

# Terminological hybridity in institutional legal translation

## A corpus-driven analysis of key genres of EU and international law

Fernando Prieto Ramos & Giorgina Cerutti

University of Geneva

The analysis of domain-specific terminology is essential for characterizing specialized discourses, and emerges as a useful means of measuring the thematic hybridity of law and legal translation in particular. This paper accordingly presents a large-scale mapping of terminological and phraseological features in a multi-genre corpus that was built as part of the LETRINT project on institutional legal translation. The corpus-driven analysis focuses on the density of legal terminology and phraseology, on the one hand, and that of terminology of other specialized domains, on the other, in nine genres that are considered representative of three central legal functions (law-making, compliance monitoring and adjudication) in three international settings (the European Union, the United Nations and the World Trade Organization). The comparative examination of density scores provides empirical evidence of the common core features of the selected genres, and reveals variations based on institutional thematic focus, primary legal function and genre specificities. These insights nuance our understanding of international legal discourses and domain specialization in institutional translation.

**Keywords:** terminology, phraseology, hybridity, international legal discourse, legal genre, institutional legal translation

### 1. Background

Legal texts are thematically diverse due to the broad scope of the law's many branches and intersecting conceptual networks, which deal with virtually every aspect of life. As noted by Mayoral Asensio, these branches of law and other fields of knowledge are interconnected; "legal texts rarely present a single subject matter

or thematic frame and almost all of them offer other information, technical, medical, etc.” (Mayoral Asensio 2007: 51). This has significant implications for drafters and translators. For example, drafting or translating financial or technical regulations may involve delving into finance or technical domains as much as employing the terminology of the relevant branches of law and the phraseological conventions of the legal genre at hand. In other words, each legal genre can be regarded as a communicative framework built around distinctive legal discursive features but able to accommodate knowledge of an endless number of domains (Prieto Ramos 2019: 34).

Terminology and phraseology therefore emerge as key discursive features and useful indicators of thematic hybridity. Legal concepts and formulations are the backbone of discourses that perform the primary functions of legal genres, and often co-exist with terminology of other domains. This also applies to international law and multilingual legal texts produced by international organizations through institutional translation. They deal with a vast array of subjects, including, among many others: human rights, environmental issues, public health, trade, finance or migration. To what extent does this thematic plurality vary between institutions and areas of international law, and how is it reflected in institutional discourses? Despite the advances of corpus linguistics, the accessibility of institutional texts and the relevance of this topic to inform thematic specialization and training among translators, the field lacks empirical data to answer the question.

To fill this gap, this paper presents a large-scale analysis of terminological and phraseological features of a multi-genre corpus that was built as part of the LETRINT project on institutional legal translation.<sup>1</sup> The paper focuses on the density of legal terminology and phraseology, on the one hand, and that of terminology of other specialized domains, on the other, in a selection of genres that are considered representative of institutional legal translation in three settings: the European Union’s (EU) main institutions, the United Nations (UN) and the World Trade Organization (WTO).

In Translation Studies (TS), the hybridity of institutional texts has, for the most part, been analyzed as a matter of deviation from standard use due to the contact and influence between drafters of diverse language and cultural backgrounds. As a result, institutional varieties of the official languages develop their own hybrid features enhanced through translation. This phenomenon has been examined particularly in the context of the EU institutions (e.g. Tirkonnen-Condit 2001; Trklja 2018). It affects not only target languages, but also, more

---

1. “Legal Translation in International Institutional Settings: Scope, Strategies and Quality Markers”, a Consolidator Grant project led by the first author with the support of the Swiss National Science Foundation (<https://transius.unige.ch/letrint>).

remarkably, the language of the source texts, most often English as a *lingua franca* (e.g. Hewson 2013; Sandrelli 2018), but also French in the case of the Court of Justice of the EU (CJEU) (e.g. McAuliffe 2011). Thematic or domain hybridity, however, has rarely been examined in TS, with the exception of Fontanet's (2018) study of the fluctuations of legal and technical terminology and sub-technical vocabulary in a small corpus of cooperation agreements, safety documents and contract adjudication proposals translated at the European Organization for Nuclear Research (CERN).

Our corpus-driven study sheds light on the wider spectrum of institutional legal functions and translation by exploring a comprehensive set of terminological and phraseological categories in three major settings of multilingual text production. The approach is further explained in Section 2, while the results are outlined in Section 3.

## 2. Methodology

As mentioned above, the LETRINT project aims to provide empirical data on the scope and features of institutional legal translation, overcoming the limitations of previous research on specific contexts or genres, most typically legislation. To this end, based on the legal contextualization of supranational and intergovernmental institutional missions and multilingual text production, three central legal functional categories were used as common pillars for the selection of settings and genres representative of institutional legal discourses and translation: law- and policy-making, compliance monitoring and adjudication (Prieto Ramos 2014a, 2019). The three organizations selected are considered exemplars of these legal functions and of the translation of international legal texts: the UN as the main umbrella intergovernmental organization, including its International Court of Justice (ICJ); the WTO, a specialized organization that deals with a wide range of trade-related rules and settles disputes regarding their application; and the four main institutions of the EU, the most important supranational organization in the world (the European Commission, the European Parliament, the Council of the EU and the CJEU).

An ambitious categorization of all textual production of the above institutions in three languages (English, French and Spanish) during three years (2005, 2010 and 2015) provided the foundation for a comprehensive characterization of monolingual and multilingual text volumes, and for identifying the most prominent translated genres within each legal functional category, both in quantitative and qualitative terms (Prieto Ramos and Guzmán 2021). In turn, this full mapping enabled the creation of several (derived) corpora through a multi-layered sequen-

tial approach (see Prieto Ramos et al. 2019), and led to the design of the LETRINT 1+ corpus used in this study.

This corpus was built in order to manually annotate discourse features and related translation difficulty levels, and subsequently analyze translation patterns and quality parameters. With a view to ensuring feasibility, representativeness, overall balance and cross-genre comparability, stratified sampling and fragmentation techniques were applied in the light of the previous mapping and according to tailored quantitative and qualitative criteria (for further details, see Prieto Ramos et al. 2019: 108–112).<sup>2</sup> The resulting components of LETRINT 1+, including the number of texts and tokens of its nine sub-corpora (one per institutional setting and legal function), are listed in Table 1.

Overall, the LETRINT 1+ corpus is composed of 256 texts and 752,061 tokens.<sup>3</sup> A cap of approximately 90,000 tokens and a minimum of 12 texts per sample were applied to all the genres selected for the nine sub-corpora, except for the WTO's binding decisions (WTO-1), for which the sample was exceptionally reduced to 32,121 tokens. This is the total retrieved from the corresponding stratum of the parent LETRINT 1 corpus for this functional category and institution (which, in turn, included one third of the total volume per genre and year); given its limited volume, no further sampling was necessary for this LETRINT 1+ sub-corpus. The other sub-corpora represented between 1.41% (EU-1) and 37.14% (UN-3) of the total number of translated texts of the respective genre initially compiled for the three years covered by the project (see Table 1). The larger the initial volume of a genre, the smaller the proportion represented by the corresponding sub-corpus in LETRINT 1+ and the more challenging the stratified sampling conducted. While quantitative adequacy exceeded previous recommendations on genre-specific corpora (e.g. Biber 1993; Oostdijk 1991),<sup>4</sup> qualitative considerations, i.e. the analyses of variation parameters per genre as mentioned

---

2. In order to enhance representativeness, variation parameters were analyzed and modulated as qualitative criteria for stratified sampling for each genre, including: (1) genre sub-classifications according to body or procedural specificities (e.g. sub-types of regulations or court proceedings); (2) subject matters; and (3) countries involved in monitoring and adjudication genres in particular. This process, which was paramount for balancing the examination of specialized discourses, was supported by verifying the related data available from institutional sources, e.g. subject matter metadata in EUR-Lex and statistics on court proceedings in institutional reports.

3. An infographic on the LETRINT corpora, as well as a corpus query interface developed for the project (LETRINT-Q), can be freely accessed at: <https://transius.unige.ch/en/research/letrint/corpora>.

4. The first author considers 10 texts and 1,000 words as adequate, while Oostdijk suggests 20,000 words as a reasonable sample for genre analysis.

**Table 1.** LETRINT 1+ sub-corpora per organization and primary legal function

Law- and policy-making					Compliance monitoring			Adjudication		
Texts <sup>*</sup> Tokens					Texts <sup>*</sup> Tokens			Texts <sup>*</sup> Tokens		
EU	Sub-corpus	EU-1	66 of 4667 (1.41%)	89,885	EU-2	24 of 151 (15.89%)	89,947	EU-3	17 of 984 (1.73%)	90,021
	Genre	Regulations			European Commission reports			CJEU judgments		
UN	Sub-corpus	UN-1	52 of 851 (6.11%)	90,028	UN-2	25 of 261 (9.58%)	89,911	UN-3	13 of 35 (37.14%)	90,087
	Genre	Resolutions			UN human rights treaty bodies' concluding observations			ICJ Judges' opinions		
WTO	Sub-corpus	WTO-1	32 of 118 (27.12%)	32,121	WTO-2	15 of 62 (24.19%)	90,014	WTO-3	12 of 47 (25.53%)	90,047
	Genre	Decisions by the Ministerial Conferences and the General Council			Reports issued by the Trade Policy Review Mechanism Secretariat			Dispute settlement reports issued by panels and the Appellate Body		

\* The percentage of sample texts over the total number of texts compiled for each genre (before stratified sampling) is indicated between brackets.

above, were key to enhancing representativeness.<sup>5</sup> In fact, as noted by Koester (2010: 68), among others, this aspect of corpus design is “more important than the actual size” of a specialized corpus.

The totals of tokens are based on the wordcounts of the source texts. They reflect the predominance of English as a drafting language in international institutional communication, except for the CJEU, where the judgments are drafted in the working language of the institution, French (EU-3). For the purpose of our study, this would have no impact on the examination of thematic hybridity. On the contrary, it proved highly relevant for testing the applicability of our approach in more than one language.

5. To double-check adequacy and representativeness, we calculated LETRINT 1+ closure / saturation values using the ReCor tool (Corpas Pastor and Seghiri Domínguez 2006), and we conducted comparative analyses of the TTR (type / token ratios) and log-likelihood of the most frequent terms in LETRINT 1 and LETRINT 1+. The results of these calculations, which were very satisfactory, will be presented elsewhere.

The manual annotation of the corpus proved time-consuming (over more than a year), not only because of the amount of text and detailed analysis required, but also, crucially, due to the methodological safeguards adopted to ensure the reliability and consistency of the process, and ultimately the validity and relevance of the results. All sub-corpora were divided and annotated by two teams of two annotators, one French speaker and one Spanish speaker, using the UAM CorpusTool (O'Donnell 2019). The bilingual composition of each team was paramount to detect any potential culture-bound divergence in the annotation process. The teams were modified only for the annotation of two genres in light of translation directionality and translation-related annotation purposes of the LETRINT project: UN-3, annotated by the French-language annotators only (as ICJ Judges' opinions are not translated into Spanish); and EU-3, annotated by Spanish-language annotators only (as the translation scenario considered for CJEU judgments was from originals drafted in French). Once again, this did not distort the validity of the approach for the analysis of thematic hybridity. Furthermore, UN-3 and EU-3 were among the last to be annotated and thus benefited from the classification of a great volume of previously annotated units.

The four annotators were selected according to specialization criteria. They all had legal translation qualifications and professional translation experience, including legal and institutional translation. A gradual validation approach was established so that all annotations were double-checked by the second annotator of each team, and then cross-checked against the annotation results of the other team in order to detect and resolve any discrepancies. Finally, LETRINT's principal investigator revised all the annotation results for their overall coherence and consistency, and proposed team revisions where appropriate. The objective was to reach consensus as a team.

This process also entailed revising, and where necessary adjusting, the boundaries between the annotation categories considering the granularity provided by the corpus itself and the analysis of borderline cases. As a result, the terminological and phraseological annotation categories developed by LETRINT underwent several refinements as the work progressed, until the full taxonomy became stable. This includes several categories of legal terminology (LEG-T), terminology of other specialized domains (SPEC-T), institutional titles (ITT) and legal phraseology (LEG-P), as outlined in Table 2.

In line with the goals of the project regarding legal translation patterns and quality, a particular focus was put on various categories of **legal terminology** (LEG-T) according to their origin and legal singularity, building on a previous analysis of legal terminology in institutional settings (Prieto Ramos 2014b: 128–129). The resulting four categories include legal terms established and standardized in the international legal order (LEG-INT), national system- or

**Table 2.** Taxonomy of discourse features annotated in the LETRINT 1+ corpus**Legal terminology (LEG-T)**

LEG-GEN	Terms that are used to refer to concepts that are perceived as generic and common to multiple legal systems (e.g. “appeal”, “legislation”, “provision”), including terms that originated in a particular legal tradition and were borrowed in international law (e.g. “prima facie evidence”).
LEG-INT	Terms created in the international legal order and recognized as established terminology within the specific scope of competence of a particular organization according to international legal instruments (e.g. “extended continental shelf”, “tariff escalation”, “acquis communautaire”).
LEG-NAT-SIN	Terms designating singular concepts that are specific to national legal systems or traditions, for example, names of judicial institutions, types of legislation or other legal singularities (e.g. “Beneš Decrees”, “Chancellor of Justice”, “magistrates’ court”).
LEG-NAT-GEN	Names of national, regional or local bodies, instruments or positions that are common to multiple legal systems and traditions (e.g. “Constitution”, “Parliament”, “Prime Minister”).

**Terminology of other specialized domains (SPEC-T)**

ECO-T	Terms related to economics, employment, trade and business (e.g. “export”, “labour market”, “supply”).
FIN-T	Terms used in finance, including fund management, banking and budgets (e.g. “default fund”, “financial market”, “stock”).
POL-T	Terms referring to politics, government structures, social and other public policies, and administrative matters (e.g. “civil society organization”, “decision-making procedure”, “family reunification”).
SCI-T	Terms used in the natural sciences, including physics, chemistry and biology (e.g. “bovine spongiform encephalopathy”, “ecosystem”, “genetically modified organism”).
TEC-T	Terms that designate technical applications of science, including machines, processes and materials used in industry, transport and communications (e.g. “chemical tanker”, “data collection system”, “lateral flow immunoassay”).

**Institutional titles (ITT)**

ITT1	Titles of international legal instruments, official documents or cases (e.g. “General Agreement on Trade in Services (GATS)”, “Convention on the Rights of Persons with Disabilities”).
ITT2	Established names of institutional bodies, positions and institutional events, programmes or processes (e.g. “European Food Safety Authority”, “Millennium Development Goals”).

Table 2. (continued)

Legal phraseology (LEG-P)

LEG-P1	Phrasemes, including prepositional phrases and other lexical collocations, that characterize legal discourses (e.g. “pursuant to”, “without prejudice to”, “within the meaning of”).
LEG-P2	Established expressions or formulas that can be identified as genre conventions and contribute to primary text functions, often of a performative nature (e.g. “shall enter into force on”, “decides to remain actively seized of the matter”).

tradition-specific terms (including singular and non-singular denominations, respectively, LEG-NAT-SIN and LEG-NAT-GEN), and more “generic” terms commonly used in legal texts across multiple legal systems (LEG-GEN). The categories within **terminology of other specialized domains (SPEC-T)** encompass key areas of expertise found in international institutional settings: economics, including employment, trade and business (ECO-T); finance and budgetary matters (FIN-T); political, social and administrative affairs (POL-T); science (SCI-T) and technical fields (TEC-T).

As in any categorization process, it was important to establish common criteria to deal with fuzzy boundaries and risks of overlap, as only one category was assigned to each unit. In the case of concepts of a multi-dimensional nature covered by a branch of law (e.g. migration terms with social and legal dimensions, or trade concepts used in commercial exchanges and business law), it was the area of specialization that prevailed over the legal categorization, unless the concept could be primarily considered as a legal construct. For example, “tariff” is associated to trade (ECO-T), whereas “tariff escalation” falls within the realm of international trade law (LEG-INT), as it emerged in that regulatory context. In turn, the categorization as LEG-INT sometimes entailed research on the origin of the term or the evolution of its semantics when adopted in the international legal system.

When a more extended analysis became necessary on a particular unit of this or any other category, a procedure was agreed to compile all the relevant information in terminological research records to support the final annotation decision. These records included: (a) a mapping of contextualized occurrences of the annotation unit in and beyond the LETRINT corpora, and, where relevant, a diachronic overview of use frequency retrieved through Google Ngram Viewer; (b) the contents provided for the term by institutional terminological databases, if any; and (c) additional information on the origin, history and related semantic network of the term or phraseme. Overall, 490 analytical records were created out of the 19,527 units annotated. As the analysis usually expanded to units that



were closely related to each term or phraseme investigated (an average of 2.5 per record), the total number of units covered by the records amounts to 1,222 (including 1,121 terms, 70 phrasemes and 31 institutional titles) or 6.26% of the units annotated.

Another important decision was to isolate **institutional titles (ITT)** as separate categories. From a translation perspective, they are treated as established units verified and reproduced as such in institutional texts. From a thematic angle, these units can be considered a hybrid group of multi-word units associated to LEG-INT, as they refer to legal and administrative institutional structures, including titles of legal instruments and other documents (ITT<sub>1</sub>) and institutional bodies and positions (ITT<sub>2</sub>). While they may refer to a diversity of topics and contain other domain-specific terminology (e.g. “Directive of the European Parliament and of the Council amending Directive 1999/62/EC on the charging of heavy goods vehicles for the use of certain infrastructures” or “TRIPS Council”), these subordinate terms (underlined in the examples) would also appear and be annotated in the body of the text, whereas the title as a whole would still be treated as a *sui generis* category for our discourse analysis and annotation.

Finally, as regards **legal phraseology (LEG-P)**, building on the parameters for defining phraseologisms put forward by Gries (2008: 4–6)<sup>6</sup> and on commonly-accepted notions of word combination lexicalization and formulaicity in legal phraseology (e.g. Kjaer 2007; Ruusila and Lindroos 2016), a pragmatic corpus-driven approach was tested and adopted for distinguishing between two categories: “simple” phrasemes regularly used in legal discourses (LEG-P<sub>1</sub>), and more complex genre-specific formulaic expressions, usually with a performative function (LEG-P<sub>2</sub>). In the case of the latter, additional research was required during the annotation process in order to confirm the formulaicity and pragmatic role of certain expressions.

Overall, the team consensus-building process employed for this large-scale manual annotation represented a unique, yet challenging, opportunity. As in any process of this kind, the categorization boundaries and implementation criteria agreed for the annotation work necessarily involved decisions that depend on research priorities and theoretical and methodological choices. The safeguards adopted guaranteed a harmonized approach for systematic annotation and reliable comparability of results between sub-corpora. The various inter-annotator verifications kept the risk of inconsistency and deviation to the minimum, as confirmed in the final revisions of the annotation results.

---

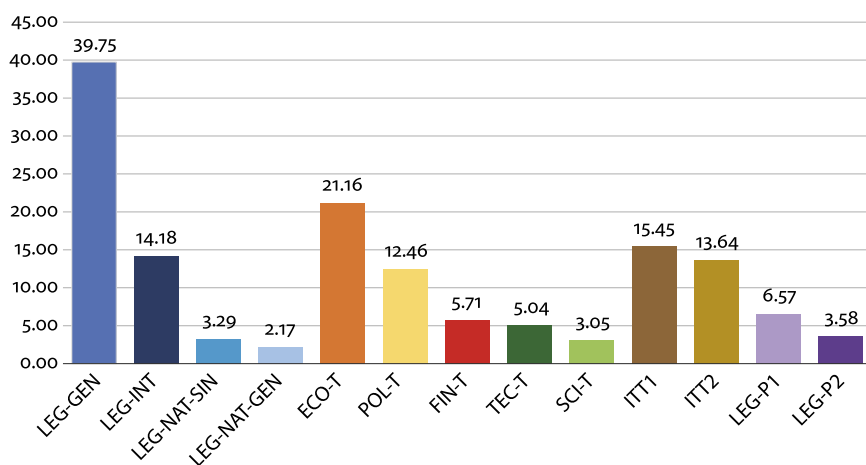
6. He defines a phraseologism as “the co-occurrence of a form or a lemma of a lexical item and one or more additional linguistic elements of various kinds which functions as one semantic unit in a clause or sentence and whose frequency of co-occurrence is larger than expected on the basis of chance” (Gries 2008: 6).

### 3. Annotation results

The distribution of the annotated units is examined in this section. The frequency results obtained through manual annotation were normalized to 1000 tokens, and descriptive statistics were used for their analysis. After a first overview of the data, the results will be presented for each set of institutional sub-corpora and then contrasted per primary legal function in order to identify the most relevant patterns and singularities.

#### 3.1 Overall distribution of terminological features

As represented in Figure 1 and Table 3, the LETRINT 1+ corpus shows a significant degree of terminological hybridity, with a global distribution of 40.66% of legal terms (59.39 annotations per 1000 tokens) and 32.47% of terminology of other specialized domains (47.42 annotations per 1000 tokens), while institutional titles and legal phraseology account for almost 20% and 7% of annotations, respectively (29.09 and 10.15 annotations per 1000 tokens). Unsurprisingly, the density of LEG-GEN terms stands out within legal terminology, with 27.22% of all annotations, followed by LEG-INT (9.71%) and LEG-NAT (2.25% of LEG-NAT-SIN and 1.49% of LEG-NAT-GEN). These data must be read in conjunction with the separate categories ITT1 (10.58%) and ITT2 (9.34%) as important referential units associated to LEG-INT within institutional texts. The combined proportion of these three categories of institution-bound units amounts to almost 30%.



**Figure 1.** Density of annotated categories in the LETRINT 1+ corpus (annotations normalized to 1000 tokens)

**Table 3.** Distribution of annotations in LETRINT 1+ (frequencies normalized to 1000 tokens)

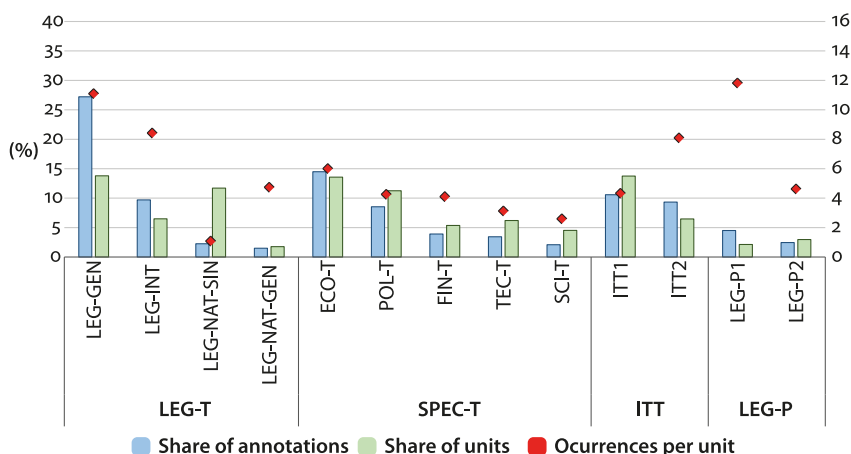
Category	Annotations per 1000 tokens	Percentage
LEG-GEN	39.75	27.22%
LEG-INT	14.18	9.71%
LEG-NAT-SIN	3.29	2.25%
LEG-NAT-GEN	2.17	1.49%
<b>LEG-T</b>	<b>59.39</b>	<b>40.66%</b>
ECO-T	21.16	14.49%
POL-T	12.46	8.53%
FIN-T	5.71	3.91%
TEC-T	5.04	3.45%
SCI-T	3.05	2.09%
<b>SPEC-T</b>	<b>47.42</b>	<b>32.47%</b>
ITT <sub>1</sub>	15.45	10.58%
ITT <sub>2</sub>	13.64	9.34%
<b>LEG-P</b>	<b>29.09</b>	<b>19.92%</b>
LEG-P <sub>1</sub>	6.57	4.50%
LEG-P <sub>2</sub>	3.58	2.45%
<b>ITT</b>	<b>10.15</b>	<b>6.95%</b>
<b>TOTAL</b>	<b>146.05</b>	<b>100.00%</b>

With 14.49% of the annotations (21.16 per 1000 tokens), ECO-T is the most frequent of non-legal categories, particularly because of its predominance in the WTO sub-corpora and, to a lesser extent, in the EU sub-corpora. The reverse applies to the second most frequent SPEC-T category, POL-T, with 8.53% of the annotations (12.46 per 1000 tokens) and a higher density in the UN texts. The remaining non-legal terminological categories are significantly less common: FIN-T (3.91%), TEC-T (3.45%) and SCI-T (2.09%). Finally, within phraseological units, LEG-P<sub>1</sub> registered approximately twice as many annotations (6.57 per 1000 tokens or 4.50%) as LEG-P<sub>2</sub> expressions (3.58 per 1000 tokens or 2.45% of annotations). As the corpus revealed, while simple phrasemes are commonly used to articulate legal discourses, LEG-P<sub>2</sub> formulations constitute the most genre-bound feature. This was confirmed by sub-corpus variations, as discussed in the next sub-sections.

If we consider the breakdown of annotated units disregarding the number of occurrences per unit, i.e. counting each unit only once within a text, the above figures change significantly for various categories. Examining the units annotated against their total annotations yields a very revealing ratio of occurrences per unit (see Table 4 and Figure 2). The average number of occurrences per unit annotated in the entire corpus is 5.63. LEG-P1 and LEG-GEN (11.82 and 11.10 annotations per unit), which can be found as common lexical resources in all legal genres, registered the highest recurrence ratio per unit. In practice, this means that the proportion of annotations for these categories (respectively, 4.50% and 27.22% of the LETRINT 1+ total) doubles their relative weight in unit counts disregarding intra-textual repetitions (2.14% and 13.80%, respectively). They are followed by international legal terms (LEG-INT) and names of institutional bodies, positions and procedures (ITT2), with 8.42 and 8.10 annotations per unit, respectively. These results reflect the repetition of such organization-specific referents to situate institutional matters within their legal and structural frameworks.

**Table 4.** Recurrence per annotated unit based on total units and their annotations for each category

Category	N° of units	% of units	Annotations	Ratio of occurrences per unit
LEG-GEN	2,694	13.80%	29,895	11.10
LEG-INT	1,267	6.49%	10,665	8.42
LEG-NAT-SIN	2,284	11.70%	2,471	1.08
LEG-NAT-GEN	345	1.77%	1,635	4.74
<b>LEG-T</b>	<b>6,590</b>	<b>33.75%</b>	<b>44,666</b>	<b>6.78</b>
ECO-T	2,649	13.56%	15,914	6.01
POL-T	2,197	11.25%	9,369	4.26
FIN-T	1,047	5.36%	4,294	4.10
TEC-T	1,211	6.20%	3,790	3.13
SCI-T	882	4.52%	2,296	2.60
<b>SPEC-T</b>	<b>7,986</b>	<b>40.90%</b>	<b>35,663</b>	<b>4.47</b>
ITT1	2,685	13.75%	11,617	4.33
ITT2	1,266	6.48%	10,261	8.10
<b>ITT</b>	<b>3,951</b>	<b>20.23%</b>	<b>21,878</b>	<b>5.54</b>
LEG-P1	418	2.14%	4,941	11.82
LEG-P2	582	2.98%	2,693	4.63
<b>LEG-P</b>	<b>1,000</b>	<b>5.12%</b>	<b>7,634</b>	<b>7.63</b>
<b>TOTAL</b>	<b>19,527</b>	<b>100%</b>	<b>109,841</b>	<b>5.63</b>



**Figure 2.** Proportions of units and annotations for each category over the total units and annotations, and their corresponding ratios of occurrences per unit

As for terminological categories of other specialized domains, the recurrence per unit decreases in the same order as the frequency of annotations per category: only ECO-T scores above average with 6.01 occurrences per unit, while the recurrence ratio falls to around 3 annotations per unit in the case of TEC-T (3.13) and SCI-T (2.60). The lowest recurrence per unit, however, is that of LEG-NAT-SIN, due to the marginal repetition of singular terms of this category (1.08).

Overall, these figures suggest that, at face value (i.e. without considering unit repetitions and context variations for translation decision-making), and excluding established institutional titles, terminology of non-legal domains (40.90% of annotated units) may trigger more translation-oriented research than the combination of legal terminology and phraseology (33.75% of LEG-T units and 5.12% of LEG-P units). Furthermore, the pronounced disparity between the share of LEG-NAT-SIN units (11.70%) and their annotations (2.25%), the largest positive difference for the first indicator among all the categories, suggests that, although these units are less often repeated across texts, they may entail as much terminological research as LEG-GEN or other categories in the specific texts where they occur. This depends on translation difficulty levels, something that is addressed by a separate LETRINT study (Prieto Ramos and Cerutti 2021).

### 3.2 Distribution per institutional setting

The results per set of institutional sub-corpora point to various commonalities and differences in the distribution of the annotated categories. Overall, the WTO sub-corpora are the richest from a terminological perspective, with 133.10 units

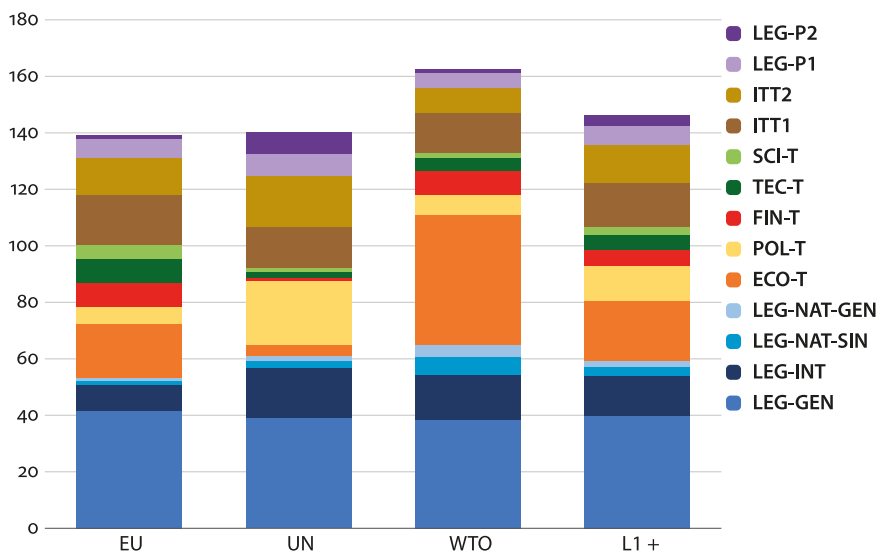
and 162.40 annotations per 1000 tokens, followed by the EU sub-corpora (100.62 units and 139.05 annotations per 1000 tokens) and the UN's (92.34 units and 140.21 annotations per 1000 tokens). The degree of hybridity is comparable in all the institutions, but with significant variations regarding non-legal domains in particular (see Table 5 and Figure 3).

**Table 5.** Distribution of annotated features in the EU, UN and WTO sub-corpora, as well as in LETRINT 1+ as a whole (frequencies normalized to 1000 tokens and their percentages)

	EU	EU%	UN	UN%	WTO	WTO%	L1+	L1+%
LEG-GEN	41.45	29.81%	39.23	27.98%	38.25	23.55%	39.75	27.22%
LEG-INT	9.21	6.62%	17.78	12.68%	15.93	9.81%	14.18	9.71%
LEG-NAT-SIN	1.76	1.27%	2.23	1.59%	6.57	4.05%	3.29	2.25%
LEG-NAT-GEN	0.75	0.54%	1.84	1.31%	4.42	2.72%	2.17	1.49%
LEG-T	53.16	38.23%	61.08	43.56%	65.17	40.13%	59.39	40.66%
ECO-T	19.10	13.74%	3.90	2.78%	45.74	28.16%	21.16	14.49%
POL-T	6.37	4.58%	22.56	16.09%	7.34	4.52%	12.46	8.53%
FIN-T	8.22	5.91%	1.18	0.84%	8.29	5.10%	5.71	3.91%
TEC-T	8.52	6.13%	1.93	1.38%	4.57	2.81%	5.04	3.45%
SCI-T	5.24	3.77%	1.69	1.21%	2.00	1.23%	3.05	2.09%
SPEC-T	47.46	34.13%	31.26	22.30%	67.93	41.83%	47.42	32.47%
ITT1	17.73	12.75%	14.40	10.27%	13.87	8.54%	15.45	10.58%
ITT2	12.74	9.17%	18.20	12.98%	8.99	5.54%	13.64	9.34%
ITT	30.48	21.92%	32.60	23.25%	22.87	14.08%	29.09	19.92%
LEG-P1	6.72	4.83%	7.50	5.35%	5.20	3.20%	6.57	4.50%
LEG-P2	1.23	0.89%	7.77	5.54%	1.23	0.76%	3.58	2.45%
LEG-P	7.95	5.72%	15.27	10.89%	6.44	3.96%	10.15	6.95%
TOTAL	139.05	100.00%	140.21	100.00%	162.40	100.00%	146.05	100.00%

The high frequency of **legal terminology** is common to all the settings, within similar proportions (43.56% in the UN, 40.13% in the WTO and 38.23% in the EU). The variation between the top and the lowest proportional values does not exceed a 6% difference for any LEG-T category, with remarkably homogeneous figures in the case of the predominant LEG-GEN annotations (41.45 per 1000 tokens in the EU, 39.23 in the UN and 38.25 in the WTO). LEG-INT stands out in the UN sub-corpora with 17.78 annotations per 1000 tokens (i.e. 12.68%, compared to 9.81% in

the WTO and 6.62% in the EU), which is not surprising considering the organization's role as the cornerstone of international law. As for national legal terms, both LEG-NAT-SIN and LEG-NAT-GEN are more frequent in the WTO sub-corpora, with a combined density of approximately 11 annotations per 1000 tokens, compared to 4 in the UN sub-corpora and 2.5 in the EU's. This can be linked to the dynamics of monitoring procedures, as will be further examined in Section 3.2.2.



**Figure 3.** Distribution of the annotated features in the EU, UN and WTO sub-corpora, as well as in LETRINT 1+ as a whole (frequencies normalized to 1000 tokens)

It is within **terminology of other specialized domains** that the most marked differences are found between institutional sub-corpora, with disparities reaching twenty percentage points between the highest and the lowest SPEC-T subtotals: 41.83% in the WTO, 34.13% in the EU and 22.30% in the UN. The most important inter-institutional disparity of all annotation results is that of ECO-T frequencies, with a 25% difference between the WTO and the UN sub-corpora sets. Due to the WTO's specialization in trade matters, ECO-T is actually the most frequent category in this setting (as opposed to LEG-GEN in the other two settings), which in turn makes it the only institutional sub-corpora set where the SPEC-T subtotal exceeds the LEG-T subtotal, even if slightly. In contrast, POL-T stands out in the UN sub-corpora, with 16.09% of annotations and the second most salient inter-institutional variation: a 11.5% difference or three times more annotations per 1000 tokens than in the other two settings. These results align with the UN's more pronounced emphasis on human rights, humanitarian work and interna-

tional governance. Otherwise, the UN sub-corpora registered the lowest results for all the other SPEC-T categories.

The EU sub-corpora, however, have the most homogeneous distribution of terminology of non-legal domains: except for ECO-T (with 13.74% of annotations per 1000 tokens), the values for SPEC-T categories range roughly between 4% and 6%. In comparison with the other sub-corpora, the EU's frequencies for SCI-T and TEC-T, albeit modest, are significantly higher, with a combined subtotal of almost 10% of annotations per 1000 tokens, as opposed to approximately 4% in WTO texts and 2.5% in the UN's. These findings are consistent with the diversity of topics addressed within the EU's policy areas.

The use of **institutional titles** is more recurrent in the UN and EU sub-corpora (with 32.60 and 30.48 annotations per 1000 tokens, respectively) than in the WTO's (with 22.87 annotations per 1000 tokens). While the EU texts registered the highest density of ITT1 annotations (17.73 per 1000 tokens), the UN sub-corpora include a more prominent proportion of names of bodies and positions (18.20 ITT2 annotations per 1000 tokens), which is in line with the myriad of agencies, structures and mandates created over time within the UN, and the resulting cross-references in its texts. Conversely, the WTO's shorter history and more reduced structures may explain why the density of ITT2 annotations in its sub-corpora is the lowest of the three settings (8.99 per 1000 tokens, i.e. half the UN's frequency for this category).

Regarding **legal phraseology**, LEG-P1 results are very similar in the three settings, within a range of 5.20 to 7.50 annotations per 1000 tokens. As with LEG-GEN, both this finding and the corresponding recurrence ratios point to the regularity of these features in building legal discourses in all the settings. The reverse applies to genre-bound LEG-P2 expressions, which are much more recurrent in the UN texts (7.77 annotations per 1000 tokens, as opposed to only 1.23 in the EU and WTO sub-corpora), as a result of the marked formulaicity of UN resolutions and UN treaty body reports. This will be further discussed below in connection with the internal distribution of features within each institutional sub-corpora set.

### 3.2.1 *Distribution in the EU sub-corpora*

If we observe the EU sub-corpora more closely (see Table 6 and Figure 4), the most striking difference between the selected genres is the much higher frequency of LEG-T in CJEU judgments (EU-3), roughly doubling the density found in the other two EU sub-corpora (83.87 annotations per 1000 tokens, as opposed to 39.16 in EU-1 and 36.42 in EU-2). The EU-3's subtotal is the highest for LEG-T in any LETRINT 1+ sub-corpus. This prominence is particularly pronounced in the case of LEG-GEN (72.17 annotations per 1000 tokens in EU-3 *versus* 28.43 in EU-1 and



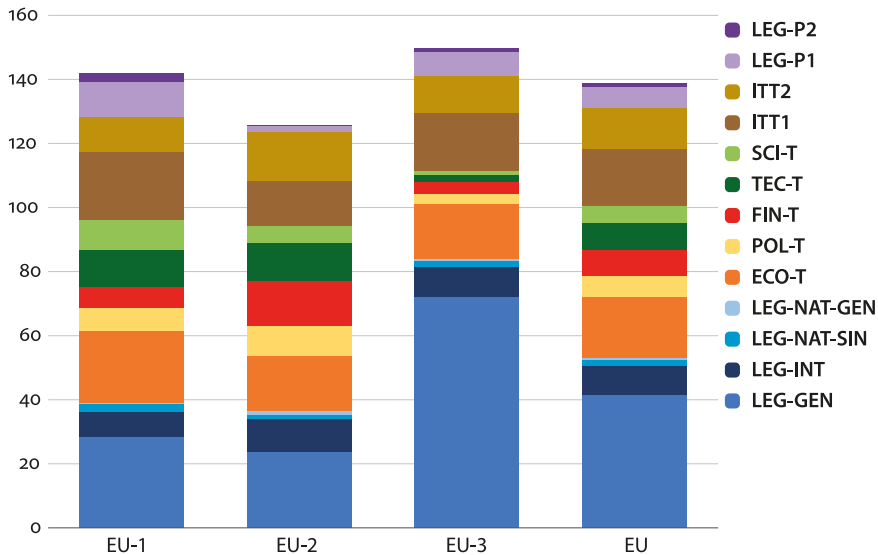
23.71 in EU-2), which is not surprising considering the recourse to common legal terms in presenting legal procedures and reasoning in judgments. These results greatly contribute to the higher total of annotations in EU-3 as a whole. In contrast, the limited frequency of LEG-NAT terminology is very similar across the EU sub-corpora, well below the results in the other two institutional settings.

**Table 6.** Distribution of annotated features in the EU sub-corpora separately and as a whole (frequencies normalized to 1000 tokens and their percentages)

	EU-1	EU-1%	EU-2	EU-2%	EU-3	EU-3%	EU	EU%
LEG-GEN	28.43	20.01%	23.71	18.89%	72.17	48.26%	41.45	29.81%
LEG-INT	7.92	5.58%	10.45	8.33%	9.24	6.18%	9.21	6.62%
LEG-NAT-SIN	2.23	1.57%	1.02	0.81%	2.03	1.36%	1.76	1.27%
LEG-NAT-GEN	0.59	0.42%	1.23	0.98%	0.42	0.28%	0.75	0.54%
<b>LEG-T</b>	<b>39.16</b>	<b>27.57%</b>	<b>36.42</b>	<b>29.01%</b>	<b>83.87</b>	<b>56.08%</b>	<b>53.16</b>	<b>38.23%</b>
ECO-T	22.60	15.91%	17.41	13.87%	17.31	11.57%	19.10	13.74%
POL-T	6.88	4.84%	9.16	7.30%	3.09	2.06%	6.37	4.58%
FIN-T	6.56	4.62%	14.14	11.27%	3.95	2.64%	8.22	5.91%
TEC-T	11.75	8.27%	11.86	9.45%	1.97	1.31%	8.52	6.13%
SCI-T	9.17	6.45%	5.19	4.14%	1.37	0.91%	5.24	3.77%
<b>SPEC-T</b>	<b>56.95</b>	<b>40.09%</b>	<b>57.77</b>	<b>46.02%</b>	<b>27.68</b>	<b>18.51%</b>	<b>47.46</b>	<b>34.13%</b>
ITT <sub>1</sub>	21.15	14.89%	14.14	11.27%	17.91	11.97%	17.73	12.75%
ITT <sub>2</sub>	11.13	7.83%	15.38	12.25%	11.73	7.84%	12.74	9.17%
<b>ITT</b>	<b>32.27</b>	<b>22.72%</b>	<b>29.52</b>	<b>23.51%</b>	<b>29.64</b>	<b>19.82%</b>	<b>30.48</b>	<b>21.92%</b>
LEG-P <sub>1</sub>	10.94	7.70%	1.81	1.44%	7.41	4.95%	6.72	4.83%
LEG-P <sub>2</sub>	2.74	1.93%	0.01	0.01%	0.96	0.64%	1.23	0.89%
<b>LEG-P</b>	<b>13.68</b>	<b>9.62%</b>	<b>1.82</b>	<b>1.45%</b>	<b>8.36</b>	<b>5.59%</b>	<b>7.95</b>	<b>5.72%</b>
<b>TOTAL</b>	<b>142.06</b>	<b>100.00%</b>	<b>125.53</b>	<b>100.00%</b>	<b>149.55</b>	<b>100.00%</b>	<b>139.05</b>	<b>100.00%</b>

SPEC-T is much more prominent in EU-1 and EU-2, where the recurrence of these categories (with approximately 57 annotations per 1000 tokens) more than doubles that registered in EU-3 (27.68), exceeding the LEG-T subtotals of the first two sub-corpora. ECO-T is the most frequent SPEC-T category in all EU genres selected, with similar values in the three sub-corpora (between 17.31 and 22.60 annotations per 1000 tokens). The results for the remaining SPEC-T categories reveal interesting variations: while they fluctuate between 1.37 (0.91%) and 3.95 (2.64%) annotations per 1000 tokens in EU-3, they range between 6.56 (4.62%)

and 11.75 (8.27%) in EU-1, and between 5.19 (4.14%) and 14.14 (11.27%) in EU-2, including the highest frequencies of SCI-T and TEC-T of all the LETRINT 1+ sub-corpora. The combined proportion of these two categories in EU-1 (14.72% of annotations) and EU-2 (13.59%) attest to the regular attention devoted to technical and scientific matters in EU regulations and their subsequent follow-up in implementation reports. The frequency of FIN-T stands out in the latter with 11.27% of annotations.



**Figure 4.** Distribution of annotated features in the EU sub-corpora separately and as a whole (frequencies normalized to 1000 tokens)

The use of ITT is quite similar in the three sub-corpora, with between 29.52 and 32.27 annotations per 1000 tokens. EU-1 and EU-3 understandably include more references to legislation and documents (ITT1), while the names of bodies and positions (ITT2) are more frequent in EU-2 (the only EU sub-corpus where ITT2 surpasses ITT1). Finally, the use of LEG-P is most recurrent in EU-1, with 13.68 annotations per 1000 tokens (including 10.94 of LEG-P1 and 2.74 of LEG-P2), compared to 8.36 in EU-3 and only 1.82 in EU-2.

3.2.2 *Distribution in the UN sub-corpora*

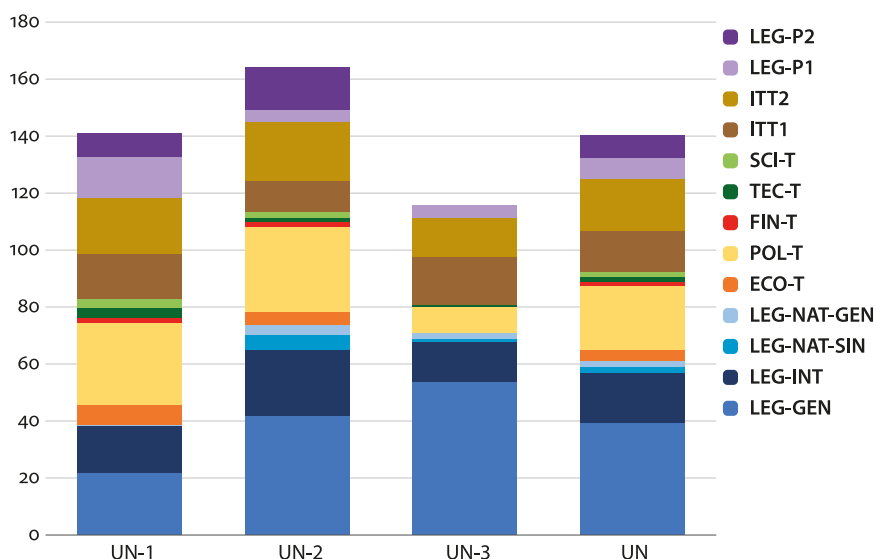
Two of the above-mentioned features of the UN sub-corpora (see Table 7 and Figure 5) are found in the three genres: the significant proportion of LEG-INT within LEG-T (between 11.59% and 13.97% of annotations), and the distinctive prominence of POL-T within SPEC-T (between 20% and 8%, the highest pro-

portions among LETRINT 1+ sub-corpora). This second feature is particularly marked in the case of UN-1, where it is the most frequent category, as opposed to UN-2 and UN-3, where LEG-GEN is more recurrent. In turn, the results for this category reflect the legal nature of each genre. UN resolutions (UN-1), as soft law instruments, contain the lowest proportion of LEG-GEN of all the LETRINT 1+ sub-corpora (21.97 annotations per 1000 tokens). In contrast, UN human rights treaty bodies' concluding observations (UN-2) score higher (42.04 annotations per 1000 tokens) than any other monitoring genre in LETRINT 1+ (see Section 3.3), which can be explained by the "quasi-judicial" function fulfilled by these bodies within the UN system (Prieto Ramos 2014a: 316). Finally, ICJ Judges' opinions (UN-3) yield the highest LEG-GEN frequency of all UN sub-corpora (53.68 annotations per 1000 tokens), in line with the prominent legal argumentative nature of this genre, and similarly to the other adjudication genres of the corpus (see Section 3.3).

**Table 7.** Distribution of annotated features in the UN sub-corpora separately and as a whole (frequencies normalized to 1000 tokens and their percentages)

	UN-1	UN-1%	UN-2	UN-2%	UN-3	UN-3%	UN	UN%
LEG-GEN	21.97	15.59%	42.04	25.61%	53.68	46.44%	39.23	27.98%
LEG-INT	16.34	11.59%	22.93	13.97%	14.08	12.18%	17.78	12.68%
LEG-NAT-SIN	0.14	0.10%	5.25	3.20%	1.30	1.12%	2.23	1.59%
LEG-NAT-GEN	0.06	0.04%	3.68	2.24%	1.78	1.54%	1.84	1.31%
<b>LEG-T</b>	<b>38.51</b>	<b>27.32%</b>	<b>73.91</b>	<b>45.03%</b>	<b>70.83</b>	<b>61.28%</b>	<b>61.08</b>	<b>43.56%</b>
ECO-T	7.34	5.21%	4.24	2.58%	0.13	0.12%	3.90	2.78%
POL-T	28.45	20.18%	30.07	18.32%	9.18	7.94%	22.56	16.09%
FIN-T	1.98	1.40%	1.51	0.92%	0.04	0.04%	1.18	0.84%
TEC-T	3.63	2.58%	1.68	1.02%	0.48	0.41%	1.93	1.38%
SCI-T	2.88	2.04%	1.97	1.20%	0.23	0.20%	1.69	1.21%
<b>SPEC-T</b>	<b>44.28</b>	<b>31.41%</b>	<b>39.47</b>	<b>24.05%</b>	<b>10.07</b>	<b>8.71%</b>	<b>31.26</b>	<b>22.30%</b>
ITT <sub>1</sub>	15.71	11.14%	10.76	6.55%	16.73	14.47%	14.40	10.27%
ITT <sub>2</sub>	19.95	14.15%	21.01	12.80%	13.64	11.80%	18.20	12.98%
<b>ITT</b>	<b>35.66</b>	<b>25.30%</b>	<b>31.76</b>	<b>19.35%</b>	<b>30.37</b>	<b>26.27%</b>	<b>32.60</b>	<b>23.25%</b>
LEG-P <sub>1</sub>	14.14	10.03%	4.03	2.45%	4.32	3.74%	7.50	5.35%
LEG-P <sub>2</sub>	8.36	5.93%	14.96	9.11%	0.00	0.00%	7.77	5.54%
<b>LEG-P</b>	<b>22.50</b>	<b>15.97%</b>	<b>18.99</b>	<b>11.57%</b>	<b>4.32</b>	<b>3.74%</b>	<b>15.27</b>	<b>10.89%</b>
<b>TOTAL</b>	<b>140.95</b>	<b>100.00%</b>	<b>164.13</b>	<b>100.00%</b>	<b>115.59</b>	<b>100.00%</b>	<b>140.21</b>	<b>100.00%</b>

A third feature previously highlighted with regard to the UN texts, namely, the high proportion of ITT2 annotations, also applies to all the sub-corpora, which registered the top scores for this category within their respective legal functions (see Section 3.3). However, UN-3 has the lowest ITT2 frequency of the three UN sub-corpora, and it is the only one where the number of ITT1 annotations per 1000 tokens exceeds that of ITT2 in this setting, which is also aligned to other adjudication genres.



**Figure 5.** Distribution of annotated features in the UN sub-corpora separately and as a whole (frequencies normalized to 1000 tokens)

Two further UN-3 singularities emerge from the results. First, the extremely low recurrence of SPEC-T, the lowest among LETRINT 1+ sub-corpora, with all values (other than POL-T) under 0.5 annotations per 1000 tokens and a SPEC-T subtotal of 10.07, compared to fluctuations between 1.98 and 7.34 in UN-1, and between 1.51 and 4.24 in UN-2, and SPEC-T subtotals of 44.28 and 39.47, respectively. This shows how, in analyzing issues of international law, ICJ Judges' opinions refer to far fewer technical details and factual elements than in other genres. Second, the absence of LEG-P2 in UN-3 is a sign of the stylistic heterogeneity found in individual opinions, where, as opposed to other judicial genres,<sup>7</sup> judges

7. ICJ judgments are co-drafted to varying degrees rather than translated from an original language that can be identified as such; therefore, they are not part of the LETRINT 1 or LETRINT 1+ corpora.

do not seem to follow pre-established formulations in structuring their texts. This is in stark contrast with the marked LEG-P2 formulaicity of UN-1 and especially UN-2, with the highest LEG-P2 scores among LETRINT 1+ sub-corpora (8.36 and 14.96 annotations per 1000 tokens, respectively). These data also point to the repetition of certain standardized segments in the translation of UN resolutions and treaty body reports.

### 3.2.3 *Distribution in the WTO sub-corpora*

The most outstanding features of the WTO sub-corpora, as mentioned above, i.e. comparatively high ECO-T and LEG-NAT frequencies and relatively low ITT scores, are accentuated in WTO-2 (see Table 8 and Figure 6). ECO-T results for this sub-corpus are the most prominent for a single SPEC-T category in the entire corpus, at 67.61 annotations per 1000 tokens (almost 40% of the total), and the global SPEC-T subtotal (104.79) outnumbers the top LEG-T subtotal in any single sub-corpus (that of EU-3 at 83.87). LEG-T categories actually account for half the proportion of SPEC-T in WTO-2 (51.70 annotations per 1000 tokens). The frequencies of LEG-GEN and LEG-INT are the lowest of the three WTO sub-corpora. The first category stands out in WTO-3 (52.43 annotations per 1000 tokens), while LEG-INT is slightly more frequent in WTO-1 (23.10) than in WTO-3 (20.49). However, LEG-NAT-SIN and LEG-NAT-GEN values reach their highest levels of the entire corpus in WTO-2 (11.50 and 6.10 annotations per 1000 tokens, respectively). Notably, this is the only LETRINT 1+ sub-corpus where the annotations of national system-bound terms outnumber those of international law terms.

The focus of trade policy reviews on measures implemented at the national level explains the above findings, as well as the higher values obtained for all SPEC-T categories, and not only ECO-T, when compared to the other WTO sub-corpora. The FIN-T score (16.09 annotations per 1000 tokens) is also the highest of this category in LETRINT 1+. In fact, WTO-2 reports cover a wide range of themes for each WTO Member under examination, including its macroeconomic situation, its trade and investment framework, and multiple other aspects of its trade-related policies (duties applied on products, customs procedures, technical regulations, sanitary and phytosanitary requirements, developments by sector, etc.).

By the same token, it is not surprising that the in-depth review of national measures and bodies in this genre turns the focus away from ITT, remarkably less frequent than in the other WTO sub-corpora (11.62 annotations per 1000 tokens, as opposed to 42.87 in WTO-1 and 26.97 in WTO-3). The ITT<sub>2</sub> scores of the three WTO sub-corpora are the lowest of all the LETRINT 1+ sub-corpora, while, interestingly, WTO-1 has the highest density of references to institutional

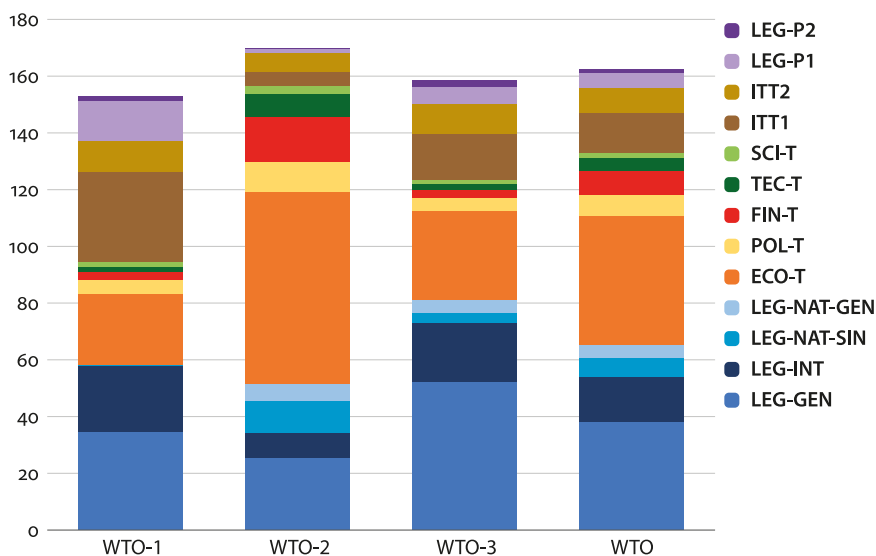
**Table 8.** Distribution of annotated features in the WTO sub-corpora separately and as a whole (frequencies normalized to 1000 tokens and their percentages)

	WTO-1	WTO-1%	WTO-2	WTO-2%	WTO-3	WTO-3%	WTO	WTO%
LEG-GEN	34.81	22.76%	25.30	14.90%	52.43	33.09%	38.25	23.55%
LEG-INT	23.10	15.11%	8.81	5.19%	20.49	12.93%	15.93	9.81%
LEG-NAT-SIN	0.50	0.33%	11.50	6.77%	3.81	2.40%	6.57	4.05%
LEG-NAT-GEN	0.00	0.00%	6.10	3.59%	4.31	2.72%	4.42	2.72%
LEG-T	58.40	38.19%	51.70	30.45%	81.04	51.15%	65.17	40.13%
ECO-T	24.72	16.16%	67.61	39.82%	31.37	19.80%	45.74	28.16%
POL-T	5.20	3.40%	10.55	6.22%	4.89	3.08%	7.34	4.52%
FIN-T	2.68	1.75%	16.09	9.48%	2.49	1.57%	8.29	5.10%
TEC-T	1.81	1.18%	7.77	4.57%	2.35	1.49%	4.57	2.81%
SCI-T	1.71	1.12%	2.78	1.64%	1.33	0.84%	2.00	1.23%
SPEC-T	36.11	23.62%	104.79	61.73%	42.43	26.79%	67.93	41.83%
ITT1	31.88	20.85%	5.04	2.97%	16.28	10.28%	13.87	8.54%
ITT2	10.99	7.19%	6.58	3.87%	10.69	6.75%	8.99	5.54%
ITT	42.87	28.03%	11.62	6.84%	26.97	17.03%	22.87	14.08%
LEG-P1	13.85	9.06%	1.61	0.95%	5.71	3.60%	5.20	3.20%
LEG-P2	1.68	1.10%	0.04	0.03%	2.27	1.43%	1.23	0.76%
LEG-P	15.54	10.16%	1.66	0.98%	7.97	5.03%	6.44	3.96%
TOTAL	152.92	100.00%	169.77	100.00%	158.42	100.00%	162.40	100.00%

legal instruments and documents (ITT1) within LETRINT 1+. Unsurprisingly, this sub-corpus also contains the largest number of LEG-P1 annotations per 1000 tokens in comparison with the other WTO sub-corpora (13.85 *versus* 5.71 in WTO-3 and 1.61 in WTO-2). The frequency of LEG-P2 is limited in all of them, particularly in WTO-2.

3.3 Comparative analysis per primary legal function and genre

The cross-cutting comparison of results per primary legal function reveals patterns that converge based on common functions, regardless of institutional thematic specificities. It also helps to discern genre-specific features that depart from shared functional trends (see Tables 9 and 10, and Figures 7 and 8). For this contrastive analysis, we will refer to three sub-corpora sets: law-making (L1+(1)), including EU-1, UN-1 and WTO-1; compliance monitoring (L1+(2)),



**Figure 6.** Distribution of annotated features in the WTO sub-corpora separately and as a whole (frequencies normalized to 1000 tokens)

which groups together EU-2, UN-2 and WTO-2; and adjudication (L1+(3)), made up of EU-3, UN-3 and WTO-3.

The most salient feature for any functional sub-corpora set is the predominance of LEG-GEN in L1+(3), which is found in all institutional settings as a common discourse feature of legal argumentation in judgments and judges' opinions. So is the concomitant prevalence of LEG-T over SPEC-T in L1+(3), with the highest sub-total differences per organization and primary function (56.19 more LEG-T than SPEC-T annotations per 1000 tokens in EU-3, +60.76 in UN-3 and +38.61 in WTO-3), even if LEG-T subtotals are similar in UN-3 and UN-2, and SPEC-T subtotals also converge in the case of WTO-3 and WTO-2. The highest densities of LEG-T as a whole, and LEG-GEN in particular, in the entire corpus are found in EU-3, while, as highlighted in Section 3.2.2, UN-3 has the lowest SPEC-T scores of the corpus.

LEG-GEN is also the most frequent single category within all the primary functions and all the genres, with the only exception of WTO-2 (due to the heightened prominence of ECO-T) and UN-1 (with more POL-T than LEG-GEN, although the difference is not so marked). Once again, these exceptions can be linked to institutional singularities, but do not undermine the overall reliance on LEG-GEN as a core legal discourse feature in all the primary functions examined. However, the SPEC-T subtotal exceeds LEG-T slightly in L1+(1) and more clearly in L1+(2). There are two exceptions to this trend: WTO-1, with

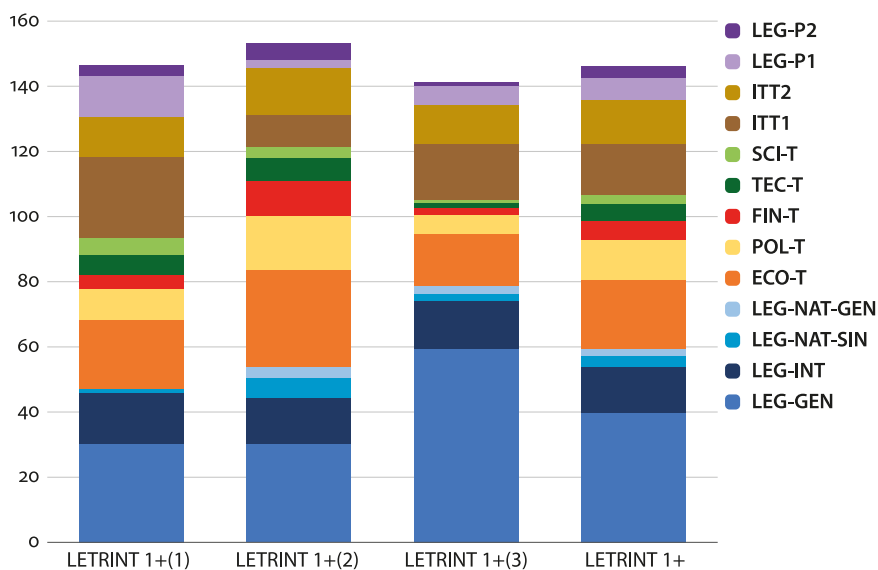
the highest density of organization-specific LEG-INT within LETRINT 1+ (23.10 annotations per 1000 tokens) and the highest LEG-GEN value and LEG-T subtotal of all L1+(1) sub-corpora; and UN-2, which registered more LEG-T than SPEC-T annotations, and higher densities for all LEG-T categories than UN-1 (see Table 10 and Figure 8).

**Table 9.** Distribution of annotated features in the three sub-corpora sets per primary legal function, as well as in LETRINT 1+ as a whole (frequencies normalized to 1000 tokens and their percentages)

	L1+(1)	L1+(1)%	L1+(2)	L1+(2)%	L1+(3)	L1+(3)%	L1+	L1+%
LEG-GEN	28.40	19.55%	30.35	19.81%	59.43	42.09%	39.75	27.22%
LEG-INT	15.79	10.86%	14.06	9.18%	14.60	10.34%	14.18	9.71%
LEG-NAT-SIN	0.96	0.66%	5.93	3.87%	2.38	1.69%	3.29	2.25%
LEG-NAT-GEN	0.22	0.15%	3.67	2.40%	2.17	1.54%	2.17	1.49%
<b>LEG-T</b>	<b>45.36</b>	<b>31.22%</b>	<b>54.01</b>	<b>35.26%</b>	<b>78.58</b>	<b>55.66%</b>	<b>59.39</b>	<b>40.66%</b>
ECO-T	18.22	12.54%	29.77	19.43%	16.28	11.53%	21.16	14.49%
POL-T	13.51	9.30%	16.59	10.83%	5.72	4.05%	12.46	8.53%
FIN-T	3.74	2.57%	10.59	6.92%	2.16	1.53%	5.71	3.91%
TEC-T	5.73	3.94%	7.10	4.64%	1.60	1.13%	5.04	3.45%
SCI-T	4.59	3.16%	3.31	2.16%	0.98	0.69%	3.05	2.09%
<b>SPEC-T</b>	<b>45.78</b>	<b>31.51%</b>	<b>67.37</b>	<b>43.99%</b>	<b>26.73</b>	<b>18.94%</b>	<b>47.42</b>	<b>32.47%</b>
ITT1	22.91	15.77%	9.98	6.52%	16.97	12.02%	15.45	10.58%
ITT2	14.02	9.65%	14.32	9.35%	12.02	8.52%	13.64	9.34%
<b>ITT</b>	<b>36.93</b>	<b>25.42%</b>	<b>24.30</b>	<b>15.86%</b>	<b>28.99</b>	<b>20.54%</b>	<b>29.09</b>	<b>19.92%</b>
LEG-P1	12.98	8.93%	2.48	1.62%	5.81	4.12%	6.57	4.50%
LEG-P2	4.26	2.93%	5.00	3.27%	1.07	0.76%	3.58	2.45%
<b>LEG-P</b>	<b>17.24</b>	<b>11.86%</b>	<b>7.49</b>	<b>4.89%</b>	<b>6.88</b>	<b>4.88%</b>	<b>10.15</b>	<b>6.95%</b>
<b>TOTAL</b>	<b>145.31</b>	<b>100.00%</b>	<b>153.16</b>	<b>100.00%</b>	<b>141.18</b>	<b>100.00%</b>	<b>146.05</b>	<b>100.00%</b>

Apart from the smaller gap between LEG-T and SPEC-T in all L1+(1) sub-corpora subtotals, another clear commonality of L1+(1) sub-corpora is the use of LEG-P1, which systematically occurs more frequently in law-making than in the other primary functions (12.98 annotations per 1000 tokens in L1+(1), compared to 5.81 in L1+(3) and 2.48 in L1+(2)). References to legal instruments and documents (ITT1) are also more frequent in law-making (22.91 annotations per 1000 tokens), although closely followed by adjudication genres (and only marginally surpassed in the case of UN-3).





**Figure 7.** Distribution of annotated features in the three sub-corpora sets per primary legal function, as well as in LETRINT 1+ as a whole (frequencies normalized to 1000 tokens)

In contrast, LEG-INT and ITT2 score similar values in all the functional sub-corpora sets (between 15.79 and 14.06 LEG-INT annotations per 1000 tokens, and between 14.32 and 12.02 in the case of ITT2, including the top scores in the UN sub-corpora as mentioned above). Interestingly, these results align with the high unit recurrence ratios for these categories (see Table 4) as core components of international legal discourses across genres.

If we look into the features of L1+(2), the salience of SPEC-T reflects inter-institutional variations of thematic specialization, but within a shared functional pattern of higher frequency of annotations compared to L1+(1) and especially L1+(3). Except for SCI-T, which registered its top score in L1+(1) due to the result for EU legislation, all the other SPEC-T categories are more frequent in L1+(2), including the highest values for ECO-T (67.61 annotations per 1000 tokens in WTO-2), POL-T (30.07 in UN-2), FIN-T (16.09 in WTO-2) and TEC-T (11.86 in EU-2).

The L1+(2) sub-corpora set also yields the highest frequency of LEG-NAT terminology, including 5.93 annotations of LEG-NAT-SIN and 3.67 of LEG-NAT-GEN per 1000 tokens, compared to 2.38 and 2.17 in L1+(3), and 0.96 and 0.22 in L1+(1), respectively. This trend is very marked in the case of the two inter-governmental organizations covered by the project, with marginal occurrences of national system-specific concepts and body names in law-making, and much

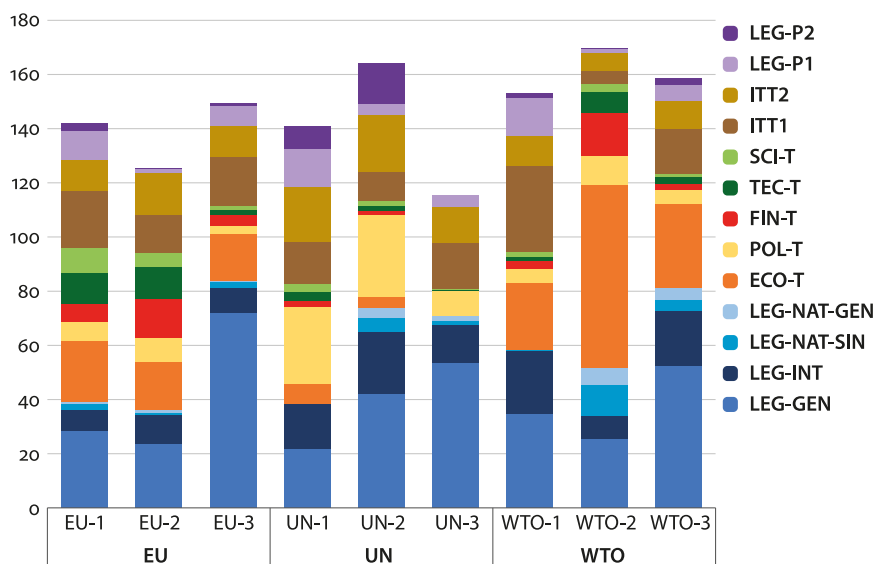
**Table 10.** Distribution of annotated features per primary legal function and institutional setting (frequencies normalized to 1000 tokens)

	LAW-MAKING (L1+(1))			MONITORING (L1+(2))			ADJUDICATION (L1+(3))		
	EU-1	UN-1	WTO-1	EU-2	UN-2	WTO-2	EU-3	UN-3	WTO-3
LEG-GEN	28.43	21.97	34.81	23.71	42.04	25.30	72.17	53.68	52.43
LEG-INT	7.92	16.34	23.10	10.45	22.93	8.81	9.24	14.08	20.49
LEG-NAT-GEN	0.59	0.06	0.00	1.23	3.68	6.10	0.42	1.78	4.31
LEG-NAT-SIN	2.23	0.14	0.50	1.02	5.25	11.50	2.03	1.30	3.81
<b>LEG-T</b>	<b>39.16</b>	<b>38.51</b>	<b>58.40</b>	<b>36.42</b>	<b>73.91</b>	<b>51.70</b>	<b>83.87</b>	<b>70.83</b>	<b>81.04</b>
ECO-T	22.60	7.34	24.72	17.41	4.24	67.61	17.31	0.13	31.37
FIN-T	6.56	1.98	2.68	14.14	1.51	16.09	3.95	0.04	2.49
POL-T	6.88	28.45	5.20	9.16	30.07	10.55	3.09	9.18	4.89
SCI-T	9.17	2.88	1.71	5.19	1.97	2.78	1.37	0.23	1.33
TEC-T	11.75	3.63	1.81	11.86	1.68	7.77	1.97	0.48	2.35
<b>SPEC-T</b>	<b>56.95</b>	<b>44.28</b>	<b>36.11</b>	<b>57.77</b>	<b>39.47</b>	<b>104.79</b>	<b>27.68</b>	<b>10.07</b>	<b>42.43</b>
LEG-P1	10.94	14.14	13.85	1.81	4.03	1.61	7.41	4.32	5.71
LEG-P2	2.74	8.36	1.68	0.01	14.96	0.04	0.96	0.00	2.27
<b>LEG-P</b>	<b>13.67</b>	<b>22.50</b>	<b>15.54</b>	<b>1.82</b>	<b>18.99</b>	<b>1.66</b>	<b>8.36</b>	<b>4.32</b>	<b>7.97</b>
ITT1	21.15	15.71	31.88	14.14	10.76	5.04	17.91	16.73	16.28
ITT2	11.13	19.95	10.99	15.38	21.01	6.58	11.73	13.64	10.69
<b>ITT</b>	<b>32.27</b>	<b>35.66</b>	<b>42.87</b>	<b>29.52</b>	<b>31.76</b>	<b>11.62</b>	<b>29.64</b>	<b>30.37</b>	<b>26.97</b>
<b>TOTAL</b>	<b>142.06</b>	<b>140.95</b>	<b>152.92</b>	<b>125.53</b>	<b>164.13</b>	<b>169.77</b>	<b>149.55</b>	<b>115.59</b>	<b>158.42</b>

higher scores in monitoring and adjudication genres, where such references are essential to describe implementation measures at the national level. As mentioned in Section 3.2.3, they are particularly frequent in WTO trade policy review reports (WTO-2).

In contrast, while the low LEG-NAT values are very similar in the three EU sub-corpora, EU-1 scores stand out within L1+(1), with a combined total of 2.82 annotations of LEG-NAT-SIN and LEG-NAT-GEN per 1000 tokens (compared to 0.50 in WTO-1 and 0.20 in UN-1). This can be explained by the closer connections between national legal systems and the supranational legal order in EU law, which means that explicit references to EU Member States' entity names are found in EU-1, often in national authority lists in annexes to regulations. However, these references are not as recurrent as the wide range of national bodies and legislation titles mentioned in WTO and UN texts in monitoring procedures.

Finally, the comparison of sub-corpora per primary legal function does not reveal any clear pattern regarding LEG-P2 fluctuations from this perspective,



**Figure 8.** Distribution of annotated features in the nine sub-corpora (frequencies normalized to 1000 tokens)

which serves to corroborate that this feature can be primarily associated to genre conventions. As mentioned in Section 3.2.2, the highest LEG-P2 frequencies are found in two different sub-corpora of the same organization, UN-1 and especially UN-2. The analysis of internal variations within and between functional sub-corpora provides further insights into the singularity of the latter genre. In fact, UN-2 contains the most significant concentration of divergent patterns compared to the other  $L_1+(2)$  sub-corpora, whereas it partially converges with  $L_1+(3)$  trends. A closer comparison between genres reveals that UN-2 scores are extremely similar to those of WTO-3, except for the distinct thematic focus of each institution (POL-T in the UN and ECO-T in the WTO) and the high LEG-P2 formulaicity of UN treaty bodies' concluding observations. All in all, our findings support the characterization of these “quasi-judicial” reports as highly hybrid from a legal functional perspective.

In order to better visualize **genre** similarities and differences in light of institutional and legal functional discursive trends, Table 11 provides a final overview of the most frequently annotated categories per sub-corpus, as well as their corresponding annotation subtotals and totals. The richest genres from a terminological angle are WTO-2 (i.e. the most technical of all genres as per SPEC-T results), with 169.77 annotations per 1000 tokens; followed by UN-2 (i.e. the  $L_1+(2)$  genre where the legal discourse features are most prevalent), with 164.13 annotations per 1000 tokens; and WTO-3 (i.e. the adjudication genre with the largest amount

**Table 11.** Top frequencies of annotated categories and annotation subtotals and totals for each sub-corpus

	EU	UN	WTO
Law-making	<i>Most frequent categories</i>	<i>Most frequent categories</i>	<i>Most frequent categories</i>
	LEG-GEN: 28.43	POL-T: 28.45	LEG-GEN: 34.81
	ECO-T: 22.60	LEG-GEN: 21.97	ITT1: 31.88
	ITT1: 21.15	ITT2: 19.95	ECO-T: 24.72
	TEC-T: 11.75	LEG-INT: 16.34	LEG-INT: 23.10
	ITT2: 11.13	ITT1: 15.71	LEG-P1: 13.85
	<i>Subtotals &amp; total</i>	<i>Subtotals &amp; total</i>	<i>Subtotals &amp; total</i>
	SPEC-T: 56.95	SPEC-T: 44.28	LEG-T: 58.40
	LEG-T: 39.16	LEG-T: 38.51	ITT: 42.87
	ITT: 32.27	ITT: 35.66	SPEC-T: 36.11
	LEG-P: 13.67	LEG-P: 22.50	LEG-P: 15.54
	TOTAL: 142.06	TOTAL: 140.95	TOTAL: 152.92
Monitoring	<i>Most frequent categories</i>	<i>Most frequent categories</i>	<i>Most frequent categories</i>
	LEG-GEN: 23.71	LEG-GEN: 42.04	ECO-T: 67.61
	ECO-T: 17.41	POL-T: 30.07	LEG-GEN: 25.3
	ITT2: 15.38	LEG-INT: 22.93	FIN-T: 16.09
	FIN-T: 14.14	ITT2: 21.01	LEG-NAT-SIN: 11.5
	ITT1: 14.14	LEG-P2: 14.96	POL-T: 10.55
	<i>Subtotals &amp; total</i>	<i>Subtotals &amp; total</i>	<i>Subtotals &amp; total</i>
	SPEC-T: 57.77	LEG-T: 73.91	SPEC-T: 104.79
	LEG-T: 36.42	SPEC-T: 39.47	LEG-T: 51.70
	ITT: 29.52	ITT: 31.76	ITT: 11.62
	LEG-P: 1.82	LEG-P: 18.99	LEG-P: 1.66
	TOTAL: 125.53	TOTAL: 164.13	TOTAL: 169.77
Adjudication	<i>Most frequent categories</i>	<i>Most frequent categories</i>	<i>Most frequent categories</i>
	LEG-GEN: 72.17	LEG-GEN: 53.68	LEG-GEN: 52.43
	ITT1: 17.91	ITT1: 16.73	ECO-T: 31.37
	ECO-T: 17.31	LEG-INT: 14.08	LEG-INT: 20.49
	ITT2: 11.73	ITT2: 13.64	ITT1: 16.28
	LEG-INT: 9.24	POL-T: 9.18	ITT2: 10.69
	<i>Subtotals &amp; total</i>	<i>Subtotals &amp; total</i>	<i>Subtotals &amp; total</i>
	LEG-T: 83.87	LEG-T: 70.83	LEG-T: 81.04
	ITT: 29.64	ITT: 30.37	SPEC-T: 42.43
	SPEC-T: 27.68	SPEC-T: 10.07	ITT: 26.97
	LEG-P: 8.36	LEG-P: 4.32	LEG-P: 7.97
	TOTAL: 149.55	TOTAL: 115.59	TOTAL: 158.42

of SPEC-T annotations but also a high density of legal discourse features), with 158.42. The other genres with the largest concentrations of legal discourse features per primary function are WTO-1 within L1+(1) and EU-3 within L1+(3). They come next in the ranking of total annotations per 1000 tokens, with 152.92 and 149.55, respectively. At the other extreme, UN-3 has the lowest density of annotations (115.59 per 1000 tokens), essentially as a result of the low SPEC-T scores.

#### 4. Conclusions

LETRINT's systematic mapping of terminological and phraseological features has yielded valuable evidence of the discursive and thematic hybridity of key genres of EU and international law. The density and recurrence scores of the annotated categories across sub-corpora suggest that several **core features** are used to articulate all the selected legal genres and to situate their functions and subject matters in the institutional frameworks at hand: common legal terminology (LEG-GEN, the most frequent feature), international legal terms (LEG-INT), associated titles of international legal instruments (ITT1) and institutional structures and positions (ITT2), and simple legal phrasemes (LEG-P1). The latter are the least frequent among these common discourse features, but also, unsurprisingly, the category with the highest ratio of recurrence per annotated unit, together with LEG-GEN.

The multiple comparisons between settings and primary legal functions have also revealed **variations** associated with thematic focus or other institutional specificities, as well as function-bound patterns and genre singularities. Overall, the most remarkable differences are found in the frequencies of terminology of non-legal domains, in line with the missions and main policy areas of each institution: predominance of trade and economic matters (ECO-T) at the WTO; more emphasis on human rights and global governance issues (POL-T) at the UN; and a more diverse distribution of thematic areas in EU texts, including the highest density of scientific and technical terminology of the entire LETRINT 1+ corpus. Other salient **institution-specific results** include the higher density of body names and position titles (ITT1) in UN texts, in line with the organization's multiple structures and longer history; and the comparatively marked prominence of national legal terminology (LEG-NAT-SIN and LEG-NAT-GEN) in the WTO due to the more recurrent references to national legal frameworks and bodies in monitoring and dispute settlement procedures.

As for the features that converge based on **primary legal function** in all the settings, our findings point to the following commonalities: the more frequent use of LEG-P1 and ITT1 in law-making; the salience of non-legal specialized terminology and the highest scores of national legal terminology in monitoring proce-

dures; and the top density values of legal terminology in adjudication genres due to the highest frequency of LEG-GEN.

Overall, the richest **genres** from a terminological perspective are the monitoring genres of the two multilateral organizations examined, WTO trade policy review reports and UN human rights treaty body reports, followed by WTO dispute settlement reports, WTO binding decisions, CJEU judgments and EU regulations. The UN sub-corpora present the most marked disparities within a single setting, and also the most singular features per primary legal function compared to the other institutions. These are explained by the distinct nature of each UN genre. UN resolutions (UN-1), the only soft law genre in the corpus, contain the lowest density of LEG-GEN of all the LETRINT 1+ sub-corpora, while UN treaty body reports (UN-2) contain the highest of all the monitoring genres, and partially converge with the terminological features of WTO dispute settlement reports. Both UN-1 and especially UN-2 have, by far, the most prominent density of genre-specific formulaic expressions (LEG-P<sub>2</sub>) of all the LETRINT 1+ sub-corpora. This contrasts with the lack of formulaicity in the more stylistically diverse ICJ Judges' opinions, which are also the least technical and the least hybrid terminologically of all the genres analyzed.

Our results nuance prototypical ideas of institutional legal discourses and what the translation of terminology of key genres of EU and international law entails, beyond the confines of legal instruments and legal terms. The terminological hybridity elicited illustrates the thematic diversity of law and legal translation more broadly, and corroborates both the relevance of legal translation competence in the contexts examined and the versatility expected of institutional translators. Interestingly, the granular data obtained on thematic specializations align with the main trends identified through a large-scale survey on institutional translation profiles and practices also conducted as part of the LETRINT project (Prieto Ramos 2020). This kind of triangulation adds empirical support and insight to our understanding of professional expectations and training needs in the field.

## Funding







Consolidator Grant 157797 (Swiss National Science Foundation) awarded to Fernando Prieto Ramos.

## Acknowledgments


The LETRINT project director and first author of this paper would like to thank the other team members who also contributed to the corpus compilation and annotation work mentioned in this article, in alphabetical order: Morgane Bona-Pellissier, Marie-Hélène Girard, Diego Guzmán and Albert Morales.

## References

- doi Biber, Douglas. 1993. "Representativeness in Corpus Design." *Literary and Linguistic Computing* 8 (4): 243–257.
- Corpas Pastor, Gloria, and Miriam Seghiri Domínguez. 2006. *ReCor*.
- doi Fontanet, Mathilde. 2018. "Translating Hybrid Legal Texts for Science and Technology Institutions: The Case of CERN." In *Institutional Translation for International Governance: Enhancing Quality in Multilingual Legal Communication*, ed. by Fernando Prieto Ramos, 126–138. London: Bloomsbury.
- doi Gries, Stefan T. 2008. "Phraseology and Linguistic Theory: A Brief Survey." In *Phraseology: An Interdisciplinary Perspective*, ed. by Sylviane Granger, and Fanny Meunier, 3–25. Amsterdam: John Benjamins.
- doi Hewson, Lance. 2013. "Is English as a lingua franca translation's defining moment?" *The Interpreter and Translator Trainer* 7 (2): 257–277.
- Kjær, Anne-Lise. 2007. "Phrasemes in legal texts." In *Phraseology. An International Handbook of Contemporary Research* (Vol. 2), ed. by Harald Burger, Dimitrij Dobrovolskij, Peter Kühn, and Neal R. Norrick, 506–515. Berlin: De Gruyter.
- doi Koester, Almut. 2010. "Building Small Specialised Corpora." In *The Routledge Handbook of Corpus Linguistics*, ed. by Michael McCarthy, and Anne O'Keeffe, 66–79. Abingdon: Routledge.
- doi Mayoral Asensio, Roberto. 2007. "Specialised translation: A concept in need of revision." *Babel* 53 (1): 48–55.
- doi McAuliffe, Karen. 2011. "Hybrid texts and uniform law? The multilingual case law of the Court of Justice of the European Union." *International Journal for the Semiotics of Law* 24: 97–115.
- O'Donnell, Mick. 2019. *UAM CorpusTool*. Retrieved from <http://www.corpustool.com/index.html>
- Oostdijk, Nelleke. 1991. *Corpus Linguistics and the Automatic Analysis of English*. Amsterdam and Atlanta: Rodopi.
- doi Prieto Ramos, Fernando. 2014a. "International and supranational law in translation: From multilingual lawmaking to adjudication." *The Translator* 20 (3): 313–331.
- Prieto Ramos, Fernando. 2014b. "Parameters for problem-solving in legal translation: Implications for legal lexicography and institutional terminology management." In *The Ashgate Handbook of Legal Translation*, ed. by Anne Wagner, Kin-Kui Sin, and Le Cheng, 121–134. Farnham: Ashgate.

-  Prieto Ramos, Fernando. 2019. "Implications of Text Categorisation for Corpus-based Legal Translation Research: The Case of International Institutional Settings." In *Research Methods in Legal Translation and Interpreting: Crossing Methodological Boundaries*, ed. by Łucja Biel, Jan Engberg, Rosario Martín Ruano, and Vilemini Sosoni, 29–47. London and New York: Routledge.
-  Prieto Ramos, Fernando. 2020. "The use of resources for legal terminological decision-making: patterns and profile variations among institutional translators." *Perspectives* 29 (2): 278–310.
-  Prieto Ramos, Fernando, Giorgia Cerutti, and Diego Guzmán. 2019. "Building representative multi-genre corpora for legal and institutional translation research: The LETRINT approach to text categorization and stratified sampling." *Translation Spaces* 8 (1): 93–116.
-  Prieto Ramos, Fernando, and Giorgia Cerutti. 2021. "Terminology as a source of difficulty in translating international legal discourses: an empirical cross-genre study." *International Journal of Legal Discourse* 6 (2): 155–179.
-  Prieto Ramos, Fernando, and Diego Guzmán. 2021. "Examining institutional translation through a legal lens: A comparative analysis of multilingual text production at international organizations." *Target* 33 (2): 254–281.
- Ruusila, Anna, and Emilia Lindroos. 2016. "Conditio sine qua non. On phraseology in legal language and its translation." *Language and Law / Linguagem e Direito* 3 (1): 120–140.
- Sandrelli, Annalisa. 2018. "Observing Eurolects: The case of English." In *Observing Eurolects: Corpus analysis of linguistic variation in EU law*, ed. by Laura Mori, 64–92. Amsterdam: John Benjamins.
-  Tirkkonen-Condit, Sonja. 2001. "EU project proposals as hybrid texts: Observations from a Finnish research project." *Across Languages and Cultures* 2 (2): 261–265.
- Trklja, Aleksandar. 2018. "A corpus investigation of formulaicity and hybridity in legal language. A case of EU case law texts." In *Phraseology in Legal and Institutional Settings*, ed. by Anne Wagner, Vijay Kumar Bhatia, Stanisław Goźdz-Roszkowski, and Gianluca Pontrandolfo, 157–169. London: Routledge.

## Address for correspondence

Fernando Prieto Ramos  
 Centre for Legal and Institutional Translation Studies (Transius)  
 Faculty of Translation and Interpreting  
 University of Geneva  
 40, boulevard du Pont-d'Arve  
 CH-1205 Geneva  
 Switzerland  
 Fernando.Prieto@unige.ch  
 <https://orcid.org/0000-0002-4314-2813>



## Co-author information

Giorgina Cerutti

Faculty of Translation and Interpreting

University of Geneva

Giorgina.Cerutti@unige.ch

<https://orcid.org/0000-0002-3260-0866>

## Publication history

Date received: 5 November 2021

Date accepted: 30 December 2021

Published online: 7 July 2022