

<https://doi.org/10.18485/primling.2025.26.12>

81'33:811.131.1

81'33:811.161.1

81'33:811.112.2

Оригинални научни рад

Примљен: 16. 8. 2025.

Прихваћен: 3. 10. 2025.

**Chiara Marchesi\***

ORCID: <https://orcid.org/0009-0009-7139-5824>

Università di Bergamo/Pavia

**Francesco Caprioli\*\***

ORCID: <https://orcid.org/0000-0001-7923-0578>

Università di Bergamo/Pavia

## **STRATEGIES OF INDETERMINACY IN LEGAL DISCOURSE: A CONTRASTIVE ANALYSIS (ITALIAN-RUSSIAN-GERMAN)**

**Abstract:** This paper investigates the use of general and specific extenders in legislative texts across three languages—Italian, Russian, and German—focusing on intellectual property law. Through a contrastive, corpus-based analysis, the study examines their frequency, structure, and degrees of modification. Particular attention is given to specific extenders, which emerge as key tools for managing vagueness. The findings suggest that extenders contribute to the balance between precision and flexibility in normative writing, revealing cross-linguistic patterns and strategies that reflect the functional needs of legal discourse.

**Key words:** legal discourse, general extenders, specific extenders, contrastive analysis, vagueness

\* chiara.marchesi@unibg.it

\*\* francesco.caprioli@unibg.it



## 1. Introduction

This paper<sup>1</sup> investigates the use of general and specific extenders (from now on: GE and SE) in legislative texts across Italian, Russian, and German, with a particular focus on intellectual property law. It aims to analyze their form, structure, and communicative function in written legal discourse, a domain in which such expressions—typically studied in spoken discourse—have received limited attention. While Italian has been more extensively studied in this regard, Russian and German lack dedicated or contrastive research on extenders. The paper highlights their primary referential function and analyses them based on a scalar model to evaluate the degree of intentional vagueness conveyed by extenders, based on their structural complexity.

### 1.1 Definition and use of extenders

The definition of GEs and SEs was developed by Overstreet (1999) and is the most widespread in the literature (Mauri, 2017). GEs are expressions of the type *and things like that*, following the structure [conjunction + pro-form + similarity marker] (cf. also Overstreet & Yule, 2021; Fiorentini, 2018), whereas SEs include more specific lexical content, such as in *and other materials*, formed by [conjunction + adjective + noun phrase], enabling the construction of context-specific, *ad hoc* categories that include both explicit and implicit list members (Frade, 2005; Mauri, 2017; Roma, 2019). Based on the type of conjunction, GEs and SEs can be adjunctive or disjunctive (Overstreet, 1999). While GEs have been widely analyzed, SEs and written occurrences remain underexplored. Most prior scholarship on GEs has focused on spoken discourse, where these expressions exhibit a range of pragmatic functions beyond the basic referential role. In particular, GEs in interaction often serve interpersonal and intersubjective functions, such as softening assertions, managing turn-taking, or expressing shared knowledge. These uses tend to be more prominent than their referential role. However, for Italian, the studies of

<sup>1</sup> The present paper is a condensed, slightly revised, and integrated English version of Marchesi/Caprioli (2025). It has been conceived by both authors and the content has been approved by both authors. Their contribution can be specified as follows: Chiara Marchesi wrote 1., 4.1, 4.2, 5. (jointly written) and 6. (jointly written); Francesco Caprioli wrote 2., 3., 4.3, 5. (jointly written) and 6. (jointly written). Chiara Marchesi is responsible for the data in Italian and in Russian; Francesco Caprioli is responsible for the data in German.

Pecorari (2022) and Roma (2019) offer significant insights into the use of extenders in written discourse, especially in formal and legal contexts.

Pecorari (2022) analyzes GEs in the PUNT-IT corpus—comprising various written genres—and finds that extenders are significantly less frequent and less varied in writing than in speech. In written texts, GEs primarily serve a referential function, marking the non-exhaustiveness of lists, unlike in speech where they often serve interactional purposes. Moreover, the planned nature of writing limits the need for *ad hoc* category construction, a key function of extenders in oral discourse.

Roma (2019)'s investigation of GEs and SEs is based on the BoLC and Jus Jurum corpora. Roma (2019) demonstrates that extenders do occur in legal language and play an important categorizing function. They help signal *ad hoc* categories, allowing implicit items to be included without exhaustive listing, which is crucial for legal interpretation. Drawing on Barotto & Mauri (2018), Roma (2019) uses the concept of property clues: contextual elements that help readers infer the intended category. In GEs, the clue is usually a preceding hypernym, while in SEs, the hypernym is embedded in the noun phrase. In this way, extenders help define legal case types, enabling specific cases to be subsumed under general legal categories while maintaining both precision and flexibility, which is essential in normative texts (cf. also Heller, 2005).

For this paper, the work of Frade (2005) is also of interest. Frade (2005) investigates a type of multinomial expression defined as “legal multinomials” in English legal texts (cf. also Heller 2005, pp. 373–374), of which *fees, disbursements, court costs or other expenses* is a possible example. These expressions are formed by a sequence of exemplars (*fees, disbursements, court costs*) followed by a vague tag (*or other expenses*), whose structure coincides with the one we here indicated for the SEs. Legal multinomials are governed by the *ejusdem generis* principle, according to which the lexical linkage works by similarity and inclusion. This means that, if a series of specific words are followed by generic phrases, these phrases should only be interpreted in reference to the aforementioned items. According to Frade (2005, p. 145), the vague tag in legal multinomials works as an economical device that limits the number of expressed exemplars, while ensuring conciseness and reducing the complexity of legal wording.

Together, the studies by Pecorari (2022), Roma (2019) and Frade (2005) underscore the importance of examining extenders in written

discourse, not only to compare them with their spoken counterparts but also to identify modality-specific features.

## 1.2 A scalar model analysis for extenders in legal discourse

Several studies that have analyzed extenders in spoken language have interpreted these expressions as strategies for encoding intentional vagueness (Fiorentini, 2018; Ghezzi, 2022; Voghera, 2012). In such cases, the speaker more or less consciously opts for a less precise expression, despite having access to a more specific alternative, for various interactional reasons (Ghezzi, 2022).

It is important to note that both GEs and SEs signal the non-exhaustiveness of the list. Barotto & Mauri (2018) link the property of non-exhaustiveness to referential vagueness, which arises when the referents of a given expression cannot be identified with certainty. However, this vagueness pertains not to the identity of the category itself, but to the identity of its members: while the individual elements may remain unspecified, the shared property underlying the category to infer must be clearly identifiable. If the shared property cannot be correctly identified, the inferential process that enables the construction of the *ad hoc* category may be jeopardised (Fiorentini, 2018; Mauri, 2017).

Following Voghera (2012), who argues that intentional vagueness is not exclusive to spontaneous spoken language but also occurs in written texts, it is reasonable to assert that GEs and SEs in legal language are clear examples of the strategic use of vagueness markers in writing (cf. also Marchesi, 2024). In fact, they enable legislators to strike a balance between precision and flexibility—a key requirement of legal texts, which must be both clear and precise and broad and inclusive (Bhatia et al., 2005; Anesa, 2014).

Through GEs and SEs, the legal drafter can include other items in the relevant *ad hoc* category, apart from those explicitly listed: they can be known elements, omitted for the sake of avoiding excessive detail that would render the law overly long and unclear (Li, 2017), or unknown future elements, which had not yet emerged at the time of drafting but may arise later. Such expressions help prevent legal texts from becoming obsolete too quickly or requiring frequent updates (Garzone, 2005; Engberg & Heller, 2008).

While both GEs and SEs can be considered markers of intentional vagueness, GEs tend to be more vague and indeterminate due to their structure, which lacks specific lexical content. This difference makes it possible to position extenders on a *continuum*, based on the degree of their internal specificity. Firstly, the extent of vagueness can vary depending on the noun following the adjective indicating alterity or similarity in the SEs (cf. Overstreet & Yule, 2021), on the basis of a hypernym-hyponym relationship: for instance, in Italian, *persona* ‘person’ is less specific than *titolare* ‘right holder’. Secondly, the presence of elements that modify the extender (restrictive relative clauses, adjectives or adjectival participles, genitives of specification, etc.) can influence its degree of vagueness. As a matter of fact, the legislative text is characterised by a dynamic interplay of strategies that both specify and de-specify, whereby vague elements are modified by components that add precision (Heller, 2006, pp. 234–238; cf. also Heller 2005), while still retaining a certain degree of vagueness (Engberg & Heller, 2008, pp. 157, 164). Thus, the scale ranges from highly vague prototypical GEs (e.g., *eccetera* ‘etcetera’, *i tak dalee* ‘and so on’, *und so weiter* ‘and so on’) to progressively more specific GEs and SEs, depending on the specificity of the noun phrase and the presence or absence of modifiers (as illustrated in Figure 1). Consequently, the *ad hoc* categories constructed via these extenders will also vary in specificity depending on the nature of the extender that triggers their inference.

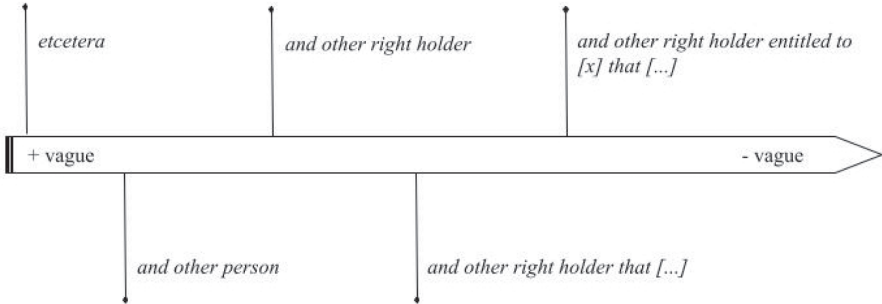


Fig. 1: The *continuum* of specificity of GEs and SEs based on their modifying elements (translated from Marchesi & Caprioli 2025, p. 30)

## 2. Corpus

For the purposes of this analysis, the focus was placed on legislation concerning intellectual property rights (IPR), a particularly relevant area due to its high demand for vagueness strategies. These strategies help prevent the need for constant legal updates in a field that is continually evolving as a result of ongoing innovation. A comparable multilingual corpus was compiled, consisting of three subcorpora, one for each language under investigation. Overall, the corpus comprises 227,202 words, distributed as follows<sup>2</sup> (see also Table 1):

- the Italian subcorpus includes the *Legge sul diritto d'autore* ('Law on Copyright') of 1941 (Law No. 633/1941) and the *Codice della proprietà industriale* ('Industrial Property Code', CPI), totalling 106,008 words;
- the Russian subcorpus is composed of Part IV of the Civil Code of the Russian Federation, titled *Prava na rezult'aty intellektual'noj dejatel'nosti i sredstva individualizacii* ('Rights to the results of intellectual activity and means of individualisation'), which is entirely dedicated to IPR (abbreviated as Čast' IV GK). It contains 71,410 words;
- the German subcorpus comprises the Federal Republic of Germany's laws on copyright (*Gesetz über Urheberrecht und verwandte Schutzrechte*, UrhG) and patents (*Patentgesetz*, PatG), totalling 49,784 words.

	Number of words
<b>Italian subcorpus</b>	106,008
<b>Russian subcorpus</b>	71,410
<b>German subcorpus</b>	49,784

Table 1: Number of words in each subcorpus

## 3. Method

Each subcorpus was automatically annotated for part-of-speech tagging using the SketchEngine platform (Kilgarriff et al., 2014), which

<sup>2</sup> All laws are retrieved in their most recent version as of September 2024.

was also used for all queries and frequency counts presented in the study. To identify GEs and SEs, adjectives expressing alterity or similarity were searched for in each language-specific subcorpus. For Italian, we searched for the forms previously categorized as GEs in the literature (e.g. *eccetera* ‘etcetera’, *ecc.* ‘etc.’, *e simili* ‘and the like’) (cf. Fiorentini, 2018), as well as the adjectives *altro* ‘other’, *simile* ‘similar’, *analogo* ‘similar’, *ulteriore* ‘additional’. For Russian, the search parameters were based on those defined in Marchesi (2024). In particular, for GEs, expressions such as *i tak dalee* ‘and so on’, *i t.d.* ‘and so on’, *i tomu podobnoe* ‘and something like that’, *i drugie* ‘and others’ were searched (cf. also Benigni, 2018; Malyuga & McCarthy, 2018); for SEs, the query concerned the adjectives *inoj* ‘other’, *drugoj* ‘other’, *pročij* ‘other’, *ostal’noj* ‘remaining’ and *analogičnyj* ‘analogous’. For German, we retrieved the forms of GEs cited in prior studies (e.g., *und so weiter* ‘etcetera’, *und/oder so (was)* ‘and/or like (that)’, *und alles* ‘and everything’, *und und und* ‘and and and’, cf. Overstreet 2005; Terraschke & Holmes, 2007), and the adjectives *andere* ‘other’, *weitere* ‘further’, *sonstig* ‘other’, *ähnlich* ‘similar’, *zusätzlich* ‘additional’, *anderweitig* ‘other’, *übrig* ‘residual’.

From the resulting data, occurrences that could structurally and functionally be classified as actual GEs or SEs were manually extracted.

## 4. Results<sup>3</sup>

In this section, the results for each language are presented in turn: Italian (4.1), Russian (4.2) and German (4.3).

### 4.1 Italian

A total of 291 extenders were identified in the Italian subcorpus: 3 GEs and 288 SEs. The small number of GEs is entirely expected, since GEs are more typical of spoken discourse. The three occurrences of GEs are: *o di altra natura* (lit. ‘or of other nature’) (2 occurrences) and *e simili* (‘and the like’) (1 occurrence). Neither Roma (2019) nor Pecorari (2022) mention the GE *o di altra natura*. In our opinion, it should be considered a GE, rather than a SE, since the noun *natura* (‘nature’) is highly generic and not semantically linked to the items it extends.

<sup>3</sup> Whereas the results are identical, the examples presented here differ from those provided in Marchesi & Caprioli (2025), where possible.

Both disjunctive and adjunctive SEs were found. The former are introduced by *o* ‘or’ (168 occurrences, 58.3%) and *ovvero* ‘or’ (1 occurrence, 0.3%); while the latter by *e* ‘and’ (107 occurrences, 37.2%) and *nonché* ‘as well as’ (4 occurrences, 1.4%). Furthermore, 8 instances (2.8%) were identified where no conjunction is used. The adjective of the SEs meaning alterity or similarity are as follows:

- *altro* ‘other’ (269 occurrences, 93.4%);
- *analogo* ‘analogous’ (15 occurrences, 5.2%);
- *simile* ‘similar’ (3 occurrences, 1%);
- *ulteriore* ‘additional’ (1 occurrence, 0.3%).

To estimate the degree of specificity of the SEs, i.e. to locate them along the *continuum* described in §1.2, we considered their possible modifying elements. The following degrees of modification were recognized, from the most indeterminate to the least indeterminate:

- SEs without modifiers (74 occurrences, 25.7%), as in (1):

(1)

Quando i titolari dei diritti chiedono al prestatore di servizi di condivisione di contenuti online di disabilitare l’accesso a loro specifiche opere **o ad altri materiali** o di rimuoverli, indicano i motivi della richiesta. (Law No. 633/1941, Art. 102-decies, p. 1)

‘When right holders request the online content sharing service provider to disable access to their concrete works **or other materials** or to remove them, they shall state the reasons for the request.’

- SEs modified by one element (145 occurrences, 50.3%). In one case only, the modifier occurs to the left and is an adjective; in the remaining cases, it occurs to the right and can be a participial construction, a relative clause, a prepositional phrase, an adjective or an adjunctive connective followed by another noun phrase. In (2) the SE *e altri segni* ‘and other signs’ is modified by the adjective *non registrabili* ‘non-registrable’.

(2)

Marchi di forma **e altri segni non registrabili** (CPI, Art. 9, heading) ‘Shape marks **and other non-registrable signs**’

- SEs modified by two elements (67 occurrences, 23.3%). The SEs are modified by at least two of the aforementioned elements and it is also possible to find a conditional clause. They can be placed either both to the right or one to the left and one to the right.

For instance, in (3) the SE *e degli altri titolari* ‘and of the other holders’ is modified on the right by the prepositional phrase *dei diritti* ‘of the rights’ and a relative clause.

(3)

Le organizzazioni di cui al comma 1 devono indicare, in qualsiasi utilizzo dell’opera orfana, nelle formule d’uso, il nome degli autori **e degli altri titolari dei diritti che sono stati individuati**. (Law No. 633/1941, Art. 69-bis, p. 4)

‘The organisations referred to in point 1 must indicate, in any use of the orphan work, in the usage formulas, the names of the authors **and other right holders that have been identified**.’

- SEs modified by three elements (2 occurrences, 0.7%), which are placed one to the left and two to the right, as shown in (4): the SE *e degli altri detentori* ‘and of the other holders’ is modified on the left by the adjective *precedenti* ‘previous’ and on the right by two specifying prepositional phrases.

(4)

Le informazioni di cui al comma 1 possono tra l’altro comprendere il nome e indirizzo dei produttori, dei fabbricanti, dei distributori, dei fornitori **e degli altri precedenti detentori dei prodotti o dei servizi** [...]. (Law No. 633/1941, Art. 156-ter, p. 2)

‘The information referred to in paragraph 1 may *inter alia* include the names and addresses of producers, manufacturers, distributors, suppliers **and other previous holders of the products or services** [...].’

## 4.2 Russian

In the Russian subcorpus, we identified 266 extenders, of which 17 were GEs and 249 were SEs. In particular, four different types of GEs were found:

- *i tomu podobnoe* ‘and something like that’ (8 occurrences);
- *i drugie* ‘and others’ (4 occurrences);
- *i tak dalee* ‘and so on’ (3 occurrences) as in (5);
- *i t.p.* ‘and something like that’ (2 occurrences).

(5)

Правила абзаца первого настоящего пункта распространяются на наследников автора, их наследников **и так далее** в пределах срока действия исключительного права. (Čast’ IV GK, Art. 1284, p. 1, par. 3)

‘The rules of the first paragraph of this point shall apply to the author’s heirs, their heirs **and so on** within the term of validity of the exclusive right.’

Among the SEs, we identified both disjunctive forms—introduced by *ili* ‘or’ (123 occurrences, 49.4%) or *libo* ‘or’ (20 occurrences, 8%)—and adjunctive forms—introduced by *i* (62 occurrences, 24.9%) or *a takže* ‘and also’ (15 occurrences, 6%). There are also 29 SEs (11.6%) that do not have a conjunction. The following adjectives are found within the SEs:

- *inoj* ‘other’ (175 occurrences, 70.3%);
- *drugoj* ‘other’ (69 occurrences, 27.7%);
- *analogičnyj* ‘analogous’ (5 occurrences, 2%).

Depending on the degree of their internal specificity, the SEs can be placed along a *continuum* of more or less determinacy. We identified the following groups:

- SEs without modifiers (56 occurrences, 22.5%), as in (6).

(6)  
Полномочия патентного поверенного **или иного представителя** удостоверяются доверенностью. (Čast’ IV GK, Art. 1247, p. 2, par. 3)  
The powers of a patent attorney **or other representative** shall be certified by a power of attorney.

- SEs modified by one element (146 occurrences, 58.6%), which can be placed either to the left or to the right. These elements can be adjectives, participial constructions, prepositional phrases, genitives that specify the noun, relative clauses, adverbs, conditional clauses. As a modifying element, we also considered the noun, which forms a compound with the noun of the SEs, as in (7), where *obladatel* ‘holder’ is modified by *patent* ‘patent’ to form a compound.

(7)  
Нарушение прав автора селекционного достижения **или иного патентообладателя** (Čast’ IV GK, Art. 1446, heading)  
The infringement of the rights of the author of a breeding achievement **or of other patent holder**

- SEs modified by two elements (40 occurrences, 16.1%), which can be placed both to the right, both to the left or one to the left and one to the right. These items can consist of two of the aforementioned elements, but an infinitive construction can also

be found. In (8) the noun of the SE *sredstva* ‘means’ is modified on the left by an adjective and on the right by a genitive.

(8)

Топология, созданная работником с использованием денежных, технических **или иных материальных средств работодателя**, но не в связи с выполнением своих трудовых обязанностей или конкретного задания работодателя, не является служебной. (Čast’ IV GK, Art. 1461, p. 5)

A layout-design created by an employee using **the employer’s financial, technical or other material means**, but not in connection with the performance of their job duties or a specific task assigned by the employer, shall not be deemed a service layout-design.

- SEs modified by three elements (6 occurrences, 2.4%), which can be placed one to the left and two to the right or vice versa. In (9) the noun *prava* ‘rights’ is modified by two adjectives on the left and by a prepositional phrase on the right.

(9)

При использовании результата интеллектуальной деятельности в составе сложного объекта за автором такого результата сохраняются право авторства **и другие личные неимущественные права на такой результат**. (Čast’ IV GK, Art. 1240, p. 3)

When using the result of intellectual activity as part of a complex object, the author of such result retains the right of authorship **and other personal non-property rights to such result**.

- SEs modified by four elements. There is just one instance (0.4%), where the noun of the SE *lico* ‘person’ is modified by two adjectives on the left and two participial constructions on the right, as in (10).

(10)

Принадлежащее Российской Федерации исключительное право на результат интеллектуальной деятельности [...] может быть передано исполнителю **или другому российскому юридическому лицу, заинтересованному во внедрении результата интеллектуальной деятельности и обладающему возможностями для его внедрения, [...]**. (Čast’ IV GK, Art. 1240.1, p. 3, par. 3)

The exclusive right of the Russian Federation to the results of intellectual activity [...] may be transferred to the performer **or another Russian legal entity interested in implementing the results of intellectual activity and having the capacity to do so, [...]**.

### 4.3 German

In the German subcorpus, no GEs were found, whilst 84 occurrences of SEs were identified. These SEs are either disjunctive—introduced by *oder* ‘or’ (59 occurrences, 70.2%)—, adjunctive—introduced by *und* ‘and’ (20 occurrences, 23.8%) or *sowie* ‘as well as’ (2 occurrences, 2.4%), or they are not introduced by any conjunction (3 occurrences, 3.6%). The adjectives indicating similarity or alterity found within the SEs are as follows:

- *andere* ‘other’ (43 occurrences, 51.2%);
- *sonstig* ‘other’ (31 occurrences, 36.9%);
- *ähnlich* ‘similar’ (8 occurrences, 9.5%);
- *weitere* ‘other’ (2 occurrences, 2.4%)

To analyse SEs along the scalar continuum of specificity, we considered the presence of modifying elements. Given the productivity of compounding in German (cf. Bosco-Coletos, 2007 among others), compounding is treated here as a structural strategy for reducing indeterminacy. Subsequently, the SEs can be grouped as follows:

- SEs without modifiers (17 occurrences, 20.2%), as in (11).

(11)

Das Deutsche Patent- und Markenamt besteht aus einem Präsidenten **und weiteren Mitgliedern**. (PatG, Art. 26, p. 2)

‘The German Patent and Trademark Office consists of a president and other members.’

- SEs modified by one element (41 occurrences, 48.8%). The modifier can occur either to the left (a compound modifier, an adjective, a participial construction) or to the right (a genitive noun, a prepositional phrase, a relative clause, an additional coordinated noun phrase). In (12), the noun of the extender *Nutzer* ‘user’ is modified by the compound modifier *Werk-* ‘work’.

(12)

Auf eine Vereinbarung, die zum Nachteil des Urhebers von Absatz 1 abweicht, können sich der Vertragspartner des Urhebers **oder andere Werknutzer** nicht berufen. (UhrG, Art. 32f, p. 2)

‘The contractual partner of the author **or other users of the work** may not invoke any agreement that deviates to the detriment of the author referred to in paragraph 1.’

- SEs modified by two elements (23 occurrences, 27.4%). At least two elements seen above are present; they can be placed one to the left and one to the right, both to the left or both to the right. In (13) the noun *Nutzungen* ‘uses’ is modified by an adjective on the left (*digitalen* ‘digital’) and a prepositional phrase on the right (*gemäß § 60a* ‘pursuant to Section 60a’). (13)

In den Fällen der öffentlichen Wiedergabe nach den §§ 46, 48, 51, 60a bis 60d, 61, 61c, 61d und 61f **sowie bei digitalen sonstigen Nutzungen gemäß § 60a** ist die Quelle einschließlich des Namens des Urhebers stets anzugeben, es sei denn, dass dies nicht möglich ist. (UhrG, Art. 63, p. 2) ‘In cases of public reproduction pursuant to Sections 46, 48, 51, 60a to 60d, 61, 61c, 61d and 61f, **as well as in cases of other digital uses pursuant to Section 60a**, the source, including the name of the author, must always be indicated, unless this is not possible.’

- SEs modified by three elements (3 occurrences, 3.6%), that can be placed two to the left and one to the right or two to the right and one to the left. In (14) the noun *Elementen* ‘elements’ is modified by an adjective on the left (*unabhängigen* ‘independent’) and by two distinct relative clauses on the right. (14)

Datenbank im Sinne dieses Gesetzes ist eine Sammlung von Werken, Daten **oder anderen unabhängigen Elementen, die systematisch oder methodisch angeordnet und einzeln mit Hilfe elektronischer Mittel oder auf andere Weise zugänglich sind und deren Beschaffung, Überprüfung oder Darstellung eine nach Art oder Umfang wesentliche Investition erfordert.** (UhrG, Art. 87a, P. 1) ‘For the purposes of this Act, a database is a collection of works, data **or other independent elements that are arranged in a systematic or methodical manner and are individually accessible by electronic means or otherwise, and whose acquisition, verification or presentation requires a substantial investment in terms of type or scope.**’

## 5. Discussion

In order to compare the data we calculated the normalized frequencies of GEs and SEs (per million words) for each subcorpus:

- IT: 28 GEs and 2,717 SEs;
- RU: 238 GEs and 3,487 SEs;
- DE: 0 GE and 1,687 SEs.

The comparison reveals notable quantitative differences in the use of extenders. Russian legal texts exhibit substantially higher use of GEs—approximately 8.5 times more than Italian legal texts; no GEs were found in German legal texts. SEs are also more frequent in Russian than in Italian or German, though the difference is less stark: SEs in Russian occur about 1.3 times more than in Italian and twice as often as in German, while Italian SEs are about 1.6 times more frequent than German SEs. The unusually high number of GEs in Russian may be attributed to legal and historical reasons: the Russian legal code was compiled rapidly to expedite Russia’s accession to the World Trade Organization (Eugster 2010), which may have led to a reliance on expressions that are less typical of legal language. Furthermore, the fact that Russian law is embedded within the Civil Code, which typically contains more general provisions, may also have played a role. Additionally, the greater similarity between Italian and German data may reflect shared legal frameworks, such as EU legislation.

Despite quantitative variation across languages, all three show a common trend in SE types according to the conjunction used. Disjunctive SEs are the most frequent in each language, always accounting for more than half of the total; conjunctive SEs occur less frequently, typically comprising about one-third; SEs without a conjunction are rarer in all languages. This widespread preference for disjunctive SEs may reflect a legislative tendency to list alternative possibilities rather than mutually inclusive ones.

Concerning the *continuum* of indeterminacy, certain tendencies can be observed. Overall, the vast majority of SEs are modified by no more than one or two elements, indicating a consistent approach to managing vagueness in legislation across the three languages. In particular, in most cases, SEs are modified by one element. In Italian and Russian, the second-largest group consists of unmodified SEs, while in German, it’s the

group consisting of SEs modified by two elements; the opposite situation is witnessed for the third-largest group. In all languages, SEs with three modifiers form the fourth-largest group, and only in Russian do we find SEs modified by four elements, though there is just one occurrence.

In all three languages, SEs typically begin with adjectives indicating alterity (IT: 93.8% of SEs; RU: 98% of SEs; DE: 90.5% of SEs), while adjectives expressing similarity (IT: 6.2% of SEs; RU: 2% of SEs; DE: 9.5% of SEs) are much less frequent.

The position of modifiers varies across the three subcorpora but generally reflects the typical syntactic patterns of each language. When the SEs are modified by one element:

- in Italian, the modifier appears almost exclusively to the right (99.3%);
- in Russian, modifier placement is evenly split (51.4% left / 48.6% right);
- in German, modifiers are predominantly on the left (63.4%).

When the SEs have two modifiers, we observed that, in Italian, they never appear on the left; instead, they nearly always appear on the right, and, just in one case, there is one modifier on the left and one on the right. In Russian, the majority of cases have both modifiers on the right (62.5%); fewer instances have one modifier on the left and one on the right and even fewer have both modifiers on the left. In German, most SEs have one modifier on the left and one the right (78.3%), while instances of both modifiers on the same side are less common.

Modifiers appear to be slightly more diverse in Russian, with ten different categories identified, compared to seven in Italian and German. This suggests a somewhat greater structural variety in how SEs are specified in the Russian subcorpus.

## 6. Conclusion

Our research suggests that, although extenders have mostly been studied in spoken language and are generally considered more typical of that modality, they—especially SEs—may also serve as a relevant tool for vagueness in written, highly controlled genres such as legal texts. While their frequency and forms differ from those found in spontaneous

speech, their presence in legislative writing should not be overlooked. In contrast to Pecorari (2022), we argue that extenders play a pivotal role in constructing *ad hoc* categories in legal texts. This observation is aligned with Roma (2019), who emphasises the importance of text genre as a key variable in understanding the use of extenders.

Our analysis tested the hypothesis that the addition of one or more modifiers can serve as a deliberate legislative strategy to modulate the degree of vagueness on a *continuum* of more or less indeterminacy. Despite some notable disparities—such as the markedly higher use of GEs and SEs in Russian legislation compared to Italian and German—and differences stemming from the underlying linguistic and legal systems, our contrastive analysis revealed consistent cross-linguistic tendencies in how extenders are employed. These findings lay the groundwork for future research, which may involve extending this type of analysis to a broader range of legal texts, diverse written genres, and additional languages, with the goal of identifying recurring patterns, differences, or shared structural features.

## References:

- Anesa, P. (2014). Defining Legal Vagueness: A Contradiction in Terms?. *Pólemos*, 8(1), 193–209.
- Barotto, A., & Mauri, C. (2018). Constructing lists to construct categories. *Italian Journal of Linguistics*, 30(1), 95–134.
- Benigni, V. (2018). Ad hoc categorization in Russian and multifunctional general ex-tenders. *Folia Linguistica Historica* 39, 97–123.
- Bhatia, V. K., Engberg, J., Gotti, M., & Heller, D. (Eds.) (2005). *Vagueness in Normative Texts*. Lang.
- Bosco-Coletos, S. (2007). *Il tedesco lingua compatta*. Edizioni dell'Orso.
- Engberg, J., & Heller, D. (2008). Vagueness and Indeterminacy in Law. In Bhatia, V. K., Candlin, C., & Engberg, J., *Legal Discourse across Cultures and Systems* (pp. 145–168). Hong Kong University Press.
- Eugster, E. (2010). Evolution and enforcement of intellectual propriety law in Russia. *Washington University Global Studies Law Review*, 9(1), 131–151.
- Fiorentini, I. (2018). Eccetera eccetera e così via di seguito. I *general extenders* dell'italiano contemporaneo. *CLUB Working Papers in Linguistics*, 2, 20–39.

- Frade, C. (2005). Legal Multinomials: Recovering Possible Meanings from Vague Tags. In Bhatia, V. K., Engberg, J., Gotti, M., & Heller, D. (Eds.), *Vagueness in Normative Texts* (pp. 133–155). Lang.
- Garzone, G. (2005). La legge e i nuovi media: vaghezza nella legislazione sulla proprietà intellettuale nell'ordinamento inglese. In Jullion, M.-C. (Ed.), *Linguistica e proprietà intellettuale* (pp. 63–75). Franco Angeli.
- Ghezzi, C. (2022). *Vagueness Markers in Italian. Age variation and pragmatic change*. Franco Angeli.
- Heller, D. (2005). Zwischen Bestimmtheit und strategischer Offenheit: Zur sprachlichen Qualifizierung deutscher und schweizerischer Sanktionsnormen. In Bhatia, V. K., Engberg, J., Gotti, M., & Heller, D. (Eds.), *Vagueness in Normative Texts* (pp. 357–378). Lang.
- Heller, D. (2006). *Regolare esplicitamente la libertà di interpretazione – Normativ wirksame Texte im internationalen Vergleich*. In Foschi Albert, M., Hepp, M., & Neuland, E. (Eds.), *Texte in Sprachforschung und Sprachunterricht. Pisaner Fachtagung 2004 zu neuen Wegen der italienisch-deutschen Kooperation* (pp. 230–242). iudicium.
- Kilgarriff, A., Baisa, V., Bušta, J., Jakubiček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine: ten years on. *Lexicography*, 1, 7–36.
- Li, S. (2017). A corpus-based study of vague language in legislative texts: Strategic use of vague terms. *English for Specific Purposes*, 45, 98–109.
- Malyuga, E., & McCarthy, M. (2018). English and Russian vague category markers in business discourse: Linguistic identity aspects. *Journal of Pragmatics*, 135, 39–52.
- Marchesi, C. (2024). *General e specific extenders nella legislazione russa sulla proprietà intellettuale: un caso di deviazione dalla norma*. *Slavia*, 3, 113–124.
- Marchesi, C., & Caprioli, F. (2025). *General e specific extenders nella legislazione sulla proprietà intellettuale in italiano, russo e tedesco*. *Linguistik Online* 135(3), 25–45.
- Mauri, C. (2017). Building and Interpreting Ad Hoc Categories: A Linguistic Analysis. In Blochowiak, C. G., Durrleman, S., & Laenzlinger, C. (Eds.), *Formal Models in the Study of Language* (pp. 297–326). Springer.
- Overstreet, M. (1999). *Whales, Candlelight, and Stuff Like That*. Oxford University Press.
- Overstreet, M. (2005). And stuff und so: Investigating pragmatic expressions in English and German. *Journal of Pragmatics*, 37, 1845–1864.
- Overstreet, M., & Yule, G. (2021). *General Extenders: The Forms and Functions of a New Linguistic Category*. Cambridge University Press.
- Pecorari, F. (2022). I general extenders nel testo scritto, tra lessico e punteggiatura. *Revue Romane*, 57(2), 311–342.

- Roma, E. (2019). Allungalista generici e specifici nei testi giuridici. *Studi Italiani di Linguistica Teorica e Applicata*, 3, 539–566.
- Terraschke, A., & Holmes, J. (2007). ‘Und Tralala’: Vagueness and General Extenders in German and New Zealand English. In Cutting, J. (Ed.): *Vague Language Explored* (pp. 198–220). Palgrave Macmillan.
- Voghera, M. (2012). Chitarre, violino, banjo e cose del genere. In Thornton, A. M., & Voghera, M. (Eds.), *Per Tullio De Mauro. Studi offerti dalle allieve in occasione del suo 80° compleanno* (pp. 341–364). Aracne Editrice.