Testing the hypothesis of "translation as a catalyst for plain legislation" on the syntactic level: a comparison of different varieties of legislative Italian

Paolo CANAVESE, PhD

University of Geneva, TRANSIUS Centre
40, boulevard du Pont-d'Arve
CH-1211 Genève 4, Switzerland
Paolo.Canavese@unige.ch

ORCID: https://orcid.org/0000-0003-3928-8905

Laura MORI, Full professor

Università degli Studi Internazionali di Roma - UNINT2
Via Cristoforo Colombo 200
00147 Rome, Italy
laura.mori@unint.eu

ORCID: https://orcid.org/0000-0002-2596-1401

Abstract: Translation is often considered as an ally of plain legal language. This corpus-based study sets out to provide empirical support for this hypothesis by comparing different varieties of legislative Italian used in a monolingual context (Italy) and in two multilingual settings (Switzerland and the European Union). The investigation relies primarily on a quantitative analysis of syntax informed by natural language processing (NLP) methods. The results suggest that translated legislation features shorter sentences, fewer nominalizations, an underuse of the passive voice, fewer non-finite clauses, less deep syntactic trees, shorter dependency links, and a preference for the SVO order with an explicit subject. Among the two multilingual contexts, Swiss legislation shows a slightly higher level of accessibility compared to EU directives. A complementary analysis using readability metrics confirms these trends. Nevertheless, in addition to the translation process in multilingual contexts, other latent external variables may have a (hardly quantifiable) impact on the level of accessibility, such as institutional language policies, legal traditions, drafting guidelines, and training programs for translators and language experts.

Keywords: legislative Italian; Eurolect; Swiss Italian; syntactic complexity; readability; plain language; accessibility; corpus-based study.

La traduzione come catalizzatore di chiarezza nei testi normativi a partire da un'analisi comparativa della sintassi in varietà di italiano legislativo

Abstract: La traduzione è spesso considerata un'alleata della chiarezza linguistica dei testi giuridici. Questo studio basato su corpora intende fornire una validazione empirica di questa ipotesi, comparando le varietà di italiano legislativo in uso in un contesto monolingue (Italia) e all'interno di due sistemi multilingui dove l'italiano è lingua ufficiale (Svizzera e Unione europea). La ricerca si basa primariamente su un'analisi quantitativa della sintassi condotta mediante strumenti di trattamento automatico del linguaggio (TAL). I risultati mostrano che la legislazione tradotta si caratterizza per frasi più brevi, un minor ricorso alla nominalizzazione, al passivo e alle subordinate implicite, alberi sintattici meno profondi, relazioni di dipendenza più brevi e una preferenza per l'ordine SVO con soggetto esplicito. Tra i due contesti multilingui, la legislazione svizzera presenta un livello di accessibilità leggermente superiore rispetto alle direttive UE. Un'analisi complementare condotta mediante metriche di leggibilità conferma queste tendenze. Tuttavia, oltre al processo di traduzione tipico dei contesti multilingui, altre variabili esterne latenti possono avere un influsso (difficilmente quantificabile) sul livello di accessibilità, quali politiche

linguistiche in campo istituzionale, tradizioni giuridiche, linee guida redazionali e attività formative per traduttori ed esperti linguistici.

Parole chiave: italiano legislativo; euroletto; italiano svizzero; complessità sintattica; leggibilità; chiarezza; accessibilità; studio su corpus.

1. Background

In recent years, both academia and institutions have been increasingly concerned with the accessibility of legal and institutional texts. This is particularly true in multilingual contexts, where a number of external factors, directly or indirectly related to legislative drafting, can influence the linguistic quality of legislation in terms of its plainness.[1] These factors are related to institutional language policies, legal traditions, as well as drafting guidelines and training programs for translators and language experts intended as operative tools (Mori 2021). Indeed, the alignment of legal contexts, be they supranational or national, with a more or less marked culture of accessibility in institutional communication influences the linguistic complexity and readability of legislative acts. In this respect, legislative multilingualism, ensured by translation, initiates virtuous dynamics that lead to the drafting of texts with fewer "obscure law" characteristics (Ainis 1997) and, therefore, greater accessibility. Unsurprisingly, the hypothesis that translation can improve the plainness of legislative texts has been frequently expressed in the legal

---

[1] The term "plainness" echoes the well-known concept of "plain language", which is defined by the International Plain Language Federation as follows: "communication is in plain language if its wording, structure, and design are so clear that the intended readers can easily find what they need, understand what they find, and use that information" (see https://www.iplfederation.org/plain-language/, accessed January 2, 2024). For the origins of the plain language movements, see, for example, Felsenfeld (1981) and Redish (1985). For a synthesis of the main features of plain language, see Adler (2012). Since the 1970s, plain language movements have expanded and spread to different countries. Today, we have a number of guidelines (see, e.g., Cutts 2013 for plain English), various organizations advocating for plain language principles (see, e.g., the above-mentioned International Plain Language Federation), an ISO standard (https://www.iso.org/standard/78907.html, last accessed January 3, 2024), and an extensive body of literature. For the latest research in this field, see, among others, Ződi (2019), van Domselaar (2022), and Williams (2023).

and linguistic literature (see, e.g., Schnyder 2001; Flückiger 2005; Egger and Ferrari 2016; Ziller 2017).

The primary objective of institutional texts is to ensure smooth communication between the State and its citizens, so that the latter can understand their duties and exercise their rights, and ultimately enjoy full inclusion in public life. With this in mind, this study sets out to investigate the differences between monolingual national and multilingual national and supranational drafting contexts.

From this perspective, Italian represents an interesting case study because it allows us to focus on three different varieties that represent the sociolinguistic outcomes of three legal contexts: one characterized by monolingual practices (domestic legislative Italian), one related to a multilingual country (Swiss legislative Italian), and one related to a supranational multilingual organization (Italian Eurolect).[2] This comprehensive view of different manifestations of legislative Italian allows us to describe the continuum of legislative varieties within the linguistic landscape of contemporary institutional Italian.

Recently, the hypothesis of translation as a "catalyst" for good practices in legislative drafting has been tested using corpus-based methods in various studies of legislative Italian across different contexts, which provided us with a robust foundation for the present research project. Previous corpus-oriented analyses aimed to uncover the degree of lexical complexity of different varieties of legislative Italian, revealing a higher level of lexical accessibility of Swiss and EU legislation compared to Italian national laws (Felici and Mori 2019; Canavese 2022a). Other previous comparative studies shed light on syntax and readability, and their results point in the same direction as the explorations of the lexical level (see Mori 2019a and Canavese 2023a).[3] Overall, these results have also highlighted remarkable

---

[2] The term Eurolect was originally proposed in French (*eurolecte*) by Goffin (1997), who argued for its status as a language for specific purposes, with specific linguistic features not limited to lexical items and terms. In the Eurolect Observatory Project (see Mori 2018a), this term is used to refer to the EU legal variety, characterized by specific co-occurring variants that differ from the homologous national legal varieties. According to this research framework, we refer to Eurolect (in this case, Italian Eurolect) to define the EU legislative language variety, in which variation occurs through the under- or over-representation of morphological, morpho-syntactic and syntactic features together with EU-related terminology.

[3] A recent study by Mori and Venturi (2021) also focused on the bidirectionality that affects the Italian language used in implementing legislation, distinguishing it as an

patterns that can be interpreted in accordance with the sociolinguistic dynamics of hybridization and harmonization in text genres within the legislative domain (see Mori 2019b).

The present study aims to further this comparison of legislative contexts where Italian is an official language, with a focus on syntax and readability. We employ the same computational linguistic method used by Mori (2019a) but expand the study to include a broader range of syntactic variables, perform statistical significance tests, and integrate the comparison with Swiss Italian.

The objective is to determine whether and to what extent the Italian versions of Swiss federal acts and EU directives in Italian resort to plainer syntactic choices compared to Italian domestic legislation and implementing laws. In doing so, we try to demonstrate the correlation between external variables and linguistic variation by providing evidence of how the contextual frame in which laws are drafted affects syntactic complexity and, consequently, readability. A special emphasis is also placed on the comparison of the two translation-mediated contexts, which exhibit both similarities (multilingual lawmaking process) and dissimilarities (direct vs. indirect applicability). These are the two research questions we attempt to answer:

RQ1   Do Swiss and EU legislative acts use plainer syntax compared to Italian domestic legislation and (EU-derived) implementing laws?

RQ2   Comparing the two varieties of translated Italian (Italian Eurolect and Swiss Italian), is there a different correlation between the independent variable "translated legislation" and the dependent variable "syntactic plainness"?

A preliminary activity focused on the theoretical operationalization of "syntactic complexity", which resulted in a list of 14 variables often associated with syntactic complexity (further literature will be provided alongside the discussion of the results). This allowed us to formalize the phenomenon under analysis and to extract quantitative measures. To facilitate the interpretation of the results, we grouped the relevant features into six categories:

---

autonomous variety that differs significantly from both Italian Eurolect and the domestic legislative variety.

1. syntactic complexity on the shallow level
2. nominal style
3. morphosyntactic features
4. non-finite subordination
5. syntactic tree complexity
6. information structure

In the following sections, we first outline our research protocol (Section 2). The quantitative results, discussed in Section 3, will be complemented by a qualitative interpretation of examples randomly selected from the four corpora to enhance the understanding of the quantitative data. Section 4 summarizes the main trends identified and presents some concluding remarks.

2. Research design

2.1 Materials

To carry out this study, we employed four corpora compiled within two different research projects. We used the three Italian corpora of the Eurolect Observatory Multilingual Corpus (EOMC, Mori 2018b, Mori 2019a), which contain EU directives (corpus A), Italian implementing laws of corpus A (corpus B), and Italian domestic legislation without any link to the supranational EU context (corpus C). For Swiss legislation, we used LEX.CH.IT (Canavese 2019), a corpus of Swiss federal acts in their Italian version, divided into three subcorpora (P1: 1974–1992, P2: 1993–2006, P3: 2007–2018). In order to ensure comparability within the selected corpora, we decided to focus only on P2. Not only is the timespan covered by the four corpora comparable, but also the corpus design criteria allowed us to easily focus on the textual sections considered as the most representative of the legal domain, i.e., the enacting terms. Table 1 provides further information on the corpora:

Table 1: Features of the corpora under comparison

Corpus	Genre/variety		Period	Size[4]
LEX.CH.IT (P2)	*Swiss federal acts*	*it_CH*	*1993– 2006*	*160 texts 497,891 tokens*
EOMC-it A	*EU directives*	*it_EU*	*1999– 2008*	*660 texts 1,439,069 tokens*
EOMC-it B	*Implementing laws of EU directives in corpus A*	*it_EU> IT*	*1999– 2013*	*275 texts 1,978,795 tokens*
EOMC-it C	*Italian domestic legislation*	*it_IT*	*1999– 2013*	*299 texts 1,511,738 tokens*

It should be emphasized that these four corpora of legislation represent distinct text genres. For the Swiss context, only federal acts passed by the legislative branch are considered; by their very nature, they lay down general principles to be later implemented by executive ordinances. EU directives and implementing laws, on the other hand, contain more technical provisions. The same applies to the corpus of Italian legislation, which includes decree-laws and legislative decrees drafted by the executive branch. These legal genres exhibit diverse levels of technicality, which may have an impact on syntactic complexity. This aspect deserves further investigation, as it may constitute a confounding variable in our study.

2.2 Methodology

To answer the research questions detailed in Section 1, we devised a three-step methodology.

---

[4] The "size" column displays the total number of texts and tokens included in the corpora. For this study, however, we decided to exclude the shortest texts. An exploratory study of the Swiss corpus revealed that short texts tend to display extreme syntactic values. For example, a text consisting of only one sentence without any subordinate clauses has 100% coordinate clauses. This value, however, is not representative of the language variety analyzed. This preliminary analysis made it possible to set a threshold to compensate for this risk of distortion; for the Swiss corpus, it seemed relevant to exclude texts with a length in the first quartile (i.e., texts with a length of less than 844 words). The same threshold was applied to the EOMC corpora.

1. Linguistic profiling

First, we performed an automatic NLP-based linguistic profiling of the four selected corpora using the annotation tools developed by the ItalianNLP Lab at the ILC of Pisa (see Montemagni 2013; Dell'Orletta, Montemagni and Venturi 2013). Following an incremental logic, the texts are first split into sentences and tokenized, then POS tagged and dependency parsed. This allowed us to measure 14 variables identified as the most relevant for our goal.

At this stage, we also calculated some new generation readability metrics using READ-IT (Dell'Orletta, Montemagni and Venturi 2011), a tool developed by the same research team that takes into account linguistic complexity parameters, both lexical and morphosyntactic.

2. Statistical analyses

To make sense of the dataset created in step 1 and to ensure a solid comparison of the four corpora, we used descriptive and inferential statistics. To identify statistically significant differences between translated and non-translated legislation (RQ1), we performed the t-test for independent samples for each variable, comparing the highest mean value between the two translation contexts (CH and EU) and the lowest between the two non-translated contexts (EU>IT and IT). We set the alpha level (significance level) at 0.05. We applied the same logic to compare the two contexts of translated legislation (RQ2). The results of the statistical tests are reported in the tables in Section 3 using different symbols. For RQ1, we marked with a tick (✓) the observations that are in line with the starting hypotheses, with a cross (×) the observations disproving the starting hypotheses. For RQ2, we indicated the context (CH or EU) that displays statistically significantly lower values of syntactic complexity. For both RQs, we used a dash (–) when no statistical significance was found. We carried out the statistical analyses in SPSS (IBM Corp. 2019).

3.   Visualization of variables

To better illustrate the cross-system trends identified, we visualized the six groups of variables by means of line plots. Since each variable is measured on a different scale, we normalized the observations by transforming mean values into z-scores.

Before analyzing each group of variables, we carried out an exploratory investigation in which we compared the four corpora by simultaneously taking into account all the quantitative variables and by visualizing their distance by means of agglomerative cluster analysis. The parameters adopted for this analysis are detailed in Section 3.1.

For this phase, we used RStudio and, more specifically, the ggplot2 (Wickham 2016) and ggdendro (de Vries and Ripley 2020) packages.

3. Results

3.1 Exploratory analysis

Before delving into a fine-grained analysis of the single variables of syntactic complexity selected, we performed a multidimensional exploratory data analysis using the agglomerative cluster analysis technique (Everitt et al. 2011). This technique is often employed in corpus linguistics (Gries 2013: 336–349) to highlight the distance between different corpora or observations by simultaneously taking into account a large number of variables. This approach allowed us to get an initial picture of the proximity and distance between the four corpora under investigation based on the 14 syntactic complexity variables mentioned in Section 1.[5]

---

[5] More specifically, we set the following parameters: z-score transformation, Manhattan distance, Ward's method for hierarchical clustering and data visualization by means of a dendrogram.
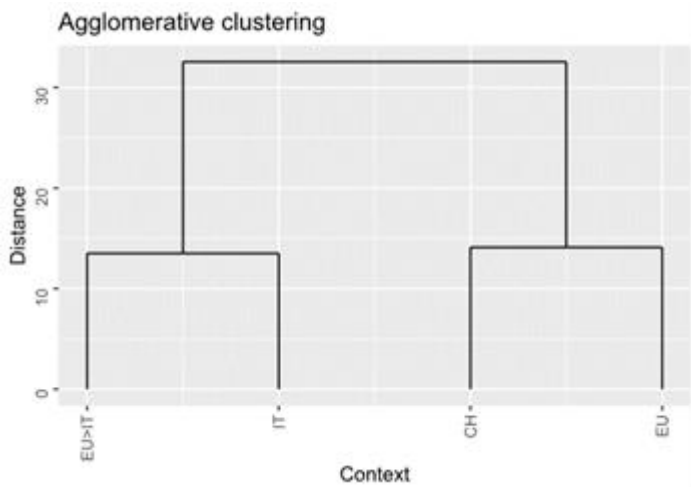
Graph 1: Agglomerative cluster analysis: distance between the four corpora

As clearly shown in Graph 1, the analysis yields two clusters. A first cluster (on the right) contains the two corpora of translated legislation (CH and EU), while a second cluster (on the left) highlights the proximity of the two monolingual corpora, i.e., legislative texts implementing EU directives in Italy and Italian domestic legislation (EU>IT and IT, respectively). The distance between the two clusters is greater than the distance between the two elements that make up each cluster, suggesting a relevant difference in the syntactic structure of translated and non-translated legislation.

This exploratory result is noteworthy, as the aim of this study is to compare the outcomes of multilingual legislative contexts with monolingual ones in terms of syntactic complexity. In what follows, we try to determine whether this difference can be explained through our starting hypothesis of plainer syntax in multilingual translation-mediated legal contexts. In order to thoroughly explore this assumption, we will first analyze the six groups of variables mentioned in Section 1 individually. The analysis of syntactic features potentially affecting linguistic accessibility will then be combined with a comparative readability assessment of the texts included in the legislative corpora under analysis.

3.2 Syntactic complexity on the shallow level

The traditional way to approximate syntactic complexity is to examine the surface of a text, and more specifically its mean sentence length. This is the basic assumption underlying first-generation readability formulae (e.g., the Flesh-Kincaid Grade Level, Kincaid et al. 1975, or the Gulpease Index for Italian, Lucisano and Piemontese 1988). The advantage of relying on mean sentence length is its simplicity of measurement. Although it is a basic metric, previous studies (e.g., Szmrecsanyi 2004) have shown that it is good predictor of syntactic complexity. It comes as no surprise that the rule of thumb to formulate short sentences is one of the first recommendations in any sets of guidelines for clear legal and institutional Italian (see, e.g., Ittig and Accademia della Crusca 2011: 19). Sometimes, a threshold is also suggested. For example, in the Swiss context, it is recommended not to exceed 15 words per sentence (OFJ 2007: 363), while for Italian administrative texts the maximum length is generally set at 20–25 words per sentence (Cortelazzo and Pellegrino 2003: 94). In addition to the mean sentence length, we also report the mean clause length:

Table 2: Syntactic complexity on the shallow level

Variable	CH	EU	EU>IT	IT	RQ1	RQ2
Sentence length	17.61	19.23	22.84	24.48	✓	CH
Clause length	12.81	13.79	15.11	16.66	✓	CH
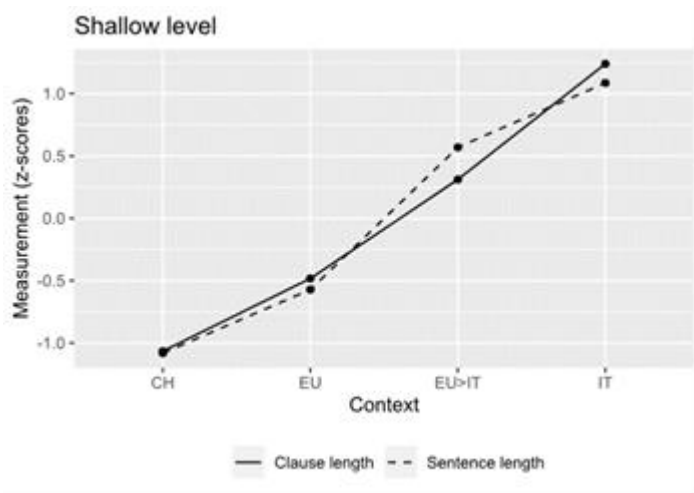
Graph 2: Syntactic complexity on the shallow level

Both variables are in line with the starting hypothesis. Indeed, translated legislation tends to pack fewer words into a sentence or clause compared to monolingual legislation. It may seem surprising that even the highest values of Italian domestic legislation comply with the plain legal language rules mentioned above. However, these data should be interpreted in the light of the macro-structure of a legislative text, which typically contains several titles or list elements that the annotation tool calculates as full sentences. While their positive influence on clarity is undeniable, they may distort the actual mean length of the provisions that make up the act.

This first group of variables provides a first approximation of syntactic complexity. However, new-generation NLP tools, such as the pipeline used in this study, allow for a much finer-grained analysis of the syntactic structure of a text. In the next sections, we will rely on the results of the automatic POS tagging and dependency parsing to explore different facets of syntactic complexity.

3.3 Nominal style

The second group of variables attempts to capture the preference of legal texts for nominal style, i.e., the overuse of nouns over verbs. Three variables were chosen to investigate this feature: the noun/verb ratio, the mean length of prepositional chains (i.e., nouns recursively modified by prepositional phrases) and the percentage of long prepositional chains (with three or more modifying phrases) in relation to the total number of prepositional chains.

Table 3: Nominal style

Variable	CH	EU	EU>IT	IT	RQ1	RQ2
Noun/verb ratio	3.00	3.24	3.50	3.96	✓	–
Prepositional chains length	1.47	1.43	1.52	1.59	✓	EU
Prepositional chains ≥3	9.18	8.87	11.58	13.71	✓	–
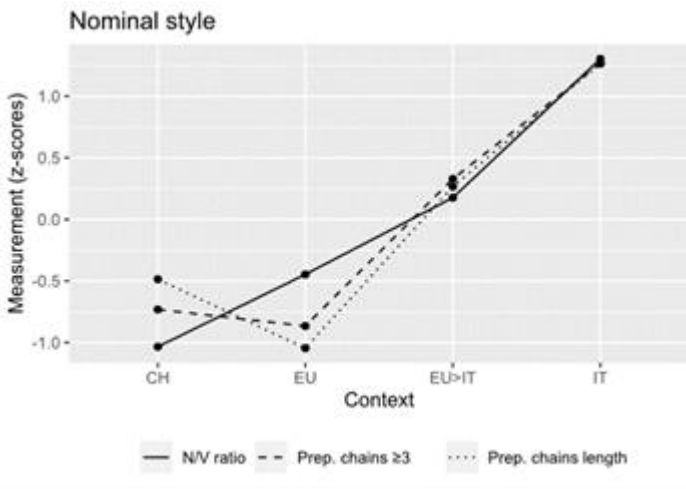
Graph 3: Nominal style

While the presence of nouns is inevitably high in specialized written texts with a strong denotative component, it is crucial to

minimize unnecessary nominalizations to produce plain legal texts. The overuse of nouns, particularly when the same content could be conveyed with a verb, generally results in a text that is denser,[6] more static, abstract, implicit, and cognitively challenging to process (Raso 2005: 112–113; Gotti 2005: 83–85). To illustrate the syntactic complexity associated with the use of the nominal style, we can look at example (1) below:

> (1) Art. 670
> Nomina nel servizio permanente
> […]
> 2. I vincitori del concorso assumono una anzianità assoluta pari a quella posseduta **nel grado di capitano o di tenente di vascello alla data del decreto di nomina in servizio permanente effettivo**, diminuita di due anni, e prendono posto nei rispettivi ruoli, in relazione a detta anzianità assoluta, nell'ordine della graduatoria del concorso, dopo l'ultimo pari grado avente la stessa anzianità assoluta. (EOMC-it C, Legislative decree 2010/14)[7]

In this paragraph of 61 words, we find only two explicit verbs. Moreover, the large part of the text highlighted in bold contains eight prepositional phrases (composed of a preposition and a noun: *nel grado*, *di capitano*, *di tenente*, *di vascello*, *alla data*, *del decreto*, *di nomina*, *in servizio*). They form complex chains of recursively embedded phrases in which one phrase modifies the previous one. Using more verbs instead of nouns can improve the reading flow, as illustrated by comparing example (2a) taken from article 7, paragraph 2 of the EU Directive laying down specific provisions for the control and eradication of bluetongue with the rephrasing of the same norm in the Italian implementing measure (2b):

> (2a) **Per coordinare** pienamente tutte le misure necessarie all'eradicazione della febbre catarrale degli ovini con la massima

---

[6] Although, in some cases, nominalization can also increase the number of words needed to express the same piece of information (see examples 1a and 1b below).

[7] Translation by the authors: "Art. 670 / Appointment to the permanent service [...] / 2. The successful candidates shall assume an absolute seniority equal to that held in the rank of captain or lieutenant at the date of the decree of appointment to the effective permanent service, minus two years, and shall take their place in their respective roles, in relation to that absolute seniority, in the order of the competition ranking list, after the last-ranking candidate with the same absolute seniority."

tempestività e per condurre l'indagine epidemiologica, viene istituita una cellula di crisi. (EOMC-it A, Directive 2000/75/EC)[8]

(2b) **Ai fini del coordinamento** di tutte le misure necessarie all'eradicazione della malattia e dell'indagine epidemiologica ci si avvale dell'unità di crisi di cui all'allegato IV al decreto del Presidente della Repubblica 17 maggio 1996, n. 362. (EOMC-it B, Legislative decree 2003/225)[9]

Here, the syntactic complexity goes beyond the mere replacement of the verb (*coordinare*) by a noun (*coordinamento*) in the transposition of the EU directive into an Italian national measure, and it extends to the prepositional phrase in which this change occurs. In fact, more noun phrases are used to convey the same piece of information, thereby raising the register. First, a complex preposition (*ai fini di*) is used instead of the simple preposition *per*, which results in a more bureaucratic wording. Second, the direct object in the EU directive (*tutte le misure necessarie*) is transformed into a prepositional phrase in the implementing act (*di tutte le misure necessarie*), making the syntactic tree heavier (see Section 3.6).

As discussed by Mori (2019a: 643) based on qualitative observations, verbal structures that occur in EU directives are often (intralingually) translated into implementing laws using nominalizations. This finding aligns with the quantitative data presented here. Compared to implementing laws and to Italian domestic legislation, EU legislation translated into Italian features a lower noun/verb ratio, shorter prepositional chains and less frequent prepositional chain accumulations. The comparison between Swiss federal acts and EU directives is more blurred; the former use slightly longer prepositional chains than the latter, while the noun/verb ratio and the use of prepositional chains longer than three phrases do not display any statistically significant difference.

---

[8] Official English version: "A crisis unit shall be established to carry out the overall coordination of all the measures necessary for ensuring the eradication of bluetongue as soon as possible and to conduct the epidemiological survey." In this example and in the next ones, the bold marking is by the authors.
[9] Translation by the authors: "For the purpose of coordinating all the measures necessary for the eradication of the disease and the epidemiological investigation, the crisis unit referred to in Annex IV to presidential decree No. 362 of May 17, 1996 will be used."

3.4 Morphosyntactic features

The use of the subjunctive mood and passive clauses are two morphosyntactic features that are often discouraged, unless necessary, by guidelines for plain legal and institutional language (see, e.g., Cortelazzo and Pellegrino 2003: 102–108; Raso 2005: 89–92, 115–117; OFJ 2007: 382).

Table 4: Morphosyntactic features

Variable	CH	EU	EU>IT	IT	RQ1	RQ2
Subjunctive/indicative ratio	9.27	20.81	18.80	13.36	×	CH
Passive clauses	10.50	10.87	11.67	11.87	✓	–
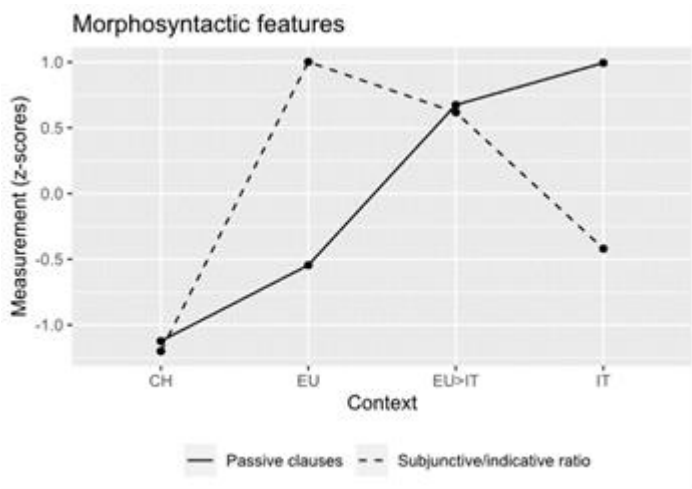
Graph 4: Morphosyntactic features

The metric "subjunctive/indicative ratio"[10] shows its lowest values in Swiss legislation, whereas it seems to be a distinctive feature of EU directives. Implementing laws do not deviate from this trait, as shown by the absence of statistical significance between the values

---

[10] Calculated as the occurrences of verbs in the subjunctive mood, divided by the occurrences of verbs in the indicative mood, multiplied by 100.

calculated on the EU and EU>IT corpora. In fact, they use even more verbs in the subjunctive mood compared to Italian domestic legislation.

This result aligns with a previous comparison between Italian Eurolect and Italian of implementing laws (see Mori 2018b) and with the investigation of lexical complexity across the above-mentioned four corpora (see Canavese 2022a: 136, 139), which showed a statistically significant overrepresentation of the conditional connectives *qualora* and *a condizione che* with obligatory subjunctive mood in EU directives and *ove* and *nel caso in cui* in implementing laws.[11] In general, the choice of a verb in the subjunctive mood rather than the indicative tends to produce morphosyntactic configurations with other related complexities. For example, conditions are often expressed using multi-word and more complex or even archaic or bureaucratic variants, such as those just mentioned, instead of the simple connective *se* (*if*) + indicative mood.

However, the overuse of the subjunctive mood in EU directives should be interpreted in the light of their legal nature and, specifically, their indirect applicability; directives set objectives without detailing how Member States are to achieve them. On the linguistic side, this means a higher frequency of syntactic structures such as *Gli Stati membri provvedono affinché...* (*Member States shall...*) with obligatory subjunctive mood:

> (3) Gli Stati membri **provvedono affinché** le carni fresche, le carni macinate e le preparazioni di carni di cui al paragrafo 1 **siano** bollate in conformità della direttiva 2002/99/CE […]. (EOMC-it A, Directive 2003/85/EC)[12]

Frequently, this syntactic structure generates a list of final clauses introduced by *affinché* (Mori 2019a: 642), which leads to a proliferation of verbs in the subjunctive mood.

---

[11] The frequency of these connectives is much lower in Swiss legislation. In Swiss federal acts, only the connective *sempreché* + subjunctive mood is overrepresented, although the frequencies are lower compared to the connectives overrepresented in EU directives and implementing laws. Moreover, the frequency of *sempreché* in Swiss legislation has dropped in the period 2007-2018 (Canavese 2022a: 140).

[12] Official English version: "Member States shall ensure that fresh meat, minced meat and meat preparations as referred to in paragraph 1, shall be marked in accordance with Directive 2002/99/EC [...]".

Regarding the use of the passive voice, our analysis reveals only small intra-corpora differences, with a slight, statistically significant underuse in multilingual contexts. The passive voice can lead to depersonalization of the provision and also to syntactic vagueness when the agent is omitted. At the same time, according to the literature (see, e.g., Höfler 2016 and Nussbaumer 2000), the use of the passive voice can sometimes enhance comprehensibility, i.e., when the patient represents a given information and is topicalized by means of the passive voice, thus contributing to a smoother thematic progression, as shown in example (4):

> (4) Art. 37 Domanda di svincolo e decisione
> [1] Ogni cittadino svizzero è, su domanda, svincolo dalla sua cittadinanza se non risiede in Svizzera e possiede o gli è stata assicurata la cittadinanza di un altro Stato. […]
> [2] **Lo svincolo è pronunciato** dall'autorità del Cantone d'origine. (LEX.CH.IT, Swiss Citizenship Act of 20 June 2014)[13]

An active formulation of the provision contained in paragraph 2 would not comply with the reader's mental representation process because it would interfere with the thematic progression. While it would remain a fully grammatical and acceptable sentence, it may not be the plainest formulation. In this respect, recent work on Swiss legislation (Canavese 2023a; 2023b) has demonstrated the importance of analyzing the impact of the passive voice from a qualitative perspective in order to gain a deeper understanding of the complexity it entails depending on the cotext.

3.5 Use of non-finite subordination

Limiting the use of subordination is a primary guideline of plain language (Cortelazzo and Pellegrino 2003: 79–81). Indeed, the use of multiple subordinate clauses within a sentence makes the syntactic

---

[13] Unofficial translation into English by the Federal Chancellery: "Art. 37 Request for relief and decision / [1] Swiss citizens shall on request be relieved of their Swiss citizenship if they are not resident in Switzerland and hold or have been assured of another nationality. [...] / [2] Relief of citizenship is granted by the authority in the canton of origin."

tree heavier (see Section 3.6) and sentences longer (see Section 3.2), thus increasing the reader's cognitive effort to process the message correctly.

At the same time, it has been argued that subordination *per se* cannot be considered a proxy for complexity: packing information into one main and one subordinate clause may be textually clearer than splitting the same piece of information into two main sentences (Mortara Garavelli 2003). This holds particularly true in legal texts, where subordination is essential to fulfill some recurring textual pattern such as "conditioning fact + legal consequence".

In addition to introducing the condition, subordinate clauses also establish a conclusive, concessive, or relative link between the elements that constitute a provision (Visconti 2013), thus hierarchizing information in foregrounds and backgrounds. For these reasons, we did not consider the mere use of subordinate over coordinate clauses as a parameter of syntactic complexity. We also excluded the use of long, recursively embedded subordination chains from our investigation, even though they are a typical vector of syntactic complexity, because preliminary explorations of our dataset revealed that this is a marginal feature in all the contexts under investigation. Instead, we decided to focus only on the use of non-finite subordinate clauses. Similar to nominalizations, non-finite clauses make a text denser and increase the potential for ambiguity (Cortelazzo and Pellegrino 2003: 81–86; Raso 2005: 116). Therefore, their use should be minimized and limited to those cases where they are justified and do not undermine clarity.

Table 5: Use of non-finite subordination

Variable		CH	EU	EU>IT	IT	RQ1	RQ2
Non-finite subordinate clauses		56.93	49.62	65.02	75.49	✓	EU
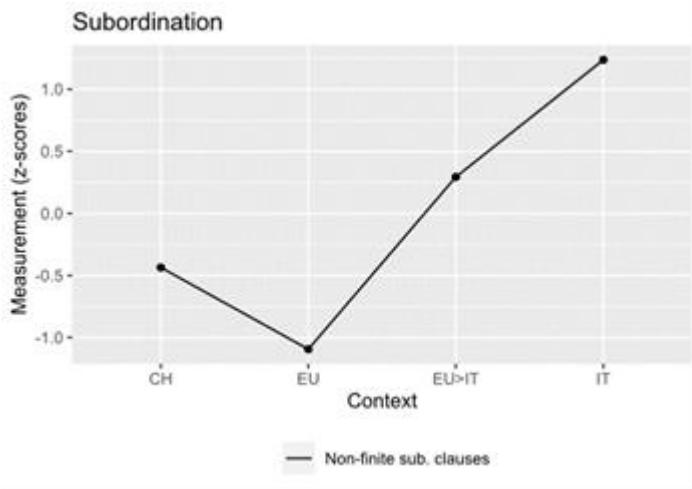
Graph 5: Use of non-finite subordinate clauses

In example (5), we find four subordinate clauses:

(5) Art. 15.
Chiusura delle indagini preliminari
1. **Ricevuta** la relazione di cui all'articolo 11, il pubblico ministero, se non richiede l'archiviazione, esercita l'azione penale, **formulando** l'imputazione e **autorizzando** la citazione dell'imputato. (EOMC-it C, Legislative decree 2000/274)[14]

---

[14] Translation into English by the authors: "Art. 15. / Completion of the preliminary investigations / 1. Upon receipt of the report referred to in Article 11, the attorney general, if she or he does not request its dismissal, shall initiate the prosecution, by formulating the charge and authorizing the summoning of the defendant."

In addition to the finite conditional clause set between the subject and the main verb, three non-finite clauses are used: a temporal participle clause at the beginning of the sentence and two coordinated adverbial clauses of means at the end. Overall, the provision is neither ambiguous nor unclear, but the use of non-finite subordination certainly packs several pieces of information into a limited number of words (24), thus creating a bureaucratic style. From this perspective, the infrequent use of non-finite subordinate clauses in the two multilingual legislative contexts under analysis can be interpreted as an indicator of greater syntactic plainness.[15]
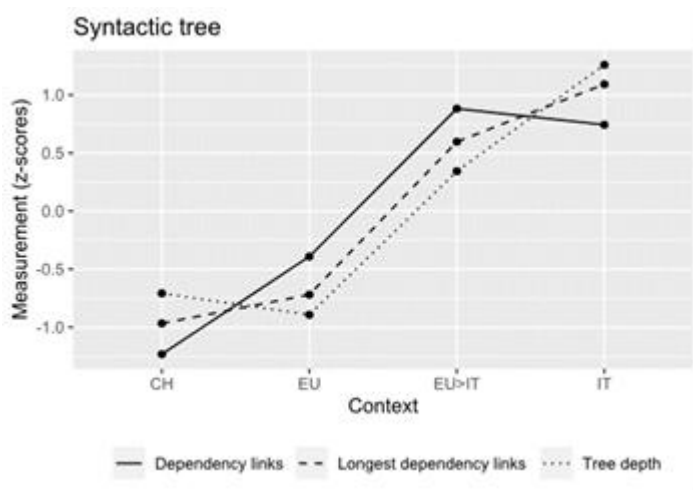
3.6 Syntactic tree complexity

The idea of packing one piece of information per sentence is a long-standing principle in plain (legal) language (see, e.g., OFJ 2007: 359). This helps readers process a sentence and avoids overloading the "working memory" (Cowan 2010), i.e., the number of elements that must be kept in mind in order to decode the sentence. Sentences that convey multiple pieces of information tend not only to be longer (see Section 3.2), but also to exhibit a deeper syntactic tree structure.

Using dependency parsing, it is possible to compute some metrics that capture and quantify the syntactic tree complexity. In this section, three metrics are analyzed: 1) the overall *tree depth*, which corresponds to the number of dependency links between the root and the furthest leaf; 2) the *dependency links length*, calculated as the mean number of tokens (excluding punctuation) between each head and each dependent; and 3) the mean length of the *longest dependency link* per sentence. The higher the values of these three variables, the higher the cognitive load required to process the sentence (Venturi 2012b: 146).

---

[15] Preliminary analyses showed that both Swiss federal acts and EU directives use more subordinating conjunctions. This is reflected in the value of non-finite subordinate clauses.

Table 6: Syntactic tree complexity

Variable	CH	EU	EU>IT	IT	RQ1	RQ2
Tree depth	5.70	5.64	6.05	6.36	✓	–
Dependency links length	2.15	2.37	2.71	2.67	✓	CH
Longest dependency links	7.23	7.58	9.47	10.18	✓	CH

Graph 6: Syntactic tree complexity

Example (6), taken from the Swiss Federal Act on Foreign Nationals and Integration, is representative of the mean sentence complexity in Swiss federal acts:

> (6) Art. 11 Soggiorno con attività lucrativa
> [1] Lo straniero che intende esercitare un'attività lucrativa in Svizzera necessita di un permesso indipendentemente dalla durata del soggiorno. […] (LEX.CH.IT, Federal Act on Foreign Nationals and Integration of 16 December 2005)[16]

---

[16] Unofficial translation into English by the Federal Chancellery: "Art. 11 Permit requirement for period of stay with gainful employment/ [1] Foreign nationals who wish to work in Switzerland require a permit irrespective of the period of stay. [...]"

Here, the syntactic tree has a depth of 6 links, corresponding to the distance between the root *necessita* and the leaf *Svizzera*. The longest dependency link is the subject link: 9 tokens separate the head *necessita* from its dependent *straniero*. The mean length of the dependency links is 1,67 tokens.

Overall, the figures reported in Table 6 may seem exceptionally low considering the genre under analysis, especially with regard to Italian domestic legislation. Previous descriptions of legislative Italian have reported its tendency to employ convoluted sentences (see, e.g., Mortara Garavelli 2001: 155–180 for a catalog of syntactic features of legal language), while the data collected here are closer to figures calculated on newspaper language (e.g., Brunato 2014). Again, this is due to the fact that legal texts contain several short titles, which are computed as full sentences by the automatic annotator (see Section 3.2). The same applies to lists, where each item is treated as a full sentence. However, the advantage of this segmentation method is that it is commonly used in quantitative studies of legal language (see Venturi 2012a), thus ensuring comparability with previous and future studies. Moreover, the legislative genres represented in the four corpora analyzed are structurally similar; this means that intra-genre comparability is not undermined and that this methodology is valid in terms of answering our research questions.

This set of variables confirms our starting hypothesis of plainer syntax in translated legislation compared to monolingual texts. Swiss federal acts and EU directives feature less deep syntactic trees and shorter dependency links. This implies that less information is packed into each sentence and that fewer parenthetical elements interrupt the reading flow. The difference between the Swiss and EU varieties of legislative Italian is small, with EU directives exhibiting slightly longer dependency links.

3.7 Information structure (syntactic categories-semantic roles)

Dependency parsing can also provide quantitative insights into aspects at the intersection of syntax and information structure, such as

constituent order. In this section, our interest laid more in the descriptive dimension, since these features cannot be automatically attributed to syntactic complexity without an in-depth qualitative analysis of the informational level.

In fact, while the subject-verb-object (SVO) order is often recommended to comply with plain language principles (Ittig and Accademia della Crusca 2011: 19), this is not always the most effectual way to pack information from a textual perspective. For example, if the subject corresponds to the new piece of information, a VOS order might be the most appropriate in terms of comprehensibility (see Canavese 2022b for an analysis of non-SVO sentences in legal language). As a result, the use of postverbal subjects – and, along the same lines, of preverbal objects – can be justified by the need to maintain topic continuity, which is a powerful device to ensure coherence, thus having a positive impact on comprehensibility.

However, since most guidelines for clear legal and administrative language, as well as previous studies in this field, consider these variables as traits of syntactic complexity, we decided to include them in this paper. Moreover, observing the positional distribution of subjects and objects within sentences can help to identify differences between the language varieties under investigation and allow us to speculate on the potential effect of source-language interferences. The same applies to null-subject sentences; as Italian is a pro-drop language, the subject can be omitted if it can be easily retrieved thanks to co-textual information.

Table 7: Information structure

Variable	CH	EU	EU>IT	IT	RQ1	RQ2
Null-subject sentences	11.00	20.82	24.89	24.48	✓	CH
Post-verbal subjects	10.59	11.94	20.07	23.35	✓	CH
Preverbal objects	3.19	2.54	2.97	2.08	×	EU
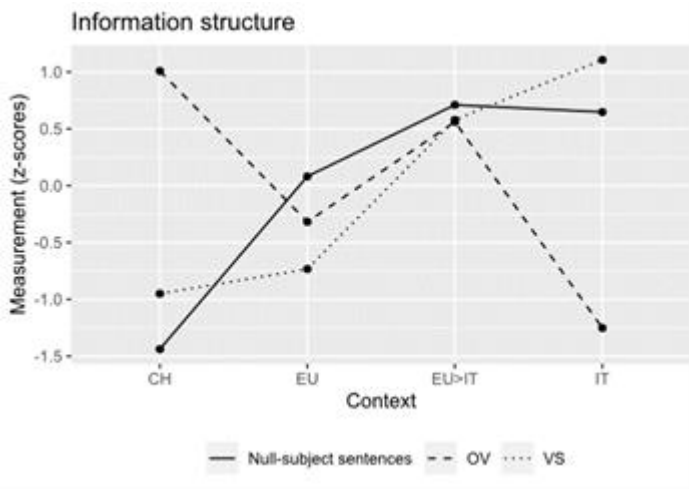
Graph 7: Information structure

These data show that translated legislation tends to adhere to the SVO structure. In particular, the frequency of post-verbal subjects is more than twice as high in the corpus of Italian domestic legislation compared to Swiss federal acts. The anteposition of objects is overall infrequent in all four language varieties, and the slightly higher figure in Swiss legislation can hardly be considered a proxy for complexity.

In the case of EU directives, the preference for SVO structures could be determined by the influence of the source texts, which are primarily drafted in English or French, both non-pro-drop languages. As argued in previous studies (e.g., Mori 2018b; 2019a), plain language solutions adopted in the English version of EU legislation may also be reproduced in the Italian translation. Similarly, for Swiss federal acts, the adherence to the SVO constituent order may depend on the deeply rooted plain language culture in Swiss legislative drafting (see Section 1).

Null subjects are much less frequent in Swiss legislation compared to all other language varieties. Once again, translated legislation displays lower values compared to domestic legislation. This finding is consistent with Schmitz's (2023) comparative study of the Italian Constitution and the Italian version of the Treaty on European Union. Qualitative analyses of Swiss federal acts show that "light" anaphoric devices in the subject position, such as pronouns, in the German and French versions are often rendered with a null-subject structure in Italian. Other anaphoric devices used to ensure cohesion include the use of repetitions or hypernyms. Null subjects usually refer back to the last subject mentioned in the text, as in example (7).

> (7) Art. 746
> Mancato superamento del corso
> 1. **Gli allievi** che non hanno superato gli esami teorici o che sono stati giudicati non idonei ad assumere il grado di sottotenente e gradi corrispondenti di complemento, pur avendo superato le prove prescritte per il conferimento del brevetto di pilota militare o del brevetto di navigatore militare, conseguono la nomina a pilota militare o navigatore militare. In tale qualità Ø sono tenuti a prestare servizio con il grado di sergente o corrispondente di complemento per un periodo di sei anni, decorrente dalla data d'inizio dei corsi di pilotaggio e dei corsi di navigatori. (Legislative decree 2010/66)[17]

Even if the null subject (Ø) in the second sentence refers back to a referent introduced over 50 words earlier, there is no ambiguity. This means that null subjects are not necessarily always a source of incomprehension. In general, however, the further the last subject, the more challenging the process of anaphora resolution. In fact, the probability that another referent mentioned between the null subject and the antecedent will be interpreted as the subject of the new sentence increases, thus slowing down the reading and comprehension process. From this perspective, the more frequent use of null subjects

---

[17] Translation into English by the authors: Art. 746 / Failure to pass the course / 1. Cadets who have failed the theoretical examinations or have been found unfit to assume the rank of second lieutenant and corresponding complement ranks, although they have passed the tests prescribed for the award of the military pilot's license or military navigator's license, shall be appointed military pilot or military navigator. In this capacity, they are required to serve with the rank of sergeant or corresponding complement ranks for a period of six years, starting from the date of the beginning of the pilot and navigator courses."

in the corpora with higher average sentence length values (see Section 3.2) can be interpreted as an indicator of more complex syntax in non-translated legislation.

3.8 Readability assessment

In this final section of the analysis, we apply the same approach proposed in Mori (2019a) to evaluate the relationship between the results of the linguistic profiling discussed so far and the readability scores computed by means of the READ-IT tool. As described in Section 2.2, this tool automatically assesses the readability level of Italian texts using a wide range of syntactic features (including those discussed in the previous sections), as well as lexical variables, and it calculates complexity by applying three metrics. The results are reported in Table 8.

Table 8: Readability metrics

Readability metric	CH	EU	EU>IT	IT	RQ1	RQ2
Base metric	46.29	75.24	92.34	91.96	✓	CH
Lexical metric	61.05	84.38	81.91	80.00	×	CH
Syntactic metric	80.55	97.37	99.86	99.73	✓	CH
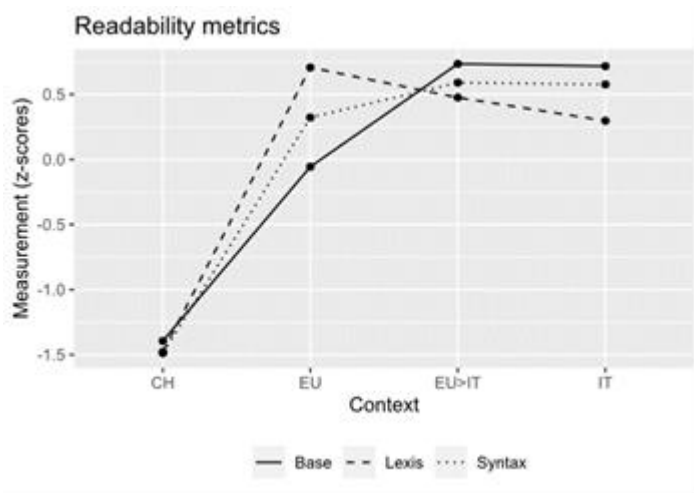
Graph 8: Readability metrics

This stage allows us to extend the scope of our analysis beyond the syntactic level, considering how and to what extent it interrelates with the results obtained on the base and lexical levels.

In general, Swiss legal texts have the lowest complexity scores and, conversely, the highest readability scores on the three levels considered here. The base metric follows the same logic as traditional readability formulae, adopting parameters such as mean sentence length and word length. The results indicate lower complexity values for the multilingual legislative contexts (CH and EU) compared to texts drafted within the monolingual Italian context.

The lexical and syntactic models show that, with the exception of Swiss legislation, the other three varieties of legislative Italian have very similar levels of complexity. The surprisingly high figure displayed by EU directives at the lexical level could depend, at least partially, on the presence of references to supranational bodies, procedures and policies (the so-called "Europeisms").[18] Since all these

---

[18] The term "Europeism" was introduced in Italian by Mori (2003) to denote lexical items and noun phrases that aim to s refer to concepts, institutions, policies, principles and practices that are specific to the EU environment. In line with this definition, this lexical category was specified into the following five subcategories: a. Semantic Europeisms for lexemes that have undergone a resemantization process in different Eurolects; b. Newly coined EU words; c. EU noun phrases, d. EU acronyms, and e.

items are not recognized by the annotation tool, which is based on the Basic Italian Vocabulary (De Mauro 2016), they are categorized as complex lexis.

The high level of syntactic complexity should be interpreted in the light of the reference corpora used by READ-IT, i.e., *DueParole* (Piemontese 1996), a corpus based on newspapers written in an easy-to-read language, on the [- complex] pole, and *La Repubblica*, a corpus containing newspaper articles written in standard language, on the [+ complex] pole along the variation continuum. One caveat must be kept in mind: laws, even in their plainest version, are specialized texts; therefore, they can hardly be less complex than newspaper texts. Despite this methodological limitation, these results are useful for comparing different varieties belonging to the same macro-genre within the same domain. Indeed, they confirm the same trend of lower linguistic complexity of legislation in translation-mediated institutional contexts compared to monolingual drafting institutions in terms of readability. They are also useful to make a comparison between the Swiss and EU varieties of legislative Italian, confirming that the former is characterized by even more accessible linguistic profiles.

4. Conclusions

This study of syntactic complexity and readability across different varieties of legislative Italian provided us with solid empirical data to address our initial research questions.

RQ1: Translation and plain language

Our findings confirm the existence of a correlation between the variables "translation" and "plain language", which is in line with the starting hypothesis. Indeed, compared to Italian domestic and EU-derived Italian legislation, Swiss and EU legislative acts are characterized by:
- shorter sentences and clauses

EU-based metaphors to convey the EU's political message (particularly abundant in EU primary law, see Menza and Mori 2022).

- fewer nouns and nominalizations
- an underuse of the passive voice
- fewer non-finite clauses
- less deep syntactic trees and shorter dependency links
- a preference for the SVO order with an explicit subject

In general, less information is packed into each sentence, thus enhancing the message processability and adhering to the "one sentence-one message" principle. However, correlation is not causation, and we are unable to establish a cause-effect relationship based on these findings. As argued in Section 1, several co-occurring contextual variables may have a positive impact on higher accessibility in multilingual legal contexts.

RQ2: Comparison between Swiss and EU legislation

The comparative analysis of Swiss and EU legislative texts revealed that the former exhibit a slightly, albeit statistically significantly, lower level of syntactic complexity. Nearly all the variables examined point to plainer syntactic solutions in Swiss acts compared to EU directives.

The influence of the supranational context emerges in the overrepresentation of the subjunctive mood, which can be interpreted as evidence of EU-rooted phenomena (Mori 2018b) related to specific principles of EU law, such as the indirect applicability of directives. This is also reflected linguistically, for example, in a higher use of conditional clauses, sometimes introduced by connectives requiring the subjunctive mood.

Peculiarities of the implementing laws of EU directives

In addition to answering the main research questions, our data also allow for further interpretations. Observing the figures concerning the implementing laws of EU directives in Italian legislation, we notice that they are overall halfway between domestic legal Italian and Eurolect. This confirms their "eccentric position": implementing laws display both features typical of the language variety from which they originate (Italian Eurolect) and of the one with which they have to comply, i.e., domestic legislative Italian (see Mori 2019a; Mori and Venturi 2021). This also highlights the

influence that legal cultures and political contexts can have on linguistic features at different levels, in relation to the underlying dynamics of hybridization and harmonization within the legal domain (see Mori 2019b).

In conclusion, our study has allowed us to empirically describe the variation within legislative texts drafted (or translated) in Italian across three different institutional contexts. This comparison shows that promoting the adoption of a plainer language to improve linguistic accessibility in a specialized field such as the legal one is not merely desirable, but also achievable, and has already been achieved in some contexts. Comparing the results of drafting and translation best practices can help to advance this goal, which has deeper socio-cultural and political implications that go beyond the mere linguistic analysis.

Notes and acknowledgment

Authors Contribution

This paper is the result of joint analyses, discussions and interpretations by both authors. However, for academic purposes, Paolo Canavese drafted Sections 2 and 3.2–3.7, whereas Laura Mori authored Sections 1, 3.1, 3.8 and 4.

Conflict of interest

The authors declare that there is no conflict of interest.

AI Use statement

AI was not used in the paper.

Bibliography

Adler, Mark. 2012. The Plain Language Movement. In: *The Oxford Handbook of Language and Law*, eds. Lawrence M. Solan and Peter M. Tiersma, 67–83. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780199572120.013.0006.

Ainis, Michele. 1997. *La legge oscura. Come e perché non funziona*. Rome/Bari: Laterza.

Brunato, Dominique. 2014. Complessità necessaria o stereotipi del burocratese? Un'indagine sulla leggibilità del linguaggio amministrativo da una prospettiva linguistico-computazionale. In: *La lingua variabile nei testi letterari, artistici e funzionali contemporanei. Analisi, interpretazione, traduzione. Atti del XIII Congresso SILFI Società Internazionale di Linguistica e Filologia Italiana. Palermo, 22–24 settembre 2014*, eds. Giovanni Ruffino and Marina Castiglione. Florence: Franco Cesati.

Canavese, Paolo. 2019. LEX.CH.IT. A Corpus for Micro-Diachronic Linguistic Investigations of Swiss Normative Acts in Italian. *Comparative Legilinguistics*, 40: 43–65. DOI: https://doi.org/10.14746/cl.2019.40.3.

Canavese, Paolo. 2022a. Plain Legal Language through Translation. A Comparison of Swiss, EU and Italian Legislative Texts. In: *Contemporary Approaches to Legal Linguistics*, eds. Daniel Green and Luke Green, 123–151. Münster: Lit.

Canavese, Paolo. 2022b. L'ordine canonico dei costituenti argomentali nella semplificazione della scrittura normativa. *Rivista internazionale di tecnica della traduzione/International Journal of Translation*, 24: 197–216. DOI: https://doi.org/10.13137/2421-6763/34297.

Canavese, Paolo. 2023a. *Atti normativi elvetici nella terza lingua ufficiale. Chiarezza de jure o de facto?* PhD dissertation, University of Geneva (Switzerland).

Canavese, Paolo. 2023b. Translation and Clarity in Swiss Multilingual Legislation. A Qualitative Study on Textuality. *Translation and Interpreting*, 15: 77–95. DOI: https://doi.org/10.12807/ti.115202.2023.a05.

Cortelazzo, Michele. A., and Federica Pellegrino. 2003. *Guida alla scrittura istituzionale*. Rome/Bari: Laterza.

Cowan, Nelson. 2010. The Magical Mystery Four. How is Working Memory Capacity Limited, and Why? *Current Directions in Psychological Science*, 19, no. 1: 51–57. DOI: https://doi.org/10.1177/0963721409359277.

Cutts, Martin. 2013. *Oxford Guide to Plain English*. 4th ed. Oxford: Oxford University Press.

De Mauro, Tullio. 2016. Il Nuovo vocabolario di base della lingua italiana. *Internazionale*. https://www.internazionale.it/opinione/tullio-de-mauro/2016/12/23/il-nuovo-vocabolario-di-base-della-lingua-italiana (accessed January 2, 2024).

de Vries, Andrie, and Brian D. Ripley. 2020. ggdendro. Create Dendrograms and Tree Diagrams Using 'ggplot2'. https://cran.r-project.org/web/packages/ggdendro/index.html (accessed January 2, 2024).

Dell'Orletta, Felice, Simonetta Montemagni, and Giulia Venturi. 2011. READ-IT. Assessing Readability of Italian Texts with a View to Text Simplification. In: *Proceedings of the 2nd Workshop on Speech and Language Processing for Assistive Technologies, Edinburgh, 30 July 2011*, ed. Norman Alm, 73–83. Stroudsburg: Association for Computational Linguistics. https://aclanthology.org/W11-2308 (accessed January 2, 2024).

Dell'Orletta, Felice, Simonetta Montemagni, and Giulia Venturi. 2013. Linguistic Profiling of Texts Across Textual Genres and Readability Levels. An Exploratory Study on Italian Fictional

Prose. In: *Proceedings of the International Conference Recent Advances in Natural Language Processing RANLP. Hissar, 7–11 September 2013*, eds. Galia Angelova, Kalina Bontcheva and Ruslan Mitkov, 189–197. https://aclanthology.org/R13-1025 (accessed January 2, 2024).

Egger, Jean-Luc, and Angela Ferrari. 2016. L'italiano federale svizzero. Elementi per una ricognizione. *Studi italiani di linguistica teorica e applicata*, XLV, no. 3: 499–523.

Everitt, Brian S., Sabine Landau, Morven Leese, and Daniel Stahl. 2011. *Cluster analysis*. 5th ed. Chichester: Wiley.

Felici, Annarita, and Laura Mori. 2019. Corpora di italiano legislativo a confronto. Dall'Unione europea alla Cancelleria svizzera. In: *Le tendenze dell'italiano contemporaneo rivisitate. Atti del LII Congresso Internazionale di Studi della Società di Linguistica Italiana. Berna, 6–8 settembre 2018*, eds. Bruno Moretti, Aline Kunz, Silvia Natale and Etna Krakenberger, 287–304. Milan: Officinaventuno. DOI: https://doi.org/10.17469/O2102SLI000017.

Felsenfeld, Carl. 1981. The Plain English Movement in the United States. *Canadian Business Law Journal*, 6: 408–421. https://ir.lawnet.fordham.edu/cgi/viewcontent.cgi?article=1475&context=faculty_scholarship (accessed January 3, 2024).

Flückiger, Alexandre. 2005. Le multilinguisme de l'Union européenne. Un défi pour la qualité de la législation. In : *Jurilinguistique. Entre langues et droits/Jurilinguistics. Between Law and Language*, eds. Jean-Claude Gémar and Nicholas Kasirer, 340–360. Bruxelles: Bruylant.

Goffin, Roger. 1997. L'eurolecte: le langage d'une Europe communautaire en devenir. *Terminologie et Traduction*, no. 1: 63–73. https://op.europa.eu/en/publication-detail/-/publication/7f9b3873-3cc8-4d21-8b5e-d589ad4851d6/language-fr/format-PDF/source-search (accessed January 2, 2024).

Gotti, Maurizio. 2005. *Investigating Specialized Discourse*. Berne: Peter Lang.

Gries, Stefan T. 2013. *Statistics for Linguistics with R. A Practical Introduction*. Berlin: De Gruyter.

Höfler, Stefan. 2016. Die Informationsstruktur von Rechtssätzen und ihre Bedeutung für die Gesetzesredaktion. *LeGes*, 27, no. 2: 225–251.

https://leges.weblaw.ch/legesissues/2016/2/20162225-251.html (accessed January 2, 2024).

IBM Corp. 2020. IBM SPSS Statistics for Windows. Version 26.0. Computer software.

Ittig, and Accademia della Crusca. 2011. *Guida alla redazione degli atti amministrativi. Regole e suggerimenti*. Florence: Istituto di teoria e tecniche dell'informazione giuridica del CNR. http://www.ittig.cnr.it/Ricerca/Testi/GuidaAttiAmministrativi.pdf (accessed January 2, 2024).

Kincaid, J. Peter, Richard L. Fishburne, Richard L. Rogers, and Brad S. Chissom. 1975. Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel. In: *Research Branch Report*, 8–75. Millington: Chief of Naval Technical Training.

Lucisano, Pietro, and Maria Emanuela Piemontese. 1988. Gulpease. Una formula per la predizione della difficoltà dei testi in lingua italiana. *Scuola e Città*, XXXIX, no. 3: 110–124.

Menza, Salvatore, and Laura Mori. 2022. La metaforizzazione del linguaggio giuridico per la costruzione discorsiva dell'identità sovranazionale nel diritto dell'Unione europea. In: *L'integrazione europea: tra sfide politiche, vincoli giuridici e legittimazione democratica*, eds. Salvatore Menza and Valentina Petralia, 9–36. Florence: Le Monnier Università.

Montemagni, Simonetta. 2013. Tecnologie linguistico-computazionali e monitoraggio della lingua italiana. *Studi Italiani di Linguistica Teorica e Applicata*, XLII, no. 1: 145–172.

Mori, Laura, and Giulia Venturi. 2021. Norme italiane di derivazione europea tra direttive dell'UE e leggi nazionali: un'analisi linguistico-computazionale della variazione intralinguistica. Poster presented at the *LIV Congresso della Società di Linguistica Italiana (SLI), "Corpora e studi linguistici", 8–10 September 2021, Università degli Studi di Firenze (Italy)*.

Mori, Laura. 2003. L'euroletto: genesi e sviluppo dell'italiano comunitario. In: *Ecologia linguistica. Atti del XXXVI Convegno internazionale di Studi della Società di Linguistica italiana*, eds. Ada Valentini, Piera Molinelli, Pierluigi Cuzzolin and Giuliano Bernini, 473–492. Roma: Bulzoni.

Mori, Laura. 2018a (ed.). *Observing Eurolects. Corpus Analysis of Linguistic Variation*. Amsterdam: Benjamins.

Mori, Laura. 2018b. Observing Eurolects. The Case of Italian. In: *Observing Eurolects. Corpus Analysis of Linguistic Variation*, ed. Laura Mori, 199–242. Amsterdam: Benjamins.

Mori, Laura. 2019a. Complessità sintattica e leggibilità. Un monitoraggio linguistico per la valutazione dell'accessibilità dei testi legislativi europei e italiani. *Studi Italiani di Linguistica Teorica e Applicata*, XLVIII, no. 3: 627–657.

Mori, Laura. 2019b. Dall'armonizzazione all'ibridazione nei testi legislativi: evidenze linguistiche e manifestazioni interculturali nell'italiano tradotto. *Entreculturasi*, 10: 377–392. https://revistas.uma.es/index.php/revtracom/article/view/12539 (accessed January 2, 2024).

Mori, Laura. 2021. Prospettive di miglioramento dell'accessibilità linguistica in italiano: verso una cultura della comunicazione istituzionale *citizen-centred*. In: *Comunicare il patrimonio culturale: accessibilità comunicativa, tecnologie, sostenibilità*, ed. Franca Orletti, 149–161. Milan: Franco Angeli. https://series.francoangeli.it/index.php/oa/catalog/view/807/653/4708 (accessed January 2, 2024).

Mortara Garavelli, Bice. 2001. *Le parole e la giustizia. Divagazioni grammaticali e retoriche su testi giuridici italiani*. Turin: Einaudi.

Mortara Garavelli, Bice. 2003. Strutture testuali e stereotipi nel linguaggio forense. In: *La lingua, la legge, la professione forense*, ed. Alarico Mariani Marini, 3–20. Milan: Giuffrè.

Nussbaumer, Markus. 2000. Textlinguistik für die Gesetzgebung. In: *Linguistica giuridica italiana e tedesca/Rechtslinguistik des Deutschen und Italienischen*, ed. Daniela Veronesi, 196–208. Padua: Unipress.

OFJ. 2007. *Guide de législation. Guide pour l'élaboration de la législation fédérale*. 3th ed. Berne: Office fédérale de la justice (OFJ). https://www.bj.admin.ch/dam/bj/fr/data/staat/legistik/hauptinstrumente/gleitf-2007-f.pdf (accessed January 2, 2024).

Piemontese, Maria Emanuela. 1996. *Capire e farsi capire. Teorie e tecniche della scrittura controllata*. Naples: Tecnodid.

Raso, Tommaso. 2005. *La scrittura burocratica. La lingua e l'organizzazione del testo*. Rome: Carocci.

Redish, Janice. 1985. The Plain English Movement. In: *The English Language Today*, ed. Sidney Greenbaum, 125–138. Oxford/New York: Pergamon.

Schmitz, Katrin. 2023. Integrating Technolects in (Micro-)Variation. Null Subjects in Italian National and EU Primary Law Texts. In: *From Formal Linguistic Theory to the Art of Historical Editions The Multifaceted Dimensions of Romance Linguistics*, eds. Natascha Pomino, Eva-Maria Remberger and Julia Zwink, 127–144. Vienna: V&R unipress, Vienna University Press.

Schnyder, Bernhard. 2001. Zur Mehrsprachigkeit der schweizerischen Gesetzgebung im Allgemeinen. *LeGes*, 12, no. 3: 33–48. https://leges.weblaw.ch/it/legesissues/2001/3/2001333-48.html (accessed January 2, 2024).

Szmrecsanyi, Benedikt. 2004. On Operationalizing Syntactic Complexity. In: *Le poids des mots. Proceedings of the 7th International Conference on Textual Data Statistical Analysis. Louvain-la-Neuve, 10–12 March 2004*, eds. Purnelle Gérald, Cédrick Fairon and Anne Dister, 1032–1039. Louvain-la-Neuve: Presses universitaires de Louvain.

van Domselaar, Iris. 2022. "Plain" Legal Language by Courts. Mere Clarity, an Expression of Civic Friendship or a Masquerade of Violence? *The Theory and Practice of Legislation*, 10, no. 1: 93–111. DOI: https://doi.org/10.1080/20508840.2022.2033946.

Venturi, Giulia. 2012a. Design and Development of TEMIS. A Syntactically and Semantically Annotated Corpus of Italian Legislative Texts. In: *Proceedings of the 4th Workshop on Semantic Processing of Legal Texts (SPLeT 2012), held in conjunction with LREC 2012, Istanbul, 27 May 2012*, 1–12. http://www.italianlp.it/wp-content/uploads/2014/10/TEMIS-Venturi.pdf (accessed January 2, 2024).

Venturi, Giulia. 2012b. Investigating Legal Language Peculiarities across Different Types of Italian Legal Texts. An NLP-Based Approach. In: *Bridging the Gap(s) between Language and the Law. Proceedings of the 3rd European Conference of the International Association of Forensic Linguistics. Porto, 15–18 October 2012*, eds. Rui Sousa-Silva, Rita Faria, Núria Gavaldà and Belinda Maia, 139–156. Porto: Faculdade de Letras da Universidade do Porto.

Visconti, Jacqueline. 2013. Uno sguardo al contesto europeo: connettivi in testi normativi dell'Unione europea. In: *Le forme linguistiche dell'ufficialità. L'italiano giuridico e amministrativo della Confederazione Svizzera*, eds. Jean-Luc Egger, Angela Ferrari and Letizia Lala, 67–81. Bellinzona: Casagrande.

Wickham, Hadley. 2016. *ggplot2. Elegant Graphics for Data Analysis*. 2nd ed. New York: Springer.

Williams, Christopher. 2023. *The Impact of Plain Language on Legal English in the United Kingdom*. London/New York: Routledge.

Ziller, Jacques. 2017. Lingue e politica linguistica nell'Unione europea. In: *Il linguaggio giuridico nell'Europa delle pluralità. Lingua italiana e percorsi di produzione e circolazione del diritto dell'Unione europea. Atti della Giornata di studio*, 21–38. Rome: Senato della Repubblica. https://iris.unito.it/bitstream/2318/1637521/2/BarberaCorinoOnesti2017_LinguisticaGiuridica_senato_compresso.pdf (accessed January 2, 2024).

Ződi, Zsolt. 2019. The Limits of Plain Legal Language. Understanding the Comprehensible Style in Law. *International Journal of Law in Context*, 15, no. 3: 246–262. DOI: https://doi.org/10.1017/S1744552319000260.