pre-positioning of prepositional phrases containing legal-bureaucratic bundles. The mentioned claims are not proven empirically, however, and MORI (2018b: 228) acknowledges further needs of investigation on the whole annotated corpora; to this end, she simply offers a quantitative indirect cue of this hypothesis by comparing the frequency of subject pronouns, referring to the following listed non-prototypical subjects: *Stati Membri, la presente direttiva, autorità competenti*. These cited items are in fact lexical subjects. MORI (2018b: 228) further claims that the syntactic subjects in Italian Eurolect are mainly re-activated through full/partial repetition of the textual referent or through anaphoric pronouns. She assumes that the observed cross-corpora difference is due to subject ellipsis (zero anaphora) used more in national legal Italian for subject re-activation. The only quantitative analysis of subject use provided in MORI (2018b: 228) relates to the use of the pronoun *esso/essa/essi/esse* in the two Italian data of the corpora used in the *Observing Eurolects* project. Corpus A contains 660 EU directives from the period 1999–2008 while corpus B contains the documents implementing the directives of corpus A at the national level. The raw frequencies (RF) in both corpora are normalised to distribution per 1 million (NF/pmw) to allow comparisons. The project used Log Likelihood (LL) values to measure statistical differences, indicating with +/- the differences from corpus A to B. The following table 1 in SCHMITZ (2023: 136) replicates table 14 from MORI (2018b: 228) revealing high LL values which indicate significant differences in favour of higher rates of the pronoun *esso/essa/essi/esse* in the EU directives.

	Corpus A		LL	Corpus B	
	RF	**NF/pmw**		**RF**	**NF/pmw**
esso	335	233	+99.08	191	97
essa	417	290	+350.42	175	88
essi	1,397	971	+1,320.16	238	120
esse	263	183	+107.49	122	62

Table 1: Use of *esso*/*essa*/*essi*/*esse* in MORI (2018b: table 14) (from SCHMITZ 2023: 136)

The very recent investigation by CANAVESE & MORI (2024) contributes a novel analysis of Italian legal texts in terms of syntactic complexity and accessibility for plain language users. The authors compare Italian versions of EU directives (from the project presented above) with 275 Italian national implementing laws, (1999-2013), 299 Italian domestic legislative acts (1999-2013) and with 160 Italian versions of Swiss federal acts (1993-2006) (cf. CANAVESE & MORI 2024: 119). Their analysis also includes null subject sentences and subject position and yields the following interesting results (CANAVESE & MORI 2024: 138): Null subjects are much less frequent in Swiss legislation (11%) compared to the Italian eurolect (20.82%), the translated documents for the implementation of the EU directives (24.89%) and finally the domestic Italian legislative acts 824.48%). CANAVESE & MORI (2024: 138) further mention a result from their qualitative analysis of the Swiss federal acts, namely that "light anaphoric devices" (as e.g. pronouns) in the German and French versions are often rendered with a null subject structure in Italian. Furthermore, the authors argue that null subjects, which usually refer to the last subject mentioned in the text, do not necessarily create ambiguity or incomprehension, even in the case of a larger distance of more than 50 words from the referent (CANAVESE & TOMBESI 2024: 138). Generally, however, they believe that the presence of another referent mentioned between the null subject and its antecedent increases the likelihood that this intervening referent will be interpreted as the subject of new sentence, thereby slowing down the reading and comprehension process (CANAVESE & TOMBESI 2024: 138).

Finally, TOMBESI (2024) offers a systematic comparative analysis of legal (national) Italian and the Italian eurolect. Her analysis of different grammatical phenomena and linguistic levels is based on a self-composed reference corpus including two corpora: her Corpus IT includes 288 complete legal texts of the Italian state (a mix of laws, directives and executive provisions), while her Corpus EU includes 273 Italian versions of EU directives. Every corpus is subdivided chronologically in five groups of documents (1951-1970, 1971-1989, 1990-2000, 2001-2007, 2008-2018) (cf. TOMBESI 2024: 10). Her manual analysis of subject pronouns in active and passive sentences yields the following results (cf. TOMBESI 2021: 126): relatively speaking, the occurrence of subject pronouns in active sentences ranges from 1.17 and 1.39% in the Corpus IT, while it is slightly higher in the Corpus EU (range of 1.91 to 2.18%). In passive sentences, the similarity of both corpora is even greater: subject pronouns occur in 0.75-1.0 % in the Corpus IT and 1.0-1.16% in the Corpus EU (cf. TOMBESI 2024: 125).

As for the absolute frequency of third person singular and plural subject pronouns, TOMBESI (2024: 126) observes a significative peak (almost 1100 tokens) in the 1990-2000 interval in the Corpus EU, while the number of occurrences in the other time frames span from 275 to approx. 600 tokens. In all periods, however, the absolute frequencies in the Corpus EU are higher than in the Corpus IT, confirming TOMBESI's assumption of an English influence in the translation process. It must be noted, however, that TOMBESI does not present rates of null subjects and they cannot be inferred from the subject pronoun rates presented so far. In qualitative terms, an interesting result concerns the absence of the subject pronouns *lui, lei* and *loro* in subject position, where exclusively *egli, esso/a, essi/e* were used.

2.3 Research questions and hypotheses of the present study

The state of research presented so far includes several claims related to the use and position of (null and overt) subjects in legal Italian. To achieve the aim of the present study (cf. section 1.2), the analysis of CONST and TEU will focus on answering the following research questions (RQ):

RQ 1: How does the distribution of null and overt subjects look like in quantitative and qualitative terms? What properties do the referents of null and overt subject pronouns have in legal Italian?

RQ2: To which extent does the sentence level restrict the option of null subjects in both CONST and TEU? What is the impact of the textual structure requirements governing EU legal acts (cf. section 1.2)?

RQ3: Can the properties of legal language referred to in RQ2 alone explain the observed distribution of null and overt subjects (as well as the position of the latter), i.e. without an impact of language contact? Alternatively, do we find significant (quantitative and qualitative) differences between the Italian CONST and the TEU that reflect a (harmonising) influence of English as the main drafting language?

The analysis will be guided by the following two hypotheses where H1 resumes aspects from the first two RQ and H2 relates more closely to R3:

H1: While specific legal drafting rules for EU secondary legislation restrict the use of null subjects, this claimed absence does not hold for both the Italian CONST and TEU version. Wherever possible, the individual legal culture and linguistic properties will be decisive, i.e. the fundamental grammatical properties of Italian (null subject property, antecedent preferences and overt subject placement options) will determine the grammar of Italian legal texts at the syntax and discourse level, allowing for null subjects within the boundaries of the specific needs of clarity of legal language.

H2: H1 does not exclude a harmonizing influence of English as main drafting language. This influence will, however, not lead to a total avoidance of null subjects but rather to an increased use of weak pronouns (instead of *pro*) and a reduction of post verbal

subjects in the TEU. Furthermore, the distance between (overt or null) pronouns and their referent will be reduced to the sentence level in the TEU.

3 Methodology

3.1 The Data

Like in SCHMITZ (2023), the analysis in the present paper is restricted to the subject expression in the main text of TEU and the functionally analogous parts of the Italian Constitution (CONST) based on its updated and annotated version from 2019)[6], i.e., those parts indicating the fundamental principles and the organization of EU and Italy, respectively. The catalogues of fundamental human rights in both CONST and CFR are less interesting for the analysis of subject omissions since formulations of fundamental rights typically contain overt strong pronouns, particularly *ognuno* 'everybody' (related to rights, permissions) and *nessuno* 'nobody' (related to prohibitions). Unlike the preliminary analysis in SCHMITZ (2023), the present study will provide a more detailed quantitative and qualitative analysis, including, among other aspects, the position of overt subjects and the type of subject.

To establish the comparability, the structure of both texts and the equivalence will be exposed in the following. The Italian CONST is composed of 139 articles (five of which were later repealed) and arranged into the following main parts: *Principi Fondamentali* 'Fundamental Principles' (articles 1–12); Part I *Diritti e Doveri dei Cittadini* 'Rights and Duties of the Citizens' (articles 13–54; excluded from the analysis for the reasons mentioned above), Part II *Ordinamento della Repubblica* 'Organisation of the Republic' (articles 55–139), followed finally by the *Disposizioni transitorie e finali* 'Transitory and Final Provisions', set apart by separate numbering. The Italian version of TEU is structured as follows: the treaty text following the preamble is divided into six parts or titles, namely Title 1 *Common Provisions* (articles 1–8), Title 2 *Provisions on democratic principles* (articles 9–12), Title 3 *Provisions on the institutions* (articles 13–19), Title 4 *Provisions on enhanced cooperations* (article 20), Title 5 *General provisions on the Union's external action and specific provisions on the Common Foreign and Security Policy* (articles 21–46) and finally, Title 6 *Final provisions* (articles 47–55). Table 2 in section 3.2 indicates the equivalences and combines this information with further relevant information for the subject analysis.

3.2 The Analysis

To start, the following text parts and sentence types were excluded from the analysis: a) the preambles, the formula of promulgation and the formula of subscription typically listing all subscribing legal representatives since these parts disallow null subjects, b) presentational

[6] During the analysis of this document, a comparison with the original version from 1947 was effectuated which revealed a very high degree of linguistic fidelity in all text parts not affected by amendments.

constructions of the type "syntagma X è" 'X is' followed by an adjectival, nominal or prepositional syntagma since they do not allow a free choice of subject type. All remaining clauses have been manually analysed and coded for the finite verbs and their (null/overt) subjects in terms of number (while person is always 3rd), type of overt subject, type of sentence matrix/subordinate clause), and voice (active/passive). The resulting analysable total numbers of finite verb forms (FV) are: CONST 509 (excl. 170 tokens of *Parte* I) and TEU 654 tokens. For these tokens, Table 3 will provide information on the number and percentage of finite main (MC) and subordinate clauses (SC) as well as finite active (AC) and passive clauses (PC) in each part.

CONST	*Principi Fondamentali* Articles 1–12	*Parte II* Articles 55–139	*Disposizioni* (17 provisions)	***Total***
	FV: 39 MC: 30 SC: 9 AC: 34 PC: 5	FV: 403 MC: 298 SC: 105 AC: 281 PC:122	FV: 67 MC: 46 SC: 21 AC: 42 PC: 25	FV: 509 (100%) MC: 374 (73.5%) SC: 135 (26.5%) AC: 357 (70.1%) PC: 152 (29.9%)
TEU	*Titles 1 and 2* Articles 1–12	*Titles 3-5* Articles 13–46	*Title 6* Articles 47–55	***Total***
	FV: 102 MC: 76 SC: 26 AC: 89 PC: 13	FV: 465 MC: 306 SC: 159 AC: 393 PC: 72	FV: 87 MC: 52 SC: 35 AC: 65 PC: 22	FV: 654 (100%) MC: 434 (66.4%) SC: 220 (33.6%) AC: 547 (83.6%) PC: 107 (16.4%)

Table 2: Functional equivalences and morphosyntactic information for the analysed text parts of CONST and TEU

Table 2 highlights some interesting differences in the general syntactic structure of the two primary law texts. Overall, the analysed parts of CONST contain a slightly lower number of subordinate clauses in relation to main clauses than TEU; in contrast, the percentage of passive clauses is much higher in CONST than in TEU. Looking more closely at the structure of complex sentences, we mostly find several coordinated main clauses with one or more embedded subordinate clauses. Typically, these sentences start with a main clause, but we also find a relatively small number of complex sentences with an initial subordinate clause including a finite verb: The TEU contains 20 such sentences (20/654 = 3.1%) and CONST only 12 (12/509 = 2,4%). In both texts, the subordinate clauses typically introduce conditional sentences (*Se*/*Qualora*… 'If…'), temporal sentences (*Quando*/*Allorché*… 'When/As soon as/Once…') or denote the validity of a specific state of affairs at a given moment (*Finché*/*Fino a quando*… 'Till/Until further notice…'). In both legal texts, the initial subordinate clauses and the following main clauses contain different overt lexical subjects. Anaphors do not occur in

this type of complex sentence. Finally, the numbers of tokens in all detailed analyses are generally low; the quantitative analysis will therefore not use inferential statistics.

4 Results

The presentation of the results is organized in two steps, providing a general quantitative picture of the distribution of null and overt subjects in Section 4.1 before delving into qualitative aspects and providing examples in Section 4.2.

4.1 Quantitative analysis

4.1.1 Distribution of null vs. overt subjects

The presentation of the results starts with the overall distribution of null vs. overt (pronominal/lexical) subjects (%) in the analyzed parts of the two primary legal texts illustrated in Figure 1:

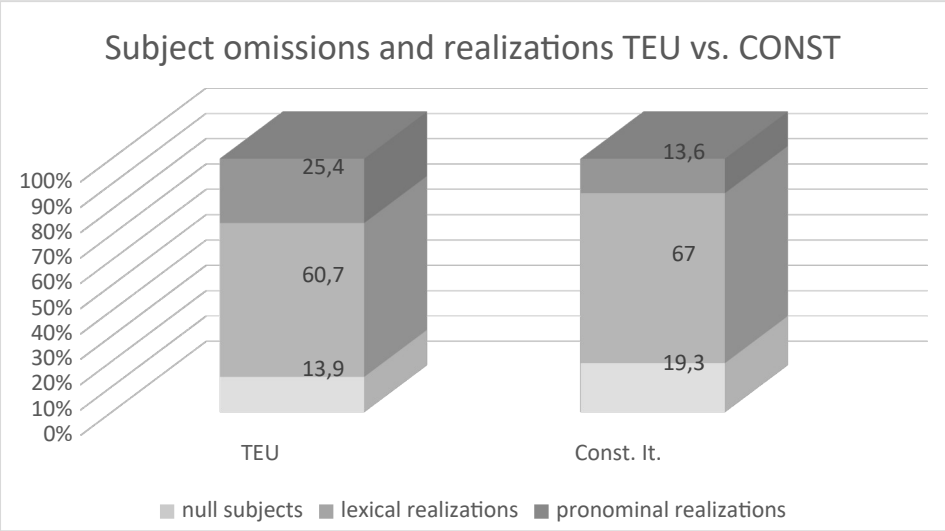


Figure 1: Null and overt subjects in TEU and CONST

Figure 1 shows a difference between the amounts of pronouns in the two texts of approx. 12%. The rates of both the null subjects and the lexical subjects in the TEU are slightly lower (approx. 6%) than in CONST in favour of overt subject pronouns. The next step involves looking more closely at the distribution of null and overt subjects in relation to the different clause types. The distribution across main vs. subordinate clauses is shown in Figure 2.

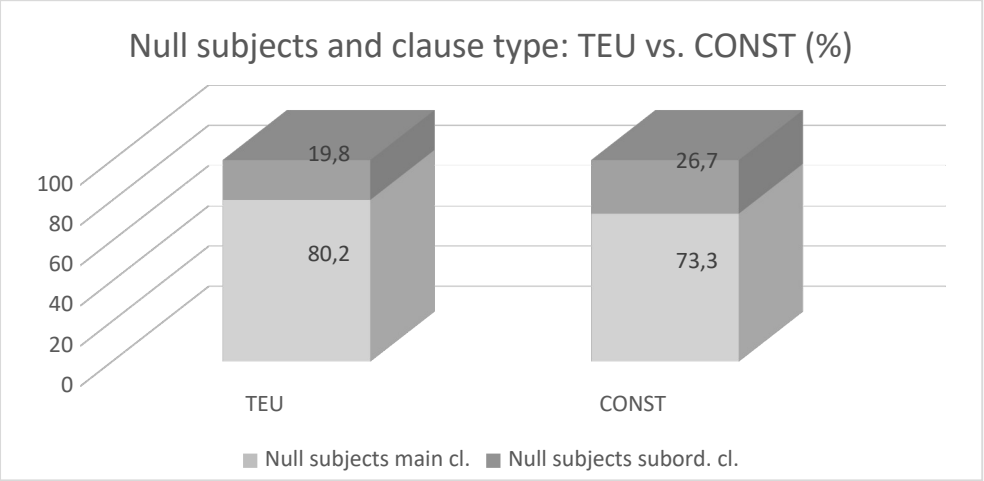


Figure 2: Null and overt subjects in main and subordinate clauses in TEU and CONST

Figure 2 shows that, although the percentages of null subjects is overall higher in CONST, most null subjects in both primary law texts are clearly found in main clauses and this preference is even more pronounced in TEU. As for the distribution of null and overt subjects across active and passive clauses, Figure 3 again shows a clear preference in both legal texts, namely the occurrence of null subjects in active clauses:

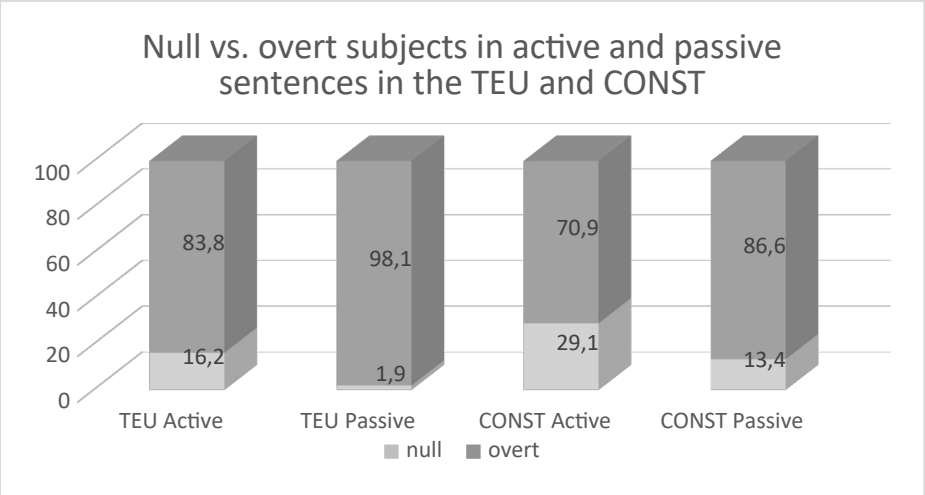


Figure 3: Null and overt subjects in active and passive clauses in TEU and CONST

Summarizing thus far, null subjects are clearly disfavoured in both Italian primary law texts in relation to overt subjects (where lexical forms dominate). Both texts also share a preference for null subjects to occur in main and active clauses. The next section will look more closely into the form and position of overt subjects.

4.1.2 Types and positions of overt subjects

The different types of pronominal and lexical subjects occurring in the analyzed parts of TEU and CONST include the following types: lexical subjects (cf. Figure 1 above) typically occur

as NP (i.e. proper nouns or other nouns without determiner) or DP. The array of pronominal subjects is larger and includes personal pronouns, *si impersonale*, demonstrative pronouns, relative pronouns and other pronouns (including e.g. indefinite pronouns like *nessuno* 'nobody' and quantifiers like *tutti* 'all'). Figure 4 shows their distribution in the two texts in absolute numbers due to their overall low frequency; the pronoun types indicated serve as examples of each category:

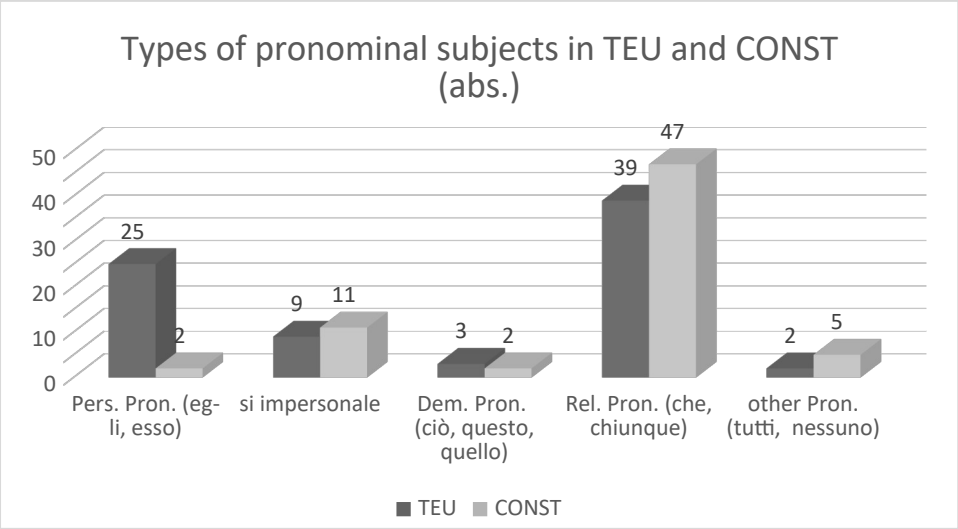


Figure 4: Distribution of subject pronouns in TEU and CONST

The comparison of the pronoun distribution in Figure 4 reveals a clear preference for the *egli/esso* series in the TEU while all other pronoun types are present but to a lesser degree than in CONST. We can further observe a preference for relative pronouns in subject function over all other subject pronouns, which might be explained with the structure of legal texts. In contrast, demonstrative and other pronouns are marginal in both documents. Finally, both documents do not contain any instance of the neo-standard subject pronouns *lui, lei* and *loro*. *Lei* is totally absent while *lui* exclusively appears in oblique PP. *Loro* is used either in an oblique PP or used as possessive element in DPs (e.g. Art. 1 para. 1 TEU […] *i loro obiettivi comuni* '[…] their common objectives').

Finally, the position of the overt subjects is illustrated in Figure 5 which shows a clear preference for preverbal subjects in both primary law texts, which is particularly pronounced in the TEU where postverbal (and always lexical) subjects are rare (examples will be provided in Section 4.2). This scarcity could be explained by an influence from the quite strict SVO order in English.

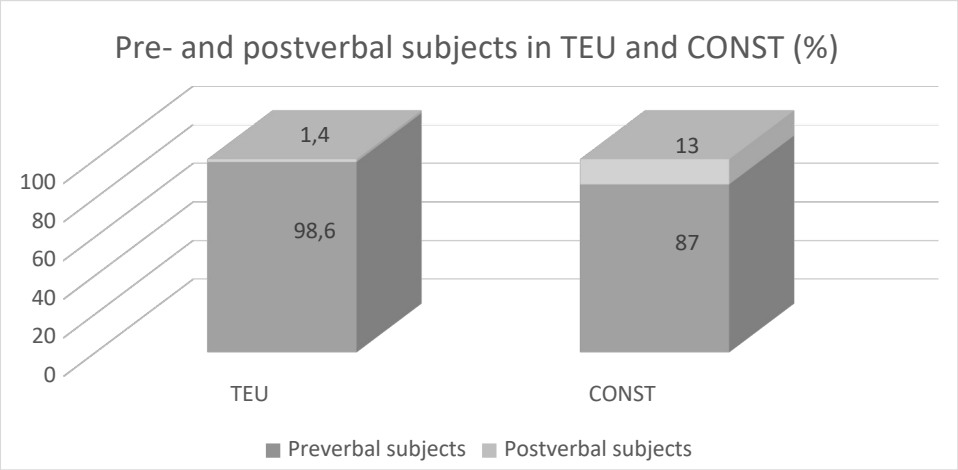


Figure 5: Distribution of pre- and postverbal subjects TEU and CONST

Summarizing the findings related to overt subjects, they are predominantly lexical and preverbal in both the TEU and CONST. The main differences concern the increased number of personal pronouns (presumably at the expense of null subjects) as well as the almost inexistent number of postverbal subjects in the TEU. To evaluate the role of semantic, syntactic and pragmatic factors in the observed distributions of null and overt subjects as well as the presumed influence of English in the Italian version of the TEU, the following section will deal in more detail with the properties and position of the referents of the null subjects.

4.2 Qualitative analysis

4.2.1 Null subjects and their referents

The quantitative analysis in section 4.1.1 has shown a preference for the placement of null subjects in main clauses. This finding raises the following questions: where are the referents and which unit in legal texts can be considered as crucial discourse unit? Before considering the possible options, it is necessary to point out that the two primary law texts are organized differently. While Articles in CONST generally consist of various sentences, which coincide with a legal paragraph or with an article's paragraph, except for very long articles, TEU articles typically contain paragraphs with various sentences (cf. also section 1.1). The possible options for the position of the referents are: (a) referent in a previous paragraph or article, (b) referent in a previous sentence of the same paragraph (TEU) or article (CONST) or (c) referent in the same (complex) sentence of the same paragraph (TEU) or article (CONST).

Looking more closely first into CONST (with an absolute number of 103 null subjects in the analyzed parts), we find that none of them has its referent in a previous article. Instead, all referents of null subjects stand in the same article (or para. in very long articles, e.g. Art. 119 CONST). In fact, 62 (60.2%) of the null subjects have their referent in the same sentence; independently from the even distribution of 52 (50.5%) [+ human] and 51 (49.5%) [- human] referents. Example (8) presents Article 87 CONST which deals with the tasks of the President

of the Repubblica (drawn from original 1947 version, unchanged to date) and provides several examples of null subjects with [+ human] referents within the same article, where the first main clause introduces the President of the Italian Republic and all further clauses refer to him and his tasks (null subjects indicated by **Ø**):

(8) Art. 87 (CONST):
Il Presidente della Repubblica è il capo dello Stato e **Ø** rappresenta l'unità nazionale.
Ø Può inviare messaggi alle Camere.
Ø Indice le elezioni delle nuove Camere e ne fissa la prima riunione.
ØAutorizza la presentazione alle Camere dei disegni di legge di iniziativa del Governo.
Ø Promulga le leggi ed ◎ emana i decreti aventi valore di legge e i regolamenti.
Ø Indice il *referendum* popolare nei casi previsti dalla Costituzione.
Ø Nomina, nei casi indicati dalla legge, i funzionari dello Stato.
Ø Accredita e **Ø** riceve i rappresentanti diplomatici, **Ø** ratifica i trattati internazionali, previa, quando **Ø** occorra, l'autorizzazione delle Camere.
Ø Ha il comando delle Forze armate, **Ø** presiede il Consiglio supremo di difesa costituito secondo la legge, **Ø** dichiara lo stato di guerra deliberato dalle Camere.
Ø Presiede il Consiglio superiore della magistratura.
Ø Può concedere grazia e commutare le pene.
Ø Conferisce le onorificenze della Repubblica.

Example (9) presents an extract of Article 119 (para. 1-2) CONST mentioned above which includes more than one sentence. Although all null subjects refer to „I comuni… le Regioni", they only appear within the paragraph (drawn from original 1947 version, unchanged to date) and the referent is repeated at the beginning of the second para.:

(9) Art. 119, para. 1-2 CONST
I Comuni, le Province, le Città metropolitane e le Regioni hanno autonomia finanziaria di entrata e di spesa, nel rispetto dell'equilibrio dei relativi bilanci, e **Ø** concorrono ad assicurare l'osservanza dei vincoli economici e finanziari derivanti dall'ordinamento dell'Unione europea.

I Comuni, le Province, le Città metropolitane e le Regioni hanno risorse autonome. **Ø** Stabiliscono e applicano tributi ed entrate propri, in armonia con la Costituzione e secondo i principi di coordinamento della finanza pubblica e del sistema tributario. **Ø** Dispongono di compartecipazioni al gettito di tributi erariali riferibile al loro territorio.

Turning to the TEU (with an absolute number of 91 null subjects), we find no referent in a previous paragraph; instead, all referents of null subjects are in the same paragraph as central unit within articles. Furthermore, a majority, namely 64 (70.3%), of the null subjects have their referent in the same sentence, independently of their [± human] reference. In contrast to CONST, we find an uneven distribution with respect to this feature in the TEU: Only 17 (18.7%) of the null subjects have [+ human] referents. The following example (10) containing para, 1 of Art. 15 TEU in Italian (with its English counterpart) shows a similar

pattern to that of the previous two examples in the Italian eurolect, whereas the English version allows only null subjects in coordination structures):

(10) IT: **Il Consiglio europeo** dà all'Unione gli impulsi necessari al suo sviluppo e **Ø** ne definisce gli orientamenti e le priorità politiche generali. **Ø** Non esercita funzioni legislative.

EN: **The European Council** shall provide the Union with the necessary impetus for its development and **Ø** shall define the general political directions and priorities thereof. **It** shall not exercise legislative functions.

Summarizing so far, the previous examples have clearly shown that both legal varieties of Italian allow for null subjects. The distance between null subject and its referent looks shorter in the TEU where the main reference space is the paragraph within an article whereas the CONST does not always provide an internal paragraph structure, often grouping individual sentences instead. In that case, the level of the article is crucial and limits the possible distance between antecedent and null subject (cf. e.g. Art. 8 CONST in example 8).

4.2.2 *Esso/egli* subjects and their referents

The quantitative analysis in Section 4.1.2 has revealed a substantial increase in the use of the personal pronouns of the *esso/egli* series. In absolute numbers, CONST contains one token of *egli* (i.e. [masc. sing.]) and two tokens of the *esso* type (1 *essi* [masc. plur.] and one *esse* [fem. plur.]), but only *essi* is a subject. The referents of the subjects *egli* and *essi* are located in the same sentence within the Articles 86, para. 1 (i.e. the *Presidente della Repubblica* 'President of the Republic'), and the two sentences of Art. 88 (i.e. the *ultimi sei mesi del suo mandato* 'last six months of his mandate') as shown in the examples (11) and (12):

(11) Art. 86, para. 1 CONST

Le funzioni del **Presidente della Repubblica**, in ogni caso che **egli** non possa adempierle, sono esercitate dal Presidente del Senato.

(12) Art. 88 CONST

Il Presidente della Repubblica può, sentiti i loro Presidenti, sciogliere le Camere o anche una sola di esse. Non può esercitare tale facoltà ne**gli ultimi sei mesi del suo mandato**, salvo che **essi** coincidano in tutto o in parte con gli ultimi sei mesi della legislatura.

These two examples are also interesting "test cases" for the claims by CORDIN (2021) (cf. section 2.1.2): in both cases the distance between the weak subject pronouns and their referent is rather short. For these cases, CORDIN (2021) suggests to apply the concept of syntactic independence when the sentences and the morphosyntactic features are different. In both examples, the weak pronouns are the subject of a subordinate clause and refer to complements in prepositional phrase in the main clause which do not agree in number with

the subject of the main clause (*le funzioni del Presidente della Repubblica* 'the functions of the President of the Republic' in ex. 11 and pro in ex. 12). In both cases, the weak pronoun could thus be omitted but the reader/writer might process the document with more effort. The presence of overt weak pronouns substantially eases comprehension, as the referents of *egli* and *essi* being embedded within a prepositional phrase are not highly prominent. In this context, prominence seems to outweigh distance and agreement features.

Proceeding to TEU, it contrasts with CONST providing 23 tokens of the *esso* type and two tokens of *egli.* All referents of these 25 personal pronouns stand in the same paragraph; out of these, three (3/25, 12%) are located also in the same sentence, while the majority (22/25, 88%) are positioned in the previous sentence. The referents of the two *egli* tokens is [+ human], namely *L'alto rappresentante dell'Unione per gli affari esteri e la politica di sicurezza* ,The High Representative of the Union for Foreign Affairs and Security Policy'. The two tokens are located in Art. 18, para. 2 and Art. 36 TEU. The first token is particularly interesting in comparison with its English counterpart, as shown in example (13):

(13) Art. 18, para. 2 TEU

IT: **L'alto rappresentante** guida la politica estera e di sicurezza comune dell'Unione. **Ø** Contribuisce con le sue proposte all'elaborazione di detta politica e la attua in qualità di mandatario del Consiglio. **Egli** agisce allo stesso modo per quanto riguarda la politica di sicurezza e di difesa comune.

EN: **The High Representative** shall conduct the Union's common foreign and security policy. **He** shall contribute by his proposals to the development of that policy, which **he** shall carry out as mandated by the Council. **The same** shall apply to the common security and defence policy.

The use of *egli* in the Italian version in the third sentence is rather unexpected since we have a case of topic continuity of the referent, namely the High Representative. It can be, however, be explained by an intervening possible referent, the *Consiglio* although less prominent given its embedding in a prepositional phrase – the proximity could, however, influence the perception of the reader/writer and therefore favour the choice of a disambiguating weak subject pronoun. Comparing the Italian version with the English one, another unexpected referential link emerges in the English text: "the same" in the third sentence does not refer to the High Representative but to his task as expressed in the previous sentence.

A final interesting result is that all 23 tokens of the *esso*-type in the TEU appear in preverbal positions in their clauses and refer exclusively to non-human entities, either EU organs, decisions or fixed political strategies of the EU. They split up as follows: seven tokens of *esso*, nine tokens of *essa* (out of which five tokens refer to *L'unione* 'The Union'), three tokens of *essi* and four tokens of *esse*.

4.2.3 Postverbal subjects in the TEU

While the preceding sections analysed the properties of the null and overt subjects in the two primary law texts, this part of the qualitative analysis answers to the claims by MORI (2018b: 228) regarding the position of overt subjects in the Italian eurolect (cf. Section 2.2), briefly repeated here. Subjects mainly occur in sentence-initial position and are generally not omitted, while national legislative Italian allows the inversion V + S (in passive sentences) when a contextual reason requires it: namely the presence of multiple subjects and the pre-positioning of prepositional phrases containing legal-bureaucratic bundles. The quantitative analysis in Section 4.1.2 has confirmed the rarity of postverbal subjects in the TEU (8 tokens/563 clauses with a finite verb, 1.4%) opposed to a higher, but still minoritarian, number in CONST (71 tokens/548 clauses with a finite verb, 13%).

From a qualitative perspective, the first observation to be made is that two of the eight tokens are identical. We find two occurrences of the sentence *È esclusa l'adozione di atti legislativi*. 'The adoption of legislative acts shall be excluded.' (Art. 24, para. 1, s. 4 and Art. 31 para. 1, s. 2 TEU). As for the subject type and clause type, the eight tokens are of the NP/DP type and mostly complex. They typically include at least an AdvP or AP adjective but also PPs, the only exception being the DP *l'Unione* 'the Union' in Art. 49, para. 2, s. 1 TEU (cf. example (17) below). While they occur evenly distributed across main and subordinate clauses, their distribution across active and passive clauses is very uneven. Seven out of the eight tokens appear in passive clauses. The following examples (14-19) present the remaining six non-doubled tokens using their English counterpart as translation. Example (14) presents the only non-passive token of a postverbal subject in a subordinate clause occurring with an intransitive verb (*partecipare* 'participate'); the pertinent clause may be read as an instance of narrow focus, i.e. dealing with who may be involved in the cooperation described, which implies that the fact that *partecipare* is not an ergative verb is irrelevant here (cf. also Section 2.2).

(14) Art. 20, para 2, s. 1 TEU

IT: La decisione che autorizza una cooperazione rafforzata è adottata dal Consiglio in ultima istanza, qualora esso stabilisca che gli obiettivi ricercati da detta cooperazione non possono essere conseguiti entro un termine ragionevole dall'Unione nel suo insieme, e a condizione che vi partecipino **almeno nove Stati membri.**

EN: The decision authorising enhanced cooperation shall be adopted by the Council as a last resort, when it has established that the objectives of such cooperation cannot be attained within a reasonable period by the Union as a whole, and provided that **at least nine Member States** participate in it.

The following examples (15-17) include the further postverbal subject tokens in subordinate clauses while examples (18-20) show the remaining tokens in main clauses; all of them include transitive verbs in periphrastic passive constructions:

(15) Art. 3, para. 2 TEU:

IT: L'Unione offre ai suoi cittadini uno spazio di libertà, sicurezza e giustizia senza frontiere interne, in cui sia assicurata **la libera circolazione delle persone** insieme a misure appropriate per quanto concerne i controlli alle frontiere esterne, l'asilo, l'immigrazione, la prevenzione della criminalità e la lotta contro quest'ultima.

EN: The Union shall offer its citizens an area of freedom, security and justice without internal frontiers, in which **the free movement of persons** is ensured in conjunction with appropriate measures with respect to external border controls, asylum, immigration and the prevention and combating of crime.

(16) Art. 7, para. 3, s. 1 TEU:

IT: Qualora sia stata effettuata **la constatazione di cui al paragrafo 2**, il Consiglio, deliberando a maggioranza qualificata, può decidere di sospendere alcuni dei diritti derivanti allo Stato membro in questione dall'applicazione dei trattati, compresi i diritti di voto del rappresentante del governo di tale Stato membro in seno al Consiglio.

EN: Where **a determination under paragraph 2** has been made, the Council, acting by a qualified majority, may decide to suspend certain of the rights deriving from the application of the Treaties to the Member State in question, including the voting rights of the representative of the government of that Member State in the Council.

(17) Art. 49, para. 2, s. 1 TEU:

IT: Le condizioni per l'ammissione e gli adattamenti dei trattati su cui è fondata **l'Unione**, da essa determinati, formano l'oggetto di un accordo tra gli Stati membri e lo Stato richiedente.

EN: The conditions of admission and the adjustments to the Treaties on which **the Union** is founded, which such admission entails, shall be the subject of an agreement between the Member States and the applicant State.

(18) Art. 14, para. 2, s. 3 TEU:

IT: A nessuno Stato membro sono assegnati **più di novantasei seggi**.

EN: No Member State shall be allocated **more than ninety-six seats.**

(19) Art. 45, para. 2, s. 4 TEU:

IT: Nell'ambito dell'Agenzia sono costituiti **gruppi specifici** che riuniscono gli Stati membri impegnati in progetti congiunti.

EN: **Specific groups** shall be set up within the Agency bringing together Member States engaged in joint projects.

In summary, the qualitative aspects of the few instances of postverbal subjects in the Italian version of the TEU are almost all complex. They occur mainly in passive clauses with different transitive verbs, only once in an active clause with an intransitive verb. Interestingly, we also find one token (cf. ex. (18)), where the English version provides a

subject in postverbal position. In all other cases the Italian version differs from the almost strict preverbal English subjects which can be interpreted as maintenance of typical Italian word order and processing preferences for focussed subjects.

5. Discussion

In a first step, the results presented in the previous section will be discussed in the light of to the hypotheses (cf. Section 2.4), starting with H1.

Both the quantitative and the qualitative analysis provide ample evidence for the use of null subjects and pronominal anaphoric reference in the selected Italian statutory texts beyond the sentence level. This clearly shows that essential properties of the Italian language are also maintained in the legal context, thereby confirming H1. The special needs of legal language are observed, among other means, by restricting the distance between anaphors and their referents to the level of the article (CONST) or paragraph (TEU) as pertinent text unit for anaphora resolution. The alleged avoidance of grammatical ties across sentence boundaries in legal texts, based on observations on English national legal documents by HILTUNEN (2012), does not hold for the Italian legal varieties.

MORI's (2018b) assumption concerning the impact of English on the Italian eurolect, has been addressed by H2, which is confirmed insofar as a clear increase in the use of subject pronouns (particularly of the *esso*-type) at the expense of possible null subjects has been observed in the TEU. Furthermore, the very low number of postverbal subjects in the TEU can be interpreted as evidence of an influence of English. H2 must be considered contradicted with respect to the restricting effect of the influence on anaphoric reference to the sentence level. As discussed above, TEU provides anaphoric reference at least at paragraph level. The difference in structural level (paragraph/article level) is an effect of the EU drafting rules presented in section 1 and not necessarily associated with an influence from English (although, of course, also the British legal tradition has been integrated).

A direct comparison of the quantitative results with those of the available studies in the legal linguistic literature (cf. section 2.1.1) is not possible since TOMBESI (2024) does not provide overall null subject rates and both the size of the database and the legal text genre in the data from CANAVESI and MORI (2024) are very different (directives in the different legal Italian varieties, and in the case of the Swiss data, a mix of documents). Abstracting away from the legal genre, the null subject rates in the Swiss data are quite close to those in the TEU (11% to approx. 13%; cf. also CANAVESE and MORI 2024: 138 for such a comparison with SCHMITZ 2023).

The present detailed analysis aimed to go beyond quantitative results by incorporating qualitative aspects such as subject type and position. In doing so, it shows that, overall, the individual properties of each language are preserved in legal discourse, including the phenomena at the syntax-pragmatics interface.

6. Concluding remarks

The present study has been dedicated to a strongly understudied phenomenon in legal language, especially in the Italian legal varieties, namely the realization of null subjects as one of the grammatical properties of Italian. The quantitative and qualitative analysis of two primary law texts, namely the complete Italian version of the Treaty of the European Union (TEU) and the functionally equivalent sections of the Italian Constitution (CONST) shows that claims on grammatical structures in legal texts based on one language (here: English) do not generally hold for another (here: Italian). This result implies that findings from English cannot be generalized to Italian legal texts which have their own tradition and clearly reflect the properties of Italian syntax. Contrary to claims by HILTUNEN (2012), the sentence level is not the crucial discourse unit for anaphoric reference for null and pronominal subjects, but the article level in CONST and the paragraph level in TEU. Claims expressed especially for the Italian eurolect by MORI (2018b) have been proven to hold only partially. A confirmation of the presumed influence from English in the Italian eurolect can be seen in the increase of subject pronouns and the rarity of postverbal subjects in the TEU. The reduction of the distance between anaphors and their referent to paragraph level in the TEU is most likely an artefact of the different text organization in CONST and TEU and not a linguistic property of the Italian eurolect. The notion of prominence of the referent has been shown to play a role in the choice of weak subject pronouns over null subjects which may facilitate the processing from the perspective of the reader/writer.

The present study could only allude to the possible impact of individual and institutional multilingualism on processing preferences among individuals involved in the collective task of translating EU legislation by way of integrating substantial psycholinguistic work in the background. To properly assess its effects in the complex interaction of legal and linguistic factors in the domain of subject expression in the legal varieties of null-subject Treaty languages, much more and systematic connection between psycholinguistics and legal linguistics as well as empirical research is required. Furthermore, and obviously, the results of the presented analysis are restricted to a limited set of texts and a small number of verb tokens and to one legal genre within the huge number of legal documents written in Italian both at national and EU level. The extension of the analysis to more texts and different legal genres is necessary. Furthermore, a detailed analysis of the Spanish TEU version and its functionally equivalent parts in the Spanish Constitution is highly desired since these two Romance null subject languages show considerable micro variation in the subject domain and different effects in reference tracking (cf. the psycholinguistic research presented in Section 2.1 and references to further comparative work cited therein.

7. References

AUER, Peter (2017). The neo-standard of Italy and elsewhere in Europe. In: Massimo CERRUTI/Claudia CROCCO/Stefania. MARZO (eds.) "*Towards a New Standard. Theoretical and Empirical Studies on the Restandardization of Italian*". Boston/Berlin: De Gruyter-Mouton, 356-375.

BERRUTO, Gaetano (2012). "*Sociolinguistica dell'italiano contemporaneo*". 2nd edition. Roma: Carocci.

BURR, Isolde (2013). "Article 55 [Languages and Deposit of the Treaty]". In: BLANKE, Hermann-Josef/MANGIAMELI, Stelio (eds.). *The Treaty on European Union (TEU).* Berlin, Heidelberg: Springer-Verlag, 1461-1525.

CALOI, Irene/BELLETTI, Adriana/POLETTO, Cecila (2018). "Multilingual competence influences answering strategies in Italian-German speakers". *Frontiers in Psychology*, 9. DOI 10.3389/fpsyg.2018.01971.

CANAVESE, Paolo/MORI, Laura (2024). "Testing the hypothesis of 'translation as a catalyst for plain legislation' on the syntactic level: a comparison of different varieties of legislative Italian". *Comparative Legilinguistics* 58, 113-150. DOI: 10.14746/cl.2024.58.3

CARDINALETTI, Anna/STARKE, Michal (1999). "The typology of structural deficiency: A case study of the three classes of pronouns". In VAN RIEMSDIJK, Hendrik (ed.), *Clitics in the languages of Europe*. Berlin: Mouton de Gruyter, 145–233.

CARDINALETTI, Anna (2021). "I pronomi deboli nel sistema pronominale dell'italiano". In: SCHNEIDER, Stefan/SALVI, Giampaolo/GARZONIO, Jacopo (eds.), *La descrizione grammaticale dell'italiano. Parte 1 – Problemi generali e italiano moderno*. Special issue of *Storie e Linguaggi*, Vol. 7(2), 131-151.

CARMINATI, Maria Nella (2002). "*The processing of Italian subject pronouns". Doctoral dissertation.* University of Massachusetts at Amherst.

CERRUTI, Massimo/CROCCO, Claudia/MARZO, Stefania (2017). Introduction. In: Massimo CERRUTI/Claudia CROCCO/Stefania. MARZO (eds.) "*Towards a New Standard. Theoretical and Empirical Studies on the Restandardization of Italian*". Boston/Berlin: De Gruyter-Mouton, 3-28.

CHIERCHIA, Gennaro (1995). "Dynamics of meaning: anaphora, presupposition, and the theory of grammar". Chicago: University of Chicago Press.

CONTEMORI, Carla/DI DOMENICO, Elisa (2021). "Microvariation in the division of labor between null- and overt-subject pronouns: the case of Italian and Spanish". *Applied Psycholinguistics* 42, 997-1028.

CORDIN, Patrizia (2021). "Soggetti pronominali deboli". In: SCHNEIDER, Stefan/SALVI, Giampaolo/GARZONIO, Jacopo (eds.), *La descrizione grammaticale dell'italiano. Parte 1 – Problemi generali e italiano moderno*. Special issue of *Storie e Linguaggi*, Vol. 7(2), 153-175.

EUROPEAN PARLIAMENT/COUNCIL OF THE EUROPEAN UNION/EUROPEAN COMMISSION (2015). "*Joint Practical Guide of the European Parliament, the Council and the Commission for persons involved in the drafting of European Union legislation".* All language versions: www.consilium.europa.eu/en/council-eu/decision-making/ordinary-legislative-procedure. (download 11.08.2024).

EUROPEAN PARLIAMENT/COUNCIL OF THE EUROPEAN UNION/EUROPEAN COMMISSION (2020). "*Joint Handbook for the presentation and drafting of acts subject to the ordinary legislative"*. https://www.consilium.europa.eu/media/47929/joint-handbook-en-31_july-2020_clean_def.pdf (English, Italian) (download 11.08.2024).

GOTTI, Maurizio (2012). "Text and Genre". In: SOLAN, Laurence M/TIERSMA, Peter M. (eds.), *The Oxford Handbook of Language and Law*. Oxford: Oxford University Press, 52-66.

HEINEMANN, Ellen (2019). "The production process of EU directives and their transposition into national law – additional considerations reviewing "Observing Eurolects. Corpus analysis of linguistic variation in EU law" by MORI, Laura (ed.). Amsterdam/Philadelphia, John Benjamins 2018". *ZERL Zeitschrift für Europäische Rechts-linguistik*, 1-59. Urn: nbn: de: hbz:38-103146.

HILTUNEN, Risto (2012). "The Grammar and Structure of Legal Texts". In: SOLAN, Laurence M/TIERSMA, Peter M. (eds.), *The Oxford Handbook of Language and Law*. Oxford: Oxford University Press, 39-51.

KOEHN, Philipp (2005). "Europarl: A Parallel Corpus for Statistical Machine Translation". *Proceedings of the 10th Machine Translation Summit* (MT Summit X).

LOZANO, Cristóbal/QUESADA, Teresa (2023). What corpus data reveal about the Position of Antecedent Strategy: anaphora resolution in Spanish monolinguals and L1 English-L2 Spanish bilinguals. *Frontiers in Psychology* 14, 1246710. DOI: 10.3389/fpsyg 2023.1246710.

MAIDEN, Martin (1995). "*A Linguistic History of Italian*". Harlow, Essex: Longman.

MATTILA, Heikki E. (2016). "*Comparative Legal Linguistics. Language of Law, Latin and Modern Lingua Francas*." 2nd edition. London/New York: Routledge.

MATTILA, Heikki E. (2019). "Review of MORI, Laura (ed.), Observing Eurolects: Corpus analysis of linguistic variation in EU law. Amsterdam/Philadelphia: John Benjamins Publishing Company 2018". *Comparative Legilinguistics,* Vol. 38, 83-94. DOI: http://dx.doi.org/10.14746/cl.2019.38.4.

MATTISSEN, Johanna (2022). "Reference-tracking in multilingualen europäischen Rechtstexten". In: WIENEN, Ursula/ REICHMANN, Tinka / SERGO, Laura (eds.): *Syntax in Fachkommunikation.* Berlin: Frank und Timme, 463–500.

MATTISSEN, Johanna (2023). "Demonstrativa in mehrsprachig verbindlichen Rechtsakten der EU". In: DUJMIC, Tatjana/MALLON, Denise (eds.), *7. Fachkonferenz Sprache und Recht Rechtsvergleichung in der Praxis. Tagungsband der 7. Fachkonferenz Sprache und Recht des Bundesverbands der Dolmetscher und Übersetzer e.V. (BDÜ), Berlin, 16./17. September 2022*. Berlin: BDÜ Fachverlag, 139-149.

MAYOL, Laia (2010). "Contrastive pronouns in null-subject Romance languages". *Lingua* 120, 2497–2514.

MORI, Laura (2003). "L'euroletto: Genesi e sviluppo dell'italiano comunitario". In: VALENTINI, Ada/Molinelli, Piera/CUZZOLIN, Pierluigi/BERNINI, Giuliano (eds.), *Ecologia linguistica. Atti del XXXVI Convegno internazionale di Studi della Societa di Linguistica italiana*. Roma: Bulzoni, 473–492.

MORI, Laura (2018a). "Introduction. The Eurolect Observatory Project". In MORI, Laura (ed.). *Observing Eurolects. Corpus analysis of linguistic variation in EU law*. Amsterdam/ Philadelphia, John Benjamins, 1-26.

MORI, Laura (2018b). "Observing Eurolects: The case of Italian". In: MORI, Laura (ed.). *Observing Eurolects. Corpus analysis of linguistic variation in EU law*. Amsterdam/ Philadelphia, John Benjamins, 199-242.

MORI, Laura/SZMRECSANYI, Benedikt (2020). "Mapping Eurolects. An aggregate perspective on similarities between legislative varieties". https://doi.org/10.1075/lic.19017.mor. (accessed May 12th, 2023).

SALVI, Giampaolo/VANELLI, Laura (2004). *Nuova grammatica italiana*. Bologna: Il Mulino.

ŠARČEVIĆ, Susan (2015). "Basic Principles of Term Formation in the Multilingual and Multicultural Context of EU Law". In: ŠARČEVIĆ, Susan. (ed.), *Language and Culture in EU Law. Multidisciplinary Perspectives*. Farnham/Surrey: Ashgate Publishing, 183-205.

SCHMITZ, Katrin/DI VENANZIO, Laura/SCHERGER, Anna-Lena (2016). Subject Omissions and Realizations in Italian and Spanish Heritage Speakers in Germany. *Lingua* 180, 101-123. https://doi.org/10.1016/J.LINGUA.2016.04.004.

SCHMITZ, Katrin (2023). "Integrating technolects in micro-variation: Null subjects in Italian national and EU primary law texts". In: POMINO, Natascha/REMBERGER, Eva-Maria/ZWINK, Julia (eds.). *From formal linguistic theory to the art of historical editions: The multifaceted dimensions of Romance linguistics*. Vandenhoeck & Ruprecht, 127-144.

SCHMITZ, Katrin/SCHERGER, Anna-Lena (2025). "I soggetti postverbali nell'italiano lingua ereditaria in Germania". In: FESENMEIER, Ludwig/PACIARONI, Tanja/DESSÍ SCHMID, Sara (eds.), *Movimenti – Bewegungen III. Akten der Sprachwissenschaftlichen Sektion des Deutschen Italienistentags 2022*. Tübingen: Narr. 37-62.

TOMBESI, Elena (2024). "*Italiano giuridico ed euroletto a confronto. Storie di parole, tendenze traduttive e interferenze*". Firenze: Franco Cesati Editore.

VENTURI, Giulia (2012). "Investigating legal language peculiarities across different types of Italian legal texts: an NLP-based approach". In: *IAFL Porto 2012 Proceedings*, 139-156.

Legal sources:

Consolidated version of the Treaty of the European Union (TEU), ABl. C 202 from 07.06.2016, 13–388.https://eur-lex.europa.eu/legal-content/DE-EN-FR-ES-IT/TXT/?from=DE&uri=CELEX%3A02016M%2FTXT-20200301. (downloads 14.11.2016)

COSTITUZIONE DELLA REPUBBLICA ITALIANA. Edizione con note. https://www.senato.it/istituzione/la-costituzione. (download 30.07.2017).