

Blini, Lorenzo/Laura Mori: Structural and functional analysis of lexical bundles in Eurolects and national legislative varieties: a Spanish-Italian contrastive study. In: www.polyphonie.at, Vol. 16, Nr. 2/2025, ISSN: 2304-7607, begutachteter Beitrag/peer-reviewed article.

Lorenzo Blini and Laura Mori (Università degli Studi Internazionali di Roma).

Structural and functional analysis of lexical bundles in Eurolects and national legislative varieties: a Spanish-Italian contrastive study¹

1. Introduction

Corpus-assisted methodologies for language analysis may be applied for the automatic extraction of lexical bundles (Biber/Barbieri 2007: 265) with a statistical tendency towards co-occurrence. This makes it possible to collect data on the formulaicity gradient of a given textual genre featured by word combinations that perform discursive functions.

This approach focuses on the use frequency of cognitively motivated and holistically stored units (Biber et al. 2003), which allows phraseology to be analysed and interpreted from a new perspective, using corpus-driven analyses (Tognini-Bonelli 2001). This may be particularly note-worthy when examining the legal language, whose high level of standardisation is evident in frequent lexical bundles that play an important cognitive, pragmatic and discursive role (Mattila 2013: 107).

The present study aims to investigate the presence of lexical bundles in three legislative corpora in Spanish and Italian, in the following texts: European directives (Corpus A) which are binding legal instruments requiring transposition into national legislation before their provisions become applicable, implementing national laws of Corpus A directives (Corpus B)² and national laws without any European derivation (Corpus C).

¹ The study was jointly conceived by the two authors who discussed any phase of the research. The analyses on Italian were led by Laura Mori and on Spanish by Lorenzo Blini. For academic purposes the paper sections are attributed as follows: Lorenzo Blini (§ 1-2-4-5.2); Laura Mori (§ 3-5.1-5.3); section 6 was written by both authors.

² Corpora B and C in Spanish include five categories of legal acts: Ley Orgánica, Ley Ordinaria, Real Decreto-Ley, Real Decreto Legislativo, Real Decreto. Corpora B and C in Italian contain two types of acts: Legge and Decreto legislativo.



Due to the different drafting processes that characterise the European multilingual context as opposed to national contexts, the intra-linguistic comparison between A and B has already revealed remarkable results at different levels, both for Spanish (Blini 2018a) and Italian (Mori 2018). In fact, in contexts such as the European Union with a multilingual regime, the recurrence and fixity of certain sequences is determined by the interlinguistic contact that takes place during the circular process of drafting-translation-revision from which the legislation originates (Mori 2019b). Adding Corpus C, as representative of the national legislative varieties that characterise texts resulting from a monolingual drafting process, adds a third element of comparison.

Previous studies have confirmed the relevance of the contextual variable when comparing EU and national legal corpora causing the over- or under-representation of variants. Therefore, the cross-corpora analyses (A vs. B vs. C) will allow us to isolate and interpret trends with regard to the distribution of lexical bundles within the European legislation in Spanish and Italian, the so-called *eurolects*, contrasted with to the national legislative variety derived from the European legislation and the national legislative variety.

The formal and functional categorisation here adopted to identify the most frequent types of lexical bundles in Spanish and Italian legislative discourse is based on some previous works (Biel 2018, Blini 2018, Mori 2019b) but it is aimed to propose a more fine-grained interpretation of linguistic variability related to the internal organisation of legislative texts; namely the difference between preambles and enacting parts.

2. Research aim

Qualitative and quantitative analyses led to date on Spanish (Blini 2018a, 2018b, 2024) and Italian (Felici/Mori 2019, Mori 2018, Mori 2020a, 2020b) have shown that there is a significant correlation between external variables and linguistic variation in laws. Drafting standards and practices in use in different contexts produce preferred linguistic features in a given textual genre. As it has already been argued (Flückinger 2005) and recently demonstrated (Felici/Griebel 2019, Canavese/Mori 2024), in situations of legislative multilingualism (such as the EU one), translation processes give rise to a greater need for standardisation, formulaicity and language plainness. Therefore, our present aim is to evaluate contrastively how these variational trends affect the textual construction as far as lexical



bundles that perform the following functions in legal discourse³ are concerned:

- 1) *content bundles*, composed of nominal and prepositional phrases, which can be extended to nominal clauses with non-prototypical subjects and to prepositional clauses playing a predominant referential function (examples in tables 5–8);
- 2) *legal framing bundles*, used to insert the norm within the universe of legal discourse and to confer the legal authority through inter- and intratextual references that play as mechanisms of legal framing (examples in tables 9–12);
- 3) *deontic and performative bundles*, which create deontic scenarios of obligation, permission and prohibition in legislative texts (examples in tables 13–15).

In the following sections we are going to present materials and our methodological approach (§ 3). Follows a quantitative analysis (§ 4) and the discussion of corpus-driven data in Spanish and Italian with a contrastive approach (§ 5) in order to highlight analogies and differences from inter and intra-linguistic perspectives (§ 6).

3. Materials and methodology

For the study Spanish and Italian data were extracted from corpora A, B and C included in the *Eurolect Observatory Multilingual Corpus* (EOMC)⁴ and organised as follows in Italian (Table 1) and in Spanish (Table 2):

³ It is possible to identify another a category, that composed of figurative lexical bundles, i.e. combinations of words characterised by the presence of metaphors, typically explited in primary laws where the founding principles of the State entity are defined (e.g. Treaties and Constitutions). As far as the European discourse is concerned, Menza/Mori (2022) focused on metaphors used for the global activation of a functional framework for the discursive construction of the European *res* (such as spatial or movement metaphors, since the free movement of people, goods and capital within a supranational geopolitical space has been at the very heart of the European Union project since its beginning).

⁴ The *Eurolect Observatory Multilingual Corpus* was developed within the Eurolect Observatory Project (www.unint.eu/eurolect-observatory/overview). For more details see Mori (ed., 2018). Data reported in Tables 1-2 refer to the whole corpora; our analyses were led on sub-corpora containing either Preambles or Enacting Parts in accordance with specific research objectives. Annexes were never considered since they are not representative of legal language per se.



	Genre	Time span	Size
Corpus A	EU directives in Italian	1999-2008	660 texts ±3.500.000 tokens
Corpus B	National laws implementing Corpus A directives	1999-2013	275 texts ±2.700.000 tokens
Corpus C	National laws (EU unrelated)	1999-2013	299 texts ±2.800.000 tokens

Table 1. Italian corpora

	Genre	Time span	Size
Corpus A	EU directives in Spanish	1999-2008	660 texts ±4.700.000 tokens
Corpus B	National laws implementing Corpus A directives	1999-2013	438 texts ±6.500.000 tokens
Corpus C	National laws (EU unrelated)	1999-2013	886 texts ±6.900.000 tokens

Table 2. Spanish corpora

The quantitative analysis of lexical bundles (§ 4) was carried out using an inductive procedure via a corpus-driven methodology: the most frequent lexical-syntactic combinations from a minimum of 3 and a maximum of 6 words were extracted using the N-Grams function of WordSmithTools 6.0 software.⁵ Results here presented refer to N-grams occurring in sub-corpora (either in preambles or in enacting parts, excluding the annexes); given their different size, frequency values of lexical bundles have been normalised by 10,000.

The qualitative analysis (Section 5), based on the automatically extracted lexical bundles, considered identifiable strings that fall into the above-mentioned heuristic categories complying with the legal genre here under examination.

The most relevant examples for different extension (ranging from 3-grams to 6-grams) were selected in accordance with their frequency and salience within the above-mentioned macro-categories leading our qualitative analyses. The main criterion adopted for the selection of salient lexical bundles was a semantic-pragmatic criterion; this explains a certain amount of variability across functional categories, structural shape of grams and languages.

Our analyses were carried out on N-grams ranging from 3 to 6 units: needless to say, in many cases 3-grams are included in larger N-grams that are repeated without variation in order to ensure consistency. Therefore, in case of lexical bundles of different gram extension clearly extracted from the same textual chunk, we are reporting the most significative ones by giving relevance to the gram length (Section 5). That is to say, a 6-gram bearing a high frequency

⁵ <https://lexically.net/wordsmith/version6/>



was preferred compared with a similar 3-gram, since the longer it is, the more salient it is as far as formulaicity is concerned.

Another crucial criterion adopted in the qualitative discussion of our results concerns the text section in which they occur, since preambles and enacting parts belong to the same regulative text type, but they differ in terms of textual function: referential the former, conative the latter. This difference may affect the distribution of the above-mentioned categories of lexical bundles and it deserves to be taken into account properly.

Thanks to the EOMC mark-up, conceived in order to allow the retrieval of linguistic data from different text sections (namely preambles, enacting parts and annexes, if included), it was possible to comply with this methodological need.

4. Quantitative analysis of lexical bundles

In the following tables (3-4) lexical bundles that characterise our corpora were grouped according to the number of lexical units (ranging from 3 to 6) with a minimum threshold of 5 occurrences.

	Corpus A		Corpus B		Corpus C	
	Preamble (619.181 tokens)	Enacting Part (1.439.069 tokens)	Preamble (72.134 tokens)	Enacting Part (1.798.795 tokens)	Preamble (46.056) tokens)	Enacting Part (1.511.738 tokens)
3-grams	216,6	236,8	248,1	266,6	220,4	235,6
4-grams	147,8	160,5	217,0	171,1	183,3	151,5
5-grams	105,6	109,6	188,1	111,2	164,6	101,5
6-grams	83,8	78,0	165,0	76,1	146,6	73,7
Total	553,8	584,9	818,2	625,0	714,9	562,3

Table 3. N-grams in Preambles and Enacting Parts (Italian)

	Corpus A		Corpus B		Corpus C	
	Preamble (716.423 tokens)	Enacting Part (1.582.213 tokens)	Preamble (597.434 tokens)	Enacting Part (4.054.428 tokens)	Preamble (1.059.557 tokens)	Enacting Part (4.799.007 tokens)
3-grams	231,1	235,8	229,8	258,5	245,7	269,6
4-grams	169,3	185,2	138,3	206,4	164,8	251,8
5-grams	122,7	133,8	83,0	140,8	113,0	206,0
6-grams	96,1	98,4	60,0	96,9	90,8	170,3
Total	619,2	635,2	511,1	702,6	614,3	897,7

Table 4. N-grams in Preambles and enacting parts (Spanish)

Both tables show that as the size of the n-grams increases, the number of occurrences



decreases. At the same time, a qualitative analysis (to be discussed thoroughly in § 5) reveals that meaningful phrase structures begin to emerge due to the concatenation of several phrases and clauses that characterise the phraseology of European and national legislative varieties, both in Italian and in Spanish. The total number of lexical bundles is similar, with a slightly higher figure in Spanish. In particular, Corpus C in Spanish has the highest number of lexical bundles, followed by A and B with similar values. In Italian, Corpus B stands out to be the most formulaic, ahead of C and A.

A cross-language similarity can be observed with regard to the greater formulaicity of EU directives (Corpus A) – both in preambles and enacting parts – for both languages.

A notable difference between the two languages lies in the distribution of lexical bundles in correlation with the two sections of the law (preambles and enacting parts) that EOMC allowed us to take into due account. In Corpus A of Spanish, both sections have similar data, while in Corpus B and especially in C the amount of lexical bundles in the enacting parts increases in a relevant way. As for Italian, a similar distribution between sections features EU directives, but the situation is reversed in national corpora, where preambles have a significantly higher number of lexical bundles.

5. Typology of lexical bundles and contrastive remarks

5.1 Content bundles

This category comprises lexical bundles (made of noun phrases and prepositional phrases) whose main textual function is referential: they activate and re-activate all along the text primary concepts which play a fundamental role for the thematic coherence.

It is note-worthy that in Corpus A of both languages the top position of the list is occupied by expected referents (such as directive, Member States and the name of EU institutions) that we excluded from our exemplification in order to focus on more unexpected data.

The following Tables (5–6) contain data extracted from preambles⁶ in Italian corpora and in the Spanish ones; Tables 7–8 present a selection of the most frequent content bundles in enacting parts in corpora for both languages. As above-described (§ 3) we conducted a two-level analysis: contrastive Italian vs. Spanish and supranational vs. national for each language.

⁶ It must be noted that in Italian preambles in national corpora are significantly smaller than in Corpus A (approximately a tenth), thus affecting the total amount of lexical bundles.



5.1.1 Remarks on preambles

In legal texts, preambles are opening sections aiming at illustrating the reason behind the enactment of a law. Given this textual objective of the sub-genre, it is highly-expected to find many frequent occurrences of referential expressions composed of noun and noun phrases referring to institutions (either supranational or national), legislative instruments, legal principles, political setting and technicalities related to the specific legal context.

Corpus A – Preamble (from 3 to 6-grams)	Corpus B – Preamble (from 3 to 6-grams)	Corpus C – Preamble (from 3 to 6-grams)
- a livello comunitario - proposta della Commissione - principio di sussidiarietà - disposizioni della direttiva - esercizio delle competenze - il parere del Comitato - la proposta della Commissione - funzionamento del mercato interno - l'attuazione della presente direttiva - il parere del Parlamento europeo - Trattato che istituisce la Comunità europea - parere del Comitato economico e sociale - ravvicinamento delle legislazioni degli Stati membri - competenze di esecuzione conferite alla Commissione - le misure previste dalla presente direttiva	- attuazione della direttiva - adempimento di obblighi - le politiche comunitarie - il parere della Conferenza - Province Autonome di Trento e Bolzano - il decreto del Ministro - parere della Commissione permanente - deliberazione del Consiglio dei Ministri - i pareri delle competenti Commissioni - decreto del Presidente della Repubblica - Ministro per le Politiche Europee - obblighi derivanti dall'appartenenza dell'Italia - disposizioni per l'adempimento di obblighi - pareri delle competenti Commissioni della Camera	- decreto del Presidente - modificazioni del decreto-legge - il seguente decreto legislativo - pareri delle competenti Commissioni - deliberazione del Consiglio dei Ministri - il parere della Conferenza - i pareri delle competenti Commissioni della Camera - Ministro dell'Economia e delle Finanze - le Regioni e le Province Autonome - rapporti tra lo Stato e le Regioni - il Ministro per la Funzione Pubblica - il parere del Consiglio di Stato - delega del Governo in materia di - disposizioni integrative e correttive del decreto

Table 5. Content bundles in preambles (Italian)

Corpus A - Preamble (from 3 to 6-grams)	Corpus B - Preamble (from 3 to 6-grams)	Corpus C - Preamble (from 3 to 6-grams)
- la(s) autoridad(es) competente(s) - los productos alimenticios - las medidas previstas - la libre circulación - las competencias de ejecución - el principio de proporcionalidad - el principio de subsidiariedad	- las Administraciones Públicas - ámbito de aplicación - la disposición final - la disposición adicional - la entrada en vigor - protección de la salud - protección de los consumidores - protección del medio ambiente	- la disposición adicional - las Administraciones Públicas - sistema de financiación - cesión de tributos - la entrada en vigor - uso de la autorización - modificación de la Ley - las entidades de crédito - financiación de las



- funcionamiento del mercado interior - el ejercicio de las competencias - los contenidos máximos de residuos - la protección de la salud - la aproximación de las legislaciones - el procedimiento establecido en el artículo - las disposiciones de la presente Directiva - el anexo I de la Directiva - las medidas necesarias para la ejecución	- texto refundido de la Ley - empresas de servicios de inversión - comercio de derechos de emisión - la aproximación de las legislaciones - todos los que la presente vieren - emisión(es) de gases de efecto invernadero - las disposiciones legales reglamentarias y administrativas - planificación general de la actividad económica	- Comunidades Autónomas - régimen de cesión de tributos - fijación del alcance y condiciones - las medidas fiscales y administrativas - la autorización contenida en el artículo - el texto refundido de la Ley - la tasa de reposición de efectivos - la resolución de las reclamaciones económico-administrativas
---	---	--

Table 6. Content bundles in preambles (Spanish)

In particular, in Corpus A of both languages content bundles make it clear the belonging to the European Union, the reference to its institutions and the adherence to their procedural principles and a wide range of nouns and noun phrases belonging to ‘europeisms’ (see Blini/Mori 2019) is attested.

Comparison A vs. B/C in Italian reveal the hybridity of the legislative variety in use in Corpus B (see Mori 2019a) where procedures and legislative instruments at EU and national levels are both referred to. Focusing on B vs. C qualitative comparison in both languages, there is an expected coincidence for those content bundles referring to the national context, legal acts and administrative procedures. Quite interestingly, the cross-corpora comparison for Spanish shows a clear-cut distribution between supranational dimension vs. the national one and Corpus B reveals a very limited use of content bundles with references to the EU. Thus, revealing a different intralinguistic trend compared to what features the intralinguistic dynamics in Italian.

5.1.2 Remarks on enacting parts

In the following tables (7–8) the most frequent nouns and adjectival phrases collected in the three corpora are reported:

Corpus A – Enacting Part (from 3 to 6-grams)	Corpus B – Enacting Part (from 3 to 6-grams)	Corpus C – Enacting Part (from 3 to 6-grams)
- le autorità competenti - atto della pubblicazione	- tutela del territorio - datore di lavoro	- entrata in vigore - Arma dei Carabinieri



- entrata in vigore - destinatari della presente direttiva - le disposizioni legislative regolamentari - un riferimento alla presente direttiva - autorizzazione all'immissione in commercio - autorità competenti dello Stato membro - settore disciplinato dalla presente direttiva - riferimento all'atto della pubblicazione ufficiale - pubblicazione nella Gazzetta ufficiale dell'Unione - testo rilevante ai fini del SEE - il ventesimo giorno successivo alla pubblicazione	- sanzione amministrativa pecuniaria - l'autorità competente - il Ministro della Salute - nuovi o maggiori oneri - data di entrata in vigore - decreto del Presidente della Repubblica - Presidente del Consiglio dei Ministri - le Regioni e le Province - decreto munito del sigillo dello Stato - rapporti tra lo Stato e le Regioni	- l'anno finanziario - Ministero dell'Economia - Ministero della Difesa - unità previsionale di base - atti normativi della Repubblica - decreto del Presidente della Repubblica - stato di previsione del Ministero - atti normativi della Repubblica Italiana - le Regioni e le Province Autonome - decreto munito del sigillo dello Stato
---	--	---

Table 7. Content bundles in enacting parts (Italian)

Corpus A – Enacting Part (from 3 to 6-grams)	Corpus B – Enacting Part (from 3 to 6-grams)	Corpus C – Enacting Part (from 3 to 6-grams)
- la(s) autoridad(es) competent(es) - empresa(s) de inversión - empresa(s) de seguros - las medidas necesarias - persona(s) física(s) o jurídica(s) - la autorización de comercialización - disposiciones de Derecho interno - titular de la autorización - lo dispuesto en el artículo - lo dispuesto en el apartado - el texto de las disposiciones - el día de su publicación - las disposiciones legales, reglamentarias y administrativas - las modalidades de la mencionada referencia - ámbito regulado por la presente Directiva - lo dispuesto en la presente Directiva	- la(s) autoridad(es) competente(s) - el apartado anterior - la siguiente redacción - el párrafo anterior - la entrada en vigor - persona(s) física(s) o jurídica(s) - la red de transporte - la prestación de servicios - lo dispuesto en el artículo - texto refundido de la Ley - lo establecido en el artículo - lo dispuesto en el apartado - órgano competente de la Comunidad Autónoma - las empresas de servicios de inversión - los términos previstos en el artículo - el territorio de aplicación del impuesto	- la disposición adicional - el apartado anterior - la siguiente redacción - el párrafo anterior - la entrada en vigor - las entregas a cuenta - el ámbito de aplicación - las entidades de crédito - texto refundido de la Ley - lo dispuesto en el artículo - lo previsto en el artículo - lo establecido en el artículo - todos los españoles particulares y autoridades - accidentes de trabajo y enfermedades profesionales - ámbito de aplicación de la Ley - los términos previstos en el artículo

Table 8. Content bundles in enacting parts (Spanish)

Content bundles occurring in enacting parts of Corpus A do mention elements involved in the



legislative process and in the transposition phase (such as competent authorities, the official publication, the coming into force of the act...). In national corpora of both languages there are more detailed references to procedural iter, economic aspects, specific legal instruments and technical formulas because they allow coming into force of the act in each Member State.

5.2 Legal framing bundles

This category comprises frequent lexical bundles typically used in legal genres aimed at instantiating the legal framing with internal and inter-textual references as well as anchoring to the supranational or national contexts.

5.2.1 Remarks on Preambles

Preambles include both a general framework and some specific provisions to be specified in the enacting parts which express the binding character of the varieties here under examination.

In the following tables (9–10) we grouped the most frequent bundles collected in preambles of Italian and Spanish corpora.

Corpus A – Preamble (from 3 to 6-grams)	Corpus B – Preamble (from 3 to 6-grams)	Corpus C – Preamble (from 3 to 6-grams)
- visto il Trattato - vista la direttiva - di cui all'articolo - previste dalla presente direttiva - secondo la procedura di cui - modificata da ultimo dalla direttiva - di cui alla presente direttiva - concernente il ravvicinamento delle legislazioni - visto il parere del Comitato economico - conformi al parere del Comitato permanente - in base al principio di sussidiarietà - in ottemperanza al principio di proporzionalità	- vista la legge - vista la direttiva - visti gli articoli - visto il decreto legislativo - recante attuazione della direttiva - sulla proposta del Ministro - vista la deliberazione del Consiglio - di concerto con i Ministri - recante disposizioni per l'adempimento - visto il decreto del Presidente - acquisito il parere della Conferenza - acquisiti i pareri delle competenti Commissioni	- vista la legge - vista la deliberazione - visti gli articoli - visto l'articolo - visto il decreto legislativo - a norma dell'articolo - acquisiti i pareri delle competenti Commissioni - di concerto con i Ministri - sulla proposta del Ministro - vista la preliminare deliberazione del Consiglio - acquisito il parere della Commissione parlamentare - acquisito il parere della Conferenza permanente

Table 9. Legal framing bundles in preambles (Italian)



Corpus A – Preamble (from 3 to 6-grams)	Corpus B – Preamble (from 3 to 6-grams)	Corpus C – Preamble (from 3 to 6-grams)
- en el artículo - visto el Tratado - considerando lo siguiente - visto el dictamen - en la presente Directiva - en particular su artículo - a efectos del EEE - establecido en el artículo - visto el dictamen del Comité - de conformidad con el procedimiento - cuya última modificación la constituye - con arreglo a la decisión - las disposiciones de la presente Directiva - de conformidad con el principio de - de acuerdo con el principio de - en lo que se refiere a	- en el artículo - en la Ley - en la Directiva - de esta Ley - por el Real Decreto - en el Real Decreto - a propuesta del Ministro - al ordenamiento jurídico español - de acuerdo con el Consejo - aprobado por el Real Decreto - por la que se modifica - por el que se aprueba - previa deliberación del Consejo de Ministros - aprobado por el Real Decreto Legislativo - para su adaptación a la Ley - de modificación de diversas leyes para	- en la Ley - por la Ley - del Real Decreto - en su artículo - en el ámbito de - por el Real Decreto - a propuesta del Ministro - de la Constitución Española - de acuerdo con el Consejo - aprobado por el Real Decreto - lo dispuesto en el artículo - por la que se regula - previa deliberación del Consejo de Ministros - aprobado por el Real Decreto Legislativo - haciendo uso de la autorización contenida - a propuesta de los Ministros de

Table 10. Legal framing bundles in preambles (Spanish)

In Corpus A in both languages these kind of bundles - especially made of complex prepositions for legal framing - refer to EU legislation, previous acts and binding principles.

Comparison A vs. B/C in Italian and Spanish reveals the hybridity of the legislative variety in use in transposition measures (B) with references to both legal frames, supranational and national, are attested to a lesser extent compared with content bundles.

Focusing on B/C comparison, national corpora in both languages are characterised by similar chunks, thus complying with the intertextual features of the language within the national traditions.

5.2.2 Remarks on enacting parts

The following bundles are aimed at leading the national legislator in the interpretation of the law, thus collocations and pseudo-technical formulas that affect the text plainness (or the lack of plainness) are here reported.



Corpus A – Enacting Parts (from 3 to 6-grams)	Corpus B – Enacting Parts (from 3 to 6-grams)	Corpus C – Enacting Parts (from 3 to 6-grams)
- in materia di - se del caso - di cui all'articolo - di cui al paragrafo - di cui all'allegato - ai sensi dell'articolo - a norma dell'articolo - corredate di un siffatto riferimento - ai fini della presente direttiva - la procedura di cui all'articolo - direttiva modificata da ultimo dalla direttiva - necessarie per conformarsi alla presente direttiva	- disposizioni di cui - ai sensi del - previsto dall'articolo - di cui all'articolo - di cui al comma - ai sensi dell'articolo - di cui all'allegato - di cui al decreto legislativo - convertito con modificazioni dalla legge - di concerto con il Ministro - di cui al presente articolo - di cui al presente comma - dalla data di entrata in vigore - di cui al decreto del Presidente - le disposizioni di cui all'articolo	- in materia di - di concerto con - di cui all'articolo - di cui al comma - ai sensi dell'articolo - di cui al decreto - di cui al presente - convertito con modificazioni dalla legge - di concerto con il Ministro - di cui al decreto legislativo - di cui al presente articolo - di cui al decreto del Presidente - con decreto del Ministro della Difesa - inserito nella raccolta ufficiale degli atti - le disposizioni di cui all'articolo

Table 11. Legal framing bundles in enacting parts (Italian)

Corpus A – Enacting Parts (from 3 to 6-grams)	Corpus B – Enacting Parts (from 3 to 6-grams)	Corpus C – Enacting Parts (from 3 to 6-grams)
- en el artículo - en el apartado - en el anexo - de las disposiciones - en la presente Directiva - a la presente Directiva - con arreglo al procedimiento - contemplado en el artículo - de conformidad con el procedimiento - cuya última modificación la constituye - de conformidad con el artículo - en el ámbito regulado por - a que se refiere el apartado - a que se refiere el artículo - con arreglo a lo dispuesto en - sin perjuicio de lo dispuesto en	- en el artículo - en el apartado - en materia de - lo previsto en - en los términos previstos - en el Real Decreto - en los términos que - conforme a lo dispuesto - lo dispuesto en el artículo - lo establecido en el artículo - lo dispuesto en el apartado - lo previsto en el artículo - a que se refiere el artículo - a que se refiere el apartado - sin perjuicio de lo dispuesto en - de conformidad con lo dispuesto en	- en el artículo - en el apartado - lo establecido en - lo previsto en - en el apartado anterior - en el párrafo anterior - en la disposición adicional - en los términos previstos - lo dispuesto en el artículo - lo previsto en el artículo - lo establecido en el artículo - de acuerdo con lo establecido - a que se refiere el artículo - a que se refiere el apartado - de acuerdo con lo establecido en - de lo dispuesto en el artículo

Table 12. Legal framing bundles in enacting parts (Spanish)

Corpus A in Italian is more characterised by very frequent 4-gram and 6-gram playing a role for legal framing bundles, thus confirming its greater formulaicity and high lexical density. Comparing A vs. B and C we note that there is a range of similar bundles varying as far as



nouns referring to sections of the law: for instance Italian directives are divided into *capitoli*, *paragrafi*, *articoli* and *lettere*, while national laws comprise *capi*, *titoli*, *articoli* and *commi* (see Mori 2018).

As for Spanish, there is a remarkable coincidence of bundles used across corpora. Furthermore, from the analysis A vs. B/C the Spanish eurolect stands out as an example of uniformity compared with the intralinguistic variability affecting the national laws (see Blini 2018a). For instance, whereas in Corpus A the noun phrase *lo dispuesto* is used to refer to the provisions adopted in the text as a whole, in B and C three variants are attested: *lo dispuesto*, *lo establecido* and *lo previsto*. Thus confirming the greater formulaicity of the Eurolect.

5.3 Deontic and performative bundles

The frequency of deontic modals and semi-modals is relatively high in legislation, because “the fundamental functions of law, that is, to impose duties and confer power [...] are realised through deontic modality – modals and related patterns which convey obligation and permission” (Biel 2017: 157).

The category here proposed concerns lexical bundles which contain prescriptive rules and performative messages, which are sometimes difficult to distinguish due to their common binding character (Mori 2020b). These comprise bundles coding the deontic modality (see Blini 2018a; Mori 2020a) which are typical in legal discourse with minimal cross-corpora variation and occurring together with legal performatives.

5.3.1 Remarks on Preambles

These lexical bundles contain noun and verbal phrases with non-prototypical subjects (such as Member States, directive) and the use of modal verbs⁷ with deontic value and forms bearing a performative effect.

In Italian legislative corpora lexical bundles belonging to this macro-category are absent in preambles of Corpus B and they are extremely limited either A or C.

The lexical bundles expressing the deontic modality in Italian do contain modal verbs especially in the conditional form of *dovere* whereas EU objectives are declared to the addressee in a mitigated way: *la Commissione dovrebbe*; *Stati membri possono*; *gli Stati*

⁷ On the cross-corpora distribution of *potere/dovere* in Italian see Mori (2020a).



membri dovrebbero; la presente direttiva dovrebbe. At the same time there are instances bearing a performative load since what is stated immediately occurs in the legal domain (such as the 5-gram *la presente direttiva si limita*).

In Corpus B this kind of lexical bundles is not attested, while preambles of Corpus C are characterised by few lexical bundles containing technical verbs, such as: *emana il seguente decreto; promulga la seguente legge; il Presidente della Repubblica promulga*.

On the contrary, in Spanish there are quite more detectable lexical bundles falling into this category and the most frequent lexical ones are not featured by the use of conditional modals: forms at the indicative, present and future, are preferred (see Table 13). Whereas in B and C there is a relevant coincidence of lexical bundles, A confirms the greater formulaicity featuring Eurolects.

Corpus A – Preamble (from 3 to 6-grams)	Corpus B – Preamble (from 3 to 6-grams)	Corpus C – Preamble (from 3 to 6-grams)
- la Comisión debe - consulta al Comité - es necesario que - es necesario establecer - los Estados miembros deben - no pueden ser alcanzados - los Estados miembros pueden - los Estados miembros podrán - la presente Directiva no excede - la presente Directiva no debe - la Comunidad puede adoptar medidas - deben aprobarse con arreglo a - se ajustan al dictamen del Comité - deben adoptarse con arreglo al procedimiento - se alienta a los Estados miembros	- se procede a - la ley establece - se establece que - se prevé que - se modifica la Ley - se modifica el texto - se modifica el artículo - se establece la obligación - sabed que las Cortes Generales - se dicta al amparo de - se modifica el texto refundido - se modifica el Real Decreto - vengo en sancionar la siguiente Ley - se dicta al amparo del artículo - se incorpora al ordenamiento jurídico interno	- se procede a - se establece que - se prevé que - la ley establece - se modifica la Ley - se modifica el artículo - se modifica el régimen - se modifica el texto - sabed que las Cortes Generales - se dicta al amparo de - en primer lugar se modifica - vengo en sancionar la siguiente Ley - se ceden a esta Comunidad Autónoma - se dispone la entrada en vigor - dispone la derogación de la Ley

Table 13. Deontic and performative bundles in preambles (Spanish)

5.3.2 Remarks on enacting parts

Enacting Parts, especially in Eurolects, are characterised by chunks of sentences repeated identically (in 5-grams and 6-grams) where the subject is frequently represented by the addressee of EU directives, that is to say each Member State, or the directive itself. The examples collected in Tables 14–15 show features related to prescriptivity of texts as well as



the constitutive performativity of legal utterances with an immediate effect.

Corpus A – Enacting Parts (from 3 to 6-grams)	Corpus B – Enacting Parts (from 3 to 6-grams)	Corpus C – Enacting Parts (from 3 to 6-grams)
- Stati membri possono - Stati membri adottano - Stati membri provvedono - è sostituito dal seguente - Stati membri sono destinatari - sono decise dagli Stati - sono destinatari della presente direttiva - presente direttiva entra in vigore - gli Stati membri provvedono affinché - sono decise dagli Stati membri - Stati membri adottano tali disposizioni - contengono un riferimento alla presente direttiva - gli Stati membri comunicano alla Commissione - gli Stati membri mettono in vigore - essi ne informano immediatamente la Commissione	- non si applicano - sono sostituite dalle seguenti - sono apportate le seguenti - sono inserite le seguenti - è inserito il seguente - è sostituita dalla presente - sono apportate le seguenti modificazioni - sarà inserito nella raccolta ufficiale - emana il seguente decreto legislativo - è fatto obbligo a chiunque spetti - di osservarlo e farlo osservare - è punito con la sanzione amministrativa - devono derivare nuovi o maggiori oneri	- applicano le disposizioni - autorizzato ad apportare - sono sostituite dalle seguenti - sono apportate le seguenti - si applicano le disposizioni - è inserito il seguente - sono inserite le seguenti - è sostituito dal seguente - ha disposto con l'articolo - sono apportate le seguenti modificazioni a - è fatto obbligo a chiunque spetti - autorizzato ad apportare con i propri - farla osservare come legge dello Stato - di osservarlo e di farlo osservare

Table 14. Deontic and performative bundles in enacting parts (Italian)

Corpus A – Enacting Parts (from 3 to 6-grams)	Corpus B – Enacting Parts (from 3 to 6-grams)	Corpus C – Enacting Parts (from 3 to 6-grams)
- se sustituye por - entrará en vigor - irán acompañadas de - se aplicará a - los Estados miembros adoptarán - los Estados miembros comunicarán - los Estados miembros garantizarán - los Estados miembros exigirán - irán acompañadas de dicha referencia - han adoptado la presente Directiva - comunicarán inmediatamente a la Comisión - se sustituirá por el texto - se sustituye por el texto	- será de aplicación - entrará en vigor - se entenderá por - serán de aplicación - queda redactado como sigue - se modifica el artículo - se modifica el apartado - tendrán la consideración de - queda redactado de la siguiente - queda redactado del siguiente modo - se dicta al amparo de - quedará redactado de la siguiente - queda redactado de la siguiente forma - queda redactado de la siguiente manera	- entrará en vigor - queda redactado de - será de aplicación - se añade un - queda redactado como sigue - se modifica el artículo - se modifica el apartado - se da nueva redacción - queda redactado de la siguiente - mando a todos los españoles - se añade un nuevo apartado - se dicta al amparo de - queda redactado en los siguientes términos - queda redactado de la siguiente forma - guarden y hagan guardar esta Ley



siguiente - los Estados miembros velarán por que - los Estados miembros establecerán las modalidades - harán referencia a la presente Directiva	- queda redactado en los siguientes términos - quedará redactado de la siguiente manera	- pasa a tener la siguiente redacción
---	--	--

Table 15. Deontic and performative bundles in enacting parts (Spanish)

Data on lexical bundles in Italian confirm trends of cross-corpora variation already reported (Mori 2020a, 2020b). In Italian there occurs the use of the active indicative present (3rd person singular and plural) which prescribes duties to be accomplished to comply with the EU directive's objectives through the use of non-prototypical subjects (such as Member States and directive). In case of inanimate subjects, the indicative present acquires a performative value within declarative utterances where some cotextual cues (such *presente, tale, come segue*, etc.) are also reported.

Furthermore, in lexical bundles in EU directives the active diathesis is used to a higher extent and this confirms a general distribution active/passive across corpora: in Corpus A in Italian active forms have a frequency which is doubled compared to passive ones. In Corpus B and C the passive is over-represented and forms at the indicative passato prossimo are mainly used to assess the legal performative value to the provisions of the law.

In Spanish what is more remarkable is the recurrence of verbs in the future indicative, especially in Corpus A. As a matter of fact, in the EU context, the use of the future tense with prescriptive value is recommended in the *Guía práctica común del Parlamento Europeo, del Consejo y de la Comisión para la redacción de textos legislativos de la Unión Europea* (2015). Moreover, Table 15 provides evidence of opposing trends between Spanish Eurolect, on the one hand, and national legislative varieties on the other. In Corpus A the range of phrasemes referring to textual changes is limited to *se sustituye/sustituirá por el texto siguiente*, while in B e C a greater variability is attested. These are the n-gram variants recorded:

se da nueva redacción
queda modificado como sigue
queda/quedará redactado como sigue
queda/quedará redactado del siguiente modo
queda/quedará redactado de la siguiente forma
queda/quedará redactado de la siguiente manera
queda/quedará redactado en los siguientes términos
pasa/pasará a tener la siguiente redacción



6. Conclusions

The context of EU legislative multilingualism and the practice of the interlinguistic translation of the European legislation favour the adoption of linguistic solutions that facilitate the comparability of the twenty-four language versions. In case of EU directives, the consequent rewriting process when implementing them at national level produces an “hybridisation” between two legal cultures and two drafting models of reference, European and national (see Mori 2019a). Therefore, Corpus B stands out as a distinct sub-genre characterised by its own variety: distinct from Eurolect (Corpus A) and from national legal variety (Corpus C). This assumption is here confirmed for Italian. Data on lexical bundles in Spanish do not confirm this trend of intralinguistic variation (Corpus B vs. C), as far as phraseology expressed via lexical bundles is concerned.

Quali-quantitative analyses here carried out, from a contrastive perspective, allowed us to categorise the most frequent lexical bundles focusing on their discursive functions as follows:

1. *Content bundles*, mainly composed of noun phrases that designate legal authority with a 3-gram core structure (core content bundles) playing a main referential function, with a higher degree of formulaicity in Eurolects for both languages.
2. *Legal framing bundles*, composed of prepositional phrases which signal the legal validity anchoring the text to the legal frame, either EU or national, playing both referential and conative function with cross-language analogies related to different textual sections (preambles vs. enacting parts).
3. *Deontic and performative bundles* comprise verbal phrases and clauses complying with the predominant conative function of regulative texts. These frequent bundles have a note-worthy salience in featuring in the legal genre and variational trends do emerge from the cross-corpora comparisons. Within this macro-categories cross-language differences are more relevant. This may rely upon the fact that these bundles are those playing the discursive function at a greater extent and cross-cultural pragmatic aspects are more rooted.

Summing up, our qualitative analysis assisted by the use of parallel and comparable corpora allowed us to focus on similarities and differences between languages considering text internal variation related to different sections with specific communicative functions (preambles vs. enacting parts). Moreover, functional categorisation of lexical bundles here outlined provides further evidence to advance the variational study on Eurolects by offering a new contrastive perspective (Italian vs. Spanish) on dynamics of variation in correlation with



Blini, Lorenzo/Laura Mori: Structural and functional analysis of lexical bundles in Eurolects and national legislative varieties: a Spanish-Italian contrastive study.

different legal contexts (supranational vs. national). In conclusion, our study proposes a replicable approach to be used in the analysis of lexical-syntactic patterns featuring different language varieties in the legal domain and for their discursive interpretation.

References

- Biber, Douglas/Susan Conrad/Viviana Cortes (2003): Lexical bundles in speech and writing: an initial taxonomy. In: A. Wilson, P. Rayson, & T. McEnery (eds.): *Corpus linguistics by the lune: A festschrift for Geoffrey Leech*. Frankfurt/Main: Peter Lang, 71–92.
- Biber, Douglas/Federica Barbieri (2007): Lexical Bundles in University Spoken and Written Registers. In: *English for Specific Purposes*, vol. 26, no. 3, 263–286.
- Biel, Lucja (2017): *Lost in the Eurofog. The Textual Fit of Translated Law*. Frankfurt/Main: Peter Lang.
- Biel, Lucja (2018): Lexical Bundles in EU Law: The Impact of Translation Process on the Patterning of Legal Language. In: Stanislaw Goźdz-Roszkowski/Gianluca Pontrandolfo (eds.): *Phraseology in Legal and Institutional Settings: A Corpus-based Interdisciplinary Perspective*. Abingdon: Routledge, 11–26.
- Biel, Lucja, et al. (2019): The formulaicity of translations across EU institutional genres. A corpus-driven analysis of lexical bundles in translated and non-translated language. In: *Translation Spaces*, vol. 8, no. 1, 67–92.
- Blini, Lorenzo (2018a): Observing Eurolects: the case of Spanish. In: Laura Mori (ed.): *Observing Eurolects. Corpus analysis of linguistic variation in EU law*. Amsterdam: John Benjamins, 330–367.
- Blini, Lorenzo (2018b): Entre traducción y reescritura: eurolecto español y discurso legislativo nacional. In: *Cuadernos AISPI*, no. 12, 15–35.
- Blini, Lorenzo (2024): Armonización lingüística y claridad legislativa: una comparación en el nivel sintáctico entre el eurolecto español y el castellano jurídico nacional. In: *Archiletras Científica*, no. 12, 95–109.
- Blini, Lorenzo/Laura Mori (2019): Los europeísmos en la lexicografía española desde la perspectiva sociolingüística. In: *Artifara*, no. 19, 249–267.
- Canavese, Paolo/Laura Mori (2024): Testing the hypothesis of “translation as a catalyst for plain legislation” on the syntactic level: a comparison of different varieties of legislative Italian. In: *Comparative Legilinguistics*, no. 58, 113–150.
- Felici, Annarita/Laura Mori (2019): Corpora di italiano legislativo a confronto: dall’Unione europea alla Cancelleria svizzera. In: Bruno Moretti et al. (eds.): *Le tendenze dell’italiano contemporaneo rivisitate. Atti del LII Congresso Internazionale di Studi della Società di Linguistica Italiana* (Berna, 6-8 settembre 2018). Milano, Officinaventuno, 287–304.
- Guía práctica común del Parlamento Europeo, del Consejo y de la Comisión para la redacción de textos legislativos de la Unión Europea (2015). Luxemburgo: Oficina de Publicaciones de la Unión europea.
- Mattila, Heikki E.S. (2013): *Comparative legal linguistics: Language of law, Latin and modern lingua francas* (2nd ed.). Farnham: Ashgate.
- Menza, Salvatore/Laura Mori (2022): La metaforizzazione del linguaggio giuridico per la costruzione discorsiva dell’identità sovranazionale nel diritto dell’Unione europea. In: Salvatore Menza/Valentina Petralia (eds.): *L’integrazione europea: tra sfide politiche, vincoli giuridici e legittimazione democratica*. Milano: Le Monnier Università, 9–36.



Blini, Lorenzo/Laura Mori: Structural and functional analysis of lexical bundles in Eurolects and national legislative varieties: a Spanish-Italian contrastive study.

Mori, Laura (2018): Observing Eurolects: The case of Italian. In: Laura Mori (ed.): Observing Eurolects. Corpus analysis of linguistic variation in EU law. Amsterdam: John Benjamins, 200–242.

Mori, Laura (2019a): Dall'armonizzazione all'ibridazione nei testi legislativi: evidenze linguistiche e manifestazioni interculturali nell'italiano tradotto. In: EntreCulturas. Revista de Traducción y Comunicación Intercultural, no. 10, 377–392.

Mori, Laura (2019b): Configurazioni strutturali e funzioni discorsive dei lexical bundles nella costruzione testuale delle leggi italiane nazionali e di derivazione europea. In: Raffaella Bombi (ed.): Saggi interlinguistici e metalinguistici. Roma: Il Calamo, 79–88.

Mori, Laura/Benedikt Szmrecsanyi (2020): Mapping eurolects: an aggregate perspective on similarities between legislative varieties. In: Languages in Contrast, vol. 21, no. 2, 186–216.

Mori, Laura (2020a): La distribuzione dei verbi modali in testi legislativi europei e italiani. Uno studio corpus-based sulla variazione intralinguistica di dovere e potere. In: AIQN-Linguistica, no. 9, 141–163.

Mori, Laura (2020b): La rappresentazione di scenari deontici e l'espressione della performatività nell'italiano delle leggi: dal diritto europeo alla legislazione nazionale. In: Linguistica e Filologia, no. 40, 45–97.

Tognini Bonelli, Elena (2001): Corpus linguistics at work. Amsterdam: John Benjamins.