

# A Structural and Grammatical Analysis of the Estonian Constitution

**Addi Rull**

Tallinn Law School,  
Tallinn University of Technology  
Akadeemia tee 15A,  
Tallinn 12618, Estonia  
E-mail: addi.rull@ttu.ee

**Tõnu Tamme**

Institute of Computer Science,  
University of Tartu  
J. Liivi 2,  
Tartu 50409, Estonia  
E-mail: tonu.tamme@ut.ee

**Leo Võhandu**

Institute of Informatics,  
Tallinn University of Technology  
Akadeemia tee 15A,  
Tallinn 12618, Estonia  
E-mail: leo.vohandu@ttu.ee

**Abstract:** *The authors propose a novel quantitative method to analyse the structure of legal texts. The method enables to determine grammatical similarity between legal texts. The authors use the external theory of fundamental rights to separate the text of fundamental rights of the Estonian Constitution into two categories of norms: constitutional rights and restrictions. Grammatical similarity between constitutional rights, restrictions and selected legal acts and case law is measured. The layer of special norms renders the best grammatical similarity with the text of fundamental rights. The same grammatical similarity tests can be replicated to cover other jurisdictions in the future. The research is experimental, but the authors believe that the method can be utilised in fields of computational linguistics and legal text*

*mining, but also in research where legal text structures are analysed for various purposes.*

**Keywords:** *chi-square test, complexity of constitutional text, grammar of legal text, interpretation of norms, Mosteller's comparison method, structure of norms*

## 1. Introduction

Legal texts are often described as complex and difficult to understand. Unlike literature, these texts are written based on the composition techniques of norms. The literal interpretation method used to understand norms is just one amongst several. Norms are linked to each other by normative reference structures and delegations. Yet, the clarity of the legal system is one of the underlying principles of constitutional law aimed at providing legal certainty.

There is little research available on the complexity of legal texts. Reinsalu (2012) has conducted a qualitative analysis of contracts and administrative acts in Estonia, focusing on the intuitive complexity of sentences and Kerge (2010) has contributed to the research of legal texts. Recently, several authors have used quantitative methods to measure the complexity of laws (e.g., Watl & Matthes, 2014) and legal systems (Smits, 2015). Smits (2015, pp. 1–16) uses four different factors to measure the complexity in law. The main focus is on the density factor. Elements of density can be the number of words or provisions in a legal text. The general proposition is that the clarity of the text can be improved by decreasing the density. Watl and Matthes (2014) show that the Flesch-Reading-Ease<sup>1</sup> (FRE) non-structural index describes only about 10 per cent of the analysed legal text structure. This finding suggests that there is no reason to apply the FRE index to laws. The authors (e.g., Tamme *et al.*, 2014) have previously analysed fictional texts by Estonian writers but also legal texts. The grammar of legal texts is as complicated as the grammar of literary texts and can be more complex because of meta-rules applicable to legal texts.

The current research analyses the provisions of fundamental rights in the Estonian Constitution (1992), which was chosen because the quantitative method

<sup>1</sup> The Flesch-Reading-Ease (FRE) index measures textual difficulty, which indicates how easy a text is to read. The FRE index can be calculated with Microsoft Word. The scale is calculated between 100–0 (100 – very easy to read, 65 – average, 30 – fairly difficult to read, 0 – very difficult to read).

is applied to texts of legal acts and case law available in the Estonian language, but the mathematical methodology used in the current article can be applied to other languages. The application of the quantitative methodology to legal acts as described in the current research is novel. The aim of the application of the methodology is to analyse the grammatical similarity of the text of fundamental rights divided into two categories of text—rights and restrictions—vis-à-vis the texts of legal acts and case law. The outcome of the experimental grammatical similarity tests was difficult to foresee and therefore a clear hypotheses could not be formulated. The authors stipulated that there could be differences in grammatical similarity between the texts of constitutional rights and restrictions measured against legal acts and case law, and if the text of restrictions did not reveal positive results, then it could be disregarded for the purpose of decreasing the density of the constitutional text in programmable applications in the future. In the end, the results rendered a grammatical similarity between constitutional rights, restrictions and the layer of *lex specialis* laws.

Before calculating the grammatical similarity between constitutional rights, restrictions and selected legal acts and case law, the authors had to conduct a thorough legal analysis to verify the use of the formal method. The grammatical similarity tests described at the end of the article reflect the structural composition of legal texts. Therefore we begin with the analysis of the structure of the Estonian Constitution in Section 2. We categorise and calculate the proportion of constitutional norms, selecting fundamental rights for further analysis. Reference structures are discussed in subsection 2.1, because references are an integral part of the structural composition of constitutions and legal interpretation techniques used in the current work are briefed in subsection 2.2.

In Section 3, we apply legal theory and interpretation to separate the text of fundamental rights into two categories of text: rights and restrictions. The external theory of fundamental rights offers a possibility to separate restrictions from the text of fundamental rights that forms 21 per cent of the text of the Constitution. Restrictions to fundamental rights are regulated in other legal acts *outside* constitutions, taking into account the principle of proportionality (e.g., Hickman, 2007). The text of restrictions in constitutions does not contain actual restrictions.

In Section 4, the application of the quantitative method to measure the grammatical similarity of rights and restrictions vis-à-vis selected legal acts is described. The method is based on Mosteller's comparison method and a chi-square test. The results are provided in Tables 10 and 11. The objective was to understand if there are differences in grammatical similarity between

constitutional rights, restrictions and different layers of law. The similarity analysis demonstrates that the layer of special laws renders better similarity than others in relation to the text of fundamental rights.

## **2. The structure of the Constitution**

Constitution is *lex fundamentalis*, meaning that it forms the basis for the hierarchy of legal acts in any given legal system (Fig. 1, see next page). Most legal systems have constitutions in written form, but some legal systems are based on unwritten constitutions (e.g., Rene & Brierley, 1993, pp. 307–396) or a collection of customary norms of superior meaning in native societies (e.g., Moore, 1984, pp. 3–9). Constitutions are the most important legal texts on the basis of which states and legal systems are formed.

The reason why constitutions are rarely changed is because they contain the most important values of a society, constitutional principles that are applied throughout a legal system, and the main state institutions. It is possible to group a constitutional text under these three categories of norms. The Commentaries of the Constitution of the Republic of Estonia provide that values represented by fundamental rights are in 48 paragraphs, which account for more than 25 per cent of the Constitution (Kask, 2012). The calculation shown in Table 1 (see next page) presents a more accurate estimation of 35 per cent for the values. Fundamental rights are not in any hierarchical relation to each other, which justifies the grouping of these norms under one category.

The Preamble of the Constitution encompasses values which mark the reasons why a constitution is enacted based on the will and the power of people, known as the concept of *pouvoir constituant* in legal theory (Maruste, 1997, pp. 51, 64). Preambles are abstract declarations that consist of words such as ‘rule of law’, ‘democracy’, ‘God’, common in many constitutions. Some preambles reflect more specific values. The Preamble of the Estonian Constitution talks about the preservation of the Estonian culture and language. If the legislator or the Supreme Court has to determine the right balance between fundamental rights by weighing one right over the other in a case concerning the Estonian language it may resort to the Preamble as the source of information for interpretation.<sup>2</sup> Values formulated in the Preamble do not strictly follow the same normative structure as fundamental rights.

<sup>2</sup> Constitutional judgement 3-4-1-7-98, 4 Nov 1998.

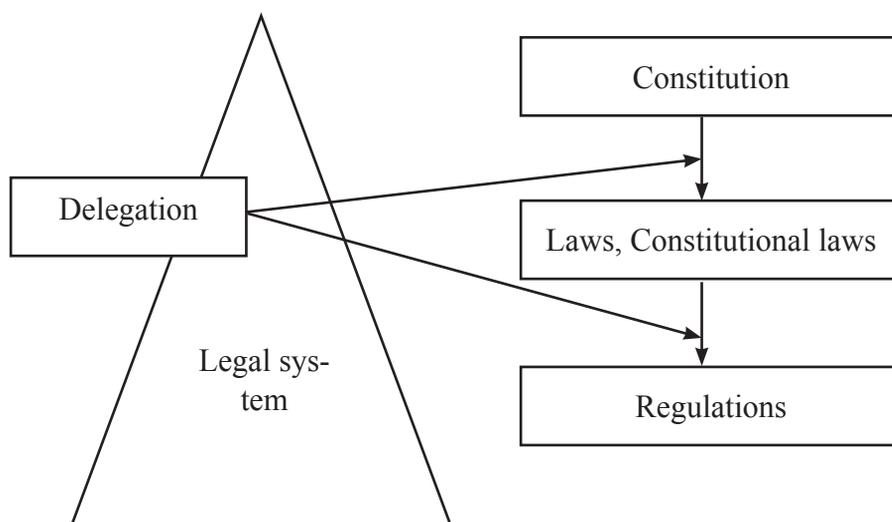
Table 1. The sizes of the Constitution

	no. of sentences	no. of words	per cent
Constitution	698	7,693	100
§ 8–55	248	2,683	35
§ 8–55 values (FR1)	119	1,043	14
§ 8–55 restrictions (FR2)	129	1,640	21

Norms establishing constitutional principles could be summoned under two basic principles: the legality and the legal certainty. The principle of legality is about a delegating norm, in other words, an authorisation derived from a constitution to legislate or execute other legal acts so that they are always in conformity with *lex fundamentalis*, as illustrated in Figure 1. Delegation connects all legal acts into the hierarchy of norms always traceable to the constitution either by express formulations in legal texts or through implied interpretations.

These principles have to be taken into account in the law-making process and applied throughout a legal system, giving the opportunity to challenge any norm which has been enacted contrary to these principles. Most, but not all, constitutional principles are enlisted in seven paragraphs (§§ 1–7) in the Estonian Constitution.

Figure 1. Hierarchy of norms



The third category of norms establishes the most important institutions of a society. In the Estonian Constitution, these are the people, the Parliament (*Riigikogu*), the President, the Government, legislation, finance and the state budget, foreign relations and international treaties, national defence, the state audit office, the legal chancellor (the Ombudsman), the courts, local government, amendment of the Constitution. These norms define the roles of central and local governments and the Parliament which all deliver certain legal acts. The President has a role of conducting a constitutional review which requires him to approve or disapprove legislation made by the Parliament. Chapter VII on legislation presents a list of laws which are constitutional.

## **2.1 Reference structures**

References are a common tool to link norms and to connect different legal acts. References to specific regulations are worded in the form of delegations. For example, the Emergency Act gives the following delegation: the Minister shall establish the statutes (a type of a regulation) and approve the staff of the regional crisis management committee in the area of its competence.<sup>3</sup> This reference is tagged and it directs to the specific list of regulations in *Riigi Teataja* ('The State Gazette').

Laws may contain internal or external references to norms. The same Act provides that restrictions or suspensions of the crossing of the state border may be imposed in the event specified in § 7 of the State Borders Act.<sup>4</sup> This is an external connection pointing to norms of another legal act. Moreover, § 28 of this act provides in subsection (1) that the involuntary treatment of infectious diseases shall be carried out on the basis of procedures established in the Communicable Diseases Prevention and Control Act. Subsection (2) of the same paragraph states that procedures specified in subsection (1) can be further specified by the Government of the Republic regulation. Paragraph 28, subsection (2) contains the internal reference to subsection (1).

The State Gazette provides connection tags only for specific regulations, but not for laws or internal references of norms. Nõlv (2016) has developed the *xLaw* application (available at [www.xlaw.eu](http://www.xlaw.eu)) for the Estonian legal system with 4,784 reference links between 373 legal acts. Karstens (2014) analysed the internal reference structures of national constitutions and suggests that it may be possible to reduce the complexity and increase the readability of the text by reducing the number of internal references.

<sup>3</sup> State of Emergency Act, RT I 2009, 39, 262; 01.09.2015, 6), § 4 (4).

<sup>4</sup> State of Emergency Act, RT I 2009, § 31 (1), 5.

Reference structures are the formations that link legal acts and norms, but only delegations are compulsory, because they establish what the government can or cannot do. Other reference structures are voluntary. Therefore the frequency and patterns of reference structures vary a lot from one legal system to another (Karstens, 2014).

## **2.2 Interpretation of the Constitution**

Different interpretation techniques are often used to understand the meaning of legal texts. These are textual, systematic, teleological and historic interpretation methods. Different complementary techniques such as logic or the requirement to maintain the integrity of the text are considered. The case law of the Constitutional Court is a source of interpretation.

Different views on how a constitution should be interpreted may emerge. Balkin (2008) studied different schools of thought called Originalism and Living Constitutionalism in the American jurisprudence. The former group looks for the original meaning of the Constitution and the latter accepts deviations from the original meaning, claiming that the Constitution is a dynamic embodiment of law. The original meaning of the Constitution is revealed in the study about the life histories of the authors of the text, as William Ewald (2008) has done with the mail correspondence of James Wilson, one of the drafters of the Constitution of the United States of America.

In the current study we use textual, systematic, teleological and historic interpretation of the constitutional text. Finally, we apply a mathematical method in order to determine grammatical similarity between the selected constitutional norms, legal acts and case law. We have categorised norms by their content, function and purpose. The study of the text and grammar (textual method) of fundamental rights is followed by systematic interpretation whereby we have established common structures of rights and thereafter analysed the purpose of fundamental rights in the context of legal theory (teleological method). We take into consideration the fact that, although historically relevant, the internal theory of fundamental rights can be disregarded.

### **3. The structure and interpretation of fundamental rights**

The interpretation of fundamental rights allows dividing them into two separate texts. The right to privacy in § 26 of the Estonian Constitution provides the following:

*Everyone has the right to the inviolability of private and family life. State agencies, local governments, and their officials shall not interfere with the private or family life of any person, except in the cases and pursuant to procedure provided by law to protect health, morals, public order, or the rights and freedoms of others, to combat a criminal offence, or to apprehend a criminal offender.*

The first sentence “[e]veryone has the right to the inviolability of private and family life” can be separated from the rest. It denotes the first level of the right, complying with the two-level interpretation of fundamental rights. Rights and freedoms in the European Convention on Human Rights (1950) are already divided by ‘1’ and ‘2’ into two texts.

In legal theory, fundamental rights are interpreted to contain single-level structure, two-level structure and two-level structure together with single-level positive obligations imposed on a state (Ruiz, 1997, pp. 92–100). Alexy (2001) describes these legal doctrines as internal (single-level) and external (two-level) approaches to the interpretation of fundamental rights.

The single-level interpretation means that the content of a fundamental right is not separated by disconnecting the positive dimension of the right provided in the first level from the negative dimension in the second level.<sup>5</sup> Fundamental rights are formulated so that the first sentence(s), or occasionally, subclause has the positive dimension. As in the example above, one has the right to privacy. If the rest of the content, which talks about the interference of one’s private life being the negative dimension, is disregarded, then it can be understood that the first level of the right does not contain restrictions. Most fundamental rights are restricted. As in this case, one’s private life can be intruded for the purpose

<sup>5</sup> The rationale of the structure of fundamental rights with its positive and negative dimensions is explained by the authors of the current article who undertook the work to separate all fundamental rights in the Estonian Constitution into two separate texts by studying the text of each fundamental right and deciding which part of the text composes a right (positive dimension) and a restriction (negative dimension), whereas the explanations of internal and external theories of fundamental rights originate from Alexy (2001) and Ruiz (1997).

of crime investigation. Restrictions are specified in different legal acts. The single-level interpretation means that there is no rigid division of the content of fundamental rights into positive and negative dimensions and separating them from the rest of the legal system when restrictions to rights are considered. In other words, fundamental rights are interpreted to contain inherent limitations. The single-level interpretation has been largely abandoned since the beginning of 1970s (Ruiz, 1997, pp. 92–100).

On the other hand, the two-level interpretation means that the positive dimension of the right remains without boundaries on the constitutional level. In the context of § 26 it means that there is a theoretical possibility to exercise the maximum degree of privacy, although this is not possible in practice. Such an interpretation is practical, because judges and legislators are daily confronted with the problem of balancing rights and they have to do it in the manner which limits rights as little as possible (e.g., Martin & Horn, 2008). The first level of the right is interpreted to have the widest scope of coverage which enables to adjust the boundaries of the right as needed. The balancing of rights takes place on the second level of the right, but outside the constitution on the functional level of the legal system, using different legal acts.

The two-level interpretation together with the single-level positive obligation adds the social dimension (Singh, 2008). One of the rights listed in the Constitution is the right to education as provided in § 37. The positive obligation of the state means that it must provide the necessary infrastructure and legal framework for people to be able to exercise their right to education. The state must build schools and provide the regulatory framework of financing education, etc. The single-level positive obligations are derived from the first level of the right.

We have taken the two-level interpretation of fundamental rights as the basis for dividing all paragraphs of fundamental rights in the Estonian Constitution into two separate texts. In Table 1 (see p. 80), the first level is labelled FR1 and the second level FR2. The first level accounts for 14 per cent and the second level 21 per cent by the word count. The theory of fundamental rights suggests that only the first level has importance, because the second level only indicates that limitations are in laws, regulations, that is, other legal acts. The elimination of the second level does not change the function of fundamental rights, because the second level expresses only one thought—limitations are regulated in laws—which is now an implied interpretation even if we only maintain the first level of the right. Such an interpretation is possible because we disregard the traditional single-level theory discussed above, and the single-level positive obligations

theory, applied together with the two-level liberal theory is only derived from the first level of the right.

In the light of the external theory of fundamental rights FR2 is redundant, because it does not contain actual restrictions. Yet, the text of restrictions may have a meaning conveyed by “entailed norms” (e.g., Navarro & Rodriguez, 2015). The commentators of the Estonian Constitution explain that some rights must be more difficult to limit than others, depending on how restrictions have been formulated or in cases they are absent (Commented edition, 2002, Ch. II, 8.2; 8.3). This interpretation is valid but also problematic because it anticipates the importance given to some values over the other. We do not question the intentions of the framers of the Estonian Constitution, but propose a thorough analysis how consistently this interpretation is applied in practice. Internationally, the common understanding is that the right to life, the prohibition of torture and slavery are rights that cannot be derogated and must not include restrictions (European Convention..., 1950, Art. 2, 3, 4, 15). In principle, academics have agreed that these are ultimate values that cannot be limited, although acts such as abortion or euthanasia raise questions about the possibility of unlimited rights and freedoms (Uibopuu, 2000, pp. 25–35).

The purpose of the analysis of theories of fundamental rights was to verify that separating fundamental rights into two categories of text does not conflict with legal logic and interpretation of constitutional norms. Additionally, it was needed to assure that the results obtained by the application of the formal method, as described in Section 4 below, are interpretable in the context of constitutional norms and different layers of legal acts and case law.

#### **4. Method of testing structural similarity**

The comparison of structural similarity is performed in order to understand whether there is any difference in the results of the text of FR1 and FR2 in relation to selected legal acts. The Law of Obligations Act (2001; abbr. *VÕS*) covers private matters and the Penal Code (2001; abbr. *KaS*) is a public law that covers all criminalised acts. The Medicinal Products Act (2005; abbr. *RavS*) and the Private Schools Act (1998; abbr. *EraKS*) have been chosen to represent *lex specialis*, regulating specific areas of life in more detail. The constitutional judgements 3-4-1-2-13 and 3-4-1-26-14 are chosen, because these judgements talk about specific norms provided in two special laws mentioned. The selection covers general norms, special norms and certain specific norms discussed in case

law. The results of the test are provided in Tables 10 and 11. The Penal Code that has 451 paragraphs and 30,890 words is chosen to explain the application of the method.

The text was analysed morphologically using the Estnltk Python package (GitHub, 2016). Each word is labelled, describing its type. Labels are defined as following: V (verb), S (noun), A (adjective), P (pronoun), D (adverb), K (adpositional), J (conjunctive), N (numeral), I (interjection), Y (abbreviation), Z (punctuation mark), X (verbal attendant), G (genitive attribute), O (ordinal numeral), H (proper noun), C (comparative adjective) and U (superlative adjective).

Each word in a sentence was replaced with its word type. For example, three sentences of the right to privacy of § 26 of the Estonian Constitution are interpreted as three sequences of word types (consisting respectively of 3, 9 and 43 types):

Y N Z.

P V S S Z J S S Z.

S Z A S J P S V V P S Z J S V D Z D S A S J S S Z S Z A S J O S S J S S Z S

S J S S Z.<sup>6</sup>

The consecutive pairs (previous word type, next word type) were counted to get a Markov frequency table (Table 2). There is one pronoun and verb pair at the beginning of the second sentence which is followed by a verb and noun pair. Thus, 1 is stored in the cells (P, V) and (V, S) of the table.

Different frequency tables were compared with the Mosteller's method (Mosteller 1968). The idea is to normalise different tables to a comparable form in which all row and column sums are equal (for example 100). We call such table representation *kernel*. A kernel can be calculated iteratively:

STEP 1. FOR every row divide each element in a row by the sum of row elements multiplied to 100.

STEP 2. FOR every column divide each element in a column by the sum of column elements multiplied to 100.

<sup>6</sup> Paragraph 26, in Estonian: "Igaühel on õigus perekonna- ja eraelu puutumatusesele. Riigiasutused, kohalikud omavalitsused ja nende ametiisikud ei tohi kellegi perekonna- ega eraellu sekkuda muidu, kui seaduses sätestatud juhtudel ja korras tervise, kõlbluse, avaliku korra või teiste inimeste õiguste ja vabaduste kaitseks, kuriteo tõkestamiseks või kurjategija tabamiseks."

STEP 3. IF all row and column sums are near enough to 100 THEN STOP ELSE GO TO STEP 1.

*Table 2. Frequency table for § 26 of the Estonian Constitution*

	V	S	A	P	D	K	J	N	I	Y	Z	X	G	O	sum=
V	1	1	0	1	1	0	0	0	0	0	0	0	0	0	4
S	2	7	1	0	0	0	5	0	0	0	8	0	0	0	23
A	0	3	0	0	0	0	0	0	0	0	0	0	0	0	3
P	1	2	0	0	0	0	0	0	0	0	0	0	0	0	3
D	0	1	0	0	0	0	0	0	0	0	1	0	0	0	2
K	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J	0	5	0	1	0	0	0	0	0	0	0	0	0	1	7
N	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
I	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Y	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
Z	0	2	2	0	1	0	2	0	0	0	0	0	0	0	7
X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
G	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
sum=	4	22	3	2	2	0	7	1	0	0	10	0	0	1	

The process normally converges to a kernel table. For example, Table 4 is the kernel of Table 3. Rows and columns that only consist of zeros or that have only one positive value are deleted. Since rows and columns use the same set of labels, the respective column is deleted together with a row and vice versa as shown in Table 5. In case two documents need to be compared, the same row and column from the other table is deleted. Now two tables have the same labels but different row and column sums. To compare the numbers in two tables, we transform the distribution of numbers in the kernel table of the second table into the cells of the first table preserving the original row and column sums.

This is done iteratively similarly to the calculation of the kernel table. We use the chi-square  $\sum \frac{(initial - predicted)^2}{predicted}$  statistics to compare the initial table with the resultant prediction table by cells.

Table 3. Mosteller's frequency table

	B	-B	sum=
A	100	10	110
-A	5	2	7
sum=	105	12	117

Table 4. Mosteller's kernel table

	B	-B	sum=
A	67	33	100
-A	33	67	100
sum=	100	100	

Table 5. Filtered table for § 26 of the Estonian Constitution

	V	S	P	D	J	Z	sum=
V	1	1	1	1	0	0	4
S	2	7	0	0	5	8	22
P	1	2	0	0	0	0	3
D	0	1	0	0	0	1	2
J	0	5	1	0	0	0	6
Z	0	2	0	1	2	0	5
sum=	4	18	2	2	7	9	

Table 6. Original table for Penal Law

	V	S	A	P	D	K	J	N	I	Y	Z	X	G	O	H	C	U	sum
V	209	1248	1045	123	134	23	93	127	0	1	138	0	0	14	12	9	1	3177
S	850	3310	1676	93	177	1702	1666	275	1	217	4025	0	0	19	50	2	3	14066
A	29	4508	594	33	49	24	193	15	0	0	233	0	0	11	5	10	1	5705
P	790	801	151	22	37	56	12	0	0	0	16	0	0	1	0	1	0	1887
D	106	312	269	600	51	20	32	34	0	1	82	0	0	16	6	4	1	1534
K	86	291	553	68	33	14	51	133	0	7	914	0	0	11	13	15	0	2189
J	112	910	444	326	87	254	2	227	0	8	15	0	0	15	9	23	0	2432
N	4	355	251	0	0	9	148	25	0	1	7445	0	0	0	0	0	0	8238
I	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Y	0	12	0	0	0	0	0	731	0	628	493	0	0	263	0	0	0	2127
Z	992	1988	556	605	879	100	228	5704	0	1135	607	0	0	24	32	13	0	12863
X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
G	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O	0	114	0	0	0	0	1	185	0	0	78	0	0	0	0	0	0	378
H	0	120	16	0	4	4	6	10	0	12	7	0	0	0	20	0	0	199
C	0	25	2	0	50	0	0	0	0	0	1	0	0	0	0	0	0	78
U	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6
sum	3179	14000	5557	1870	1501	2206	2432	7466	1	2010	14054	0	0	374	147	77	6	

The initial table for the Penal Law is given in Table 6. In the first row of the table we can see that in 1,248 times a verb is followed by a noun. The total number of verbs in Penal Law is more than 3,000. From the second row we can see that in 3,310 cases a noun is followed by a noun and in 1,666 cases by a conjunctive. We can see also that in 4,508 cases an adjective is followed by a noun, in 790 cases a pronoun is followed by a verb and in 910 cases a conjunctive is followed by a noun. Numerals appear in the text 8,238 times but verbal attendants and genitive attributes do not appear at all. So we must filter out the corresponding rows and columns. But as we are comparing two texts the filtering also depends on the table of the other text. The initial table for Constitution FR1 is given in Table 7.

*Table 7. Original table for Constitution FR1*

	V	S	A	P	D	K	J	N	I	Y	Z	X	G	O	H	C	sum=
V	29	58	21	18	8	1	6	0	0	0	9	0	1	1	8	2	162
S	50	52	13	7	13	18	49	0	0	0	80	0	0	1	2	1	286
A	4	37	4	2	2	0	7	0	0	0	9	0	0	1	2	0	68
P	45	33	5	2	2	5	1	0	0	0	5	0	0	0	1	0	99
D	11	5	5	5	5	0	2	0	0	0	1	0	0	0	4	0	38
K	3	3	2	0	1	0	1	0	0	0	13	0	1	0	0	0	24
J	14	32	9	12	1	0	0	0	0	0	2	0	0	0	0	0	70
N	0	0	0	0	0	0	0	0	0	0	48	0	0	0	0	0	48
I	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Y	0	0	0	0	0	0	0	48	0	0	0	0	0	0	0	0	48
Z	5	22	5	10	4	0	4	0	0	0	0	0	0	0	0	0	50
X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
G	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
O	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
H	1	16	3	0	1	0	0	0	0	0	2	0	0	0	0	0	23
C	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
sum=	162	266	67	56	37	24	70	48	0	0	169	0	2	3	17	3	

In the first row of the Table, we can see that in Constitution FR1 a verb is 29 times followed by a verb. In the Penal Law, the respective number was 209 and the most dominant was the verb and noun pair with the frequency 1,248.

Next, the filtered tables of the Penal Law and Constitution FR1 are normalised. This is done by iterative formulas described above. Normalisation of the Penal Law table gives kernel Table 8.

Table 8. Kernel table for Penal Law

	V	S	A	P	D	K	J	Z	H	sum=
V	9	12	27	7	17	3	13	6	5	100
S	5	4	5	1	3	30	29	21	3	100
A	1	39	14	2	6	3	24	9	2	100
P	55	12	6	2	8	12	3	1	0	100
D	7	4	10	51	9	4	6	5	4	100
K	5	3	17	5	5	2	9	47	7	100
J	5	9	12	20	12	37	0	1	4	100
Z	14	6	5	12	36	4	10	8	5	100
H	0	9	3	0	4	4	7	2	70	100
sum=	100	100	100	100	100	100	100	100	100	

As the sum of elements in each row and column is always 100, each element of the table can be handled as a percentage. The percentages in the kernel table for Penal Law are different from the percentages in the initial table, as some percentages have grown and some have decreased.

To calculate the prediction table for the filtered Penal Law table with respect to Constitution FR1 table we use the percentages in the Constitution FR1 kernel table to calculate the distribution of frequencies in the Penal Law table using the initial row sums. After that operation, all the row sums in the Penal Law table are correct but the column sums can be different from the initial ones. We change the numbers in columns proportionally to obtain the initial column sums. We repeat this operation alternately for rows and columns until all the row and column sums are the same as in the initial filtered Penal Law table. The process is similar to the calculation of kernel tables and usually it converges quickly. The prediction table for the filtered Penal Law table with respect to the Constitution FR1 kernel table is given in Table 9.

We are interested in the differences between the filtered Penal Law table and the calculated prediction table. The sum of elements in the first row in both tables is 3,025. At the beginning of the verb line in the prediction table there are frequencies 289 and 1,293, whereas in the filtered Penal Law table there are frequencies 209 and 1,248. The tables do not coincide ideally. The similarity of the two tables can be estimated by using the chi-square test. It seems that calculated chi-square values are not comparable as the values are affected by the size of the legal acts. To improve the comparison of the legal acts, the elements in the tables of longer acts can be changed proportionally so that the sums of

Table 9. Prediction table for Penal Law with respect to Constitution FR1

	V	S	A	P	D	K	J	Z	H	sum=
V	289	1293	769	241	121	41	69	166	36	3025
S	1294	3012	1237	243	510	1932	1463	3834	23	13549
A	171	3532	627	115	129	0	344	711	39	5668
P	486	798	199	29	33	224	12	100	5	1885
D	267	272	447	163	184	0	56	45	44	1478
K	139	310	340	0	70	0	53	1112	0	2023
J	216	1104	510	248	23	0	0	57	0	2159
Z	309	3048	1138	831	375	0	286	0	0	5987
H	3	119	37	0	5	0	0	12	0	177
sum=	3174	13488	5304	1870	1451	2197	2283	6037	147	

elements in every table are equal to the sum of elements of the smallest table. In the current case we use the number of 1,043 words in the text of FR1 as the sum. Otherwise tables with bigger elements give bigger chi-square similarity values. After such reduction the chi-square similarity value of the filtered Penal Law table and its Constitution FR1 prediction table is 210, which is higher than the reference value 84 for 9x9 tables. So these tables are statistically different. As shown in Table 10, FR2 provides better results in relation to the Penal Code. The results provided in Table 11 show that FR1 and FR2 are different except that both levels equally predict legal acts *lex specialis* and do not predict case law. In other words, the results are deemed positive if the number in a cell is smaller than the number in the parentheses next to it, as encircled numbers 75(84) and 119(124) in Table 11 and negative if the number in a cell is bigger than the number in the parentheses next to it, as encircled numbers 210(84) and 207(124) in Table 10.

Table 10. Comparison of legal acts by the size of the smallest act using chi-square test (critical thresholds of chi-square for zero hypothesis are in the brackets)

	VÕS	KarS	Constitution	FR1	FR2
VÕS*		230 (173)	164 (148)	193 (84)	131 (124)
KarS**	300 (173)		531 (148)	210 (84)	207 (124)
Constitution	247 (148)	249 (148)		92 (84)	99 (124)
FR1	630 (84)	352 (84)	104 (84)		128 (84)
FR2	940 (124)	347 (124)	87 (124)	136 (84)	

\* VÕS – Law of Obligations Act

\*\* KarS – Penal Law

Table 11. Comparison table incl. legal acts *lex specialis* and judgements

	Constitution	FR1	FR2
<i>EraKS*</i>	595 (148)	75 (84)	119 (124)
<i>RavS**</i>	382 (148)	83 (84)	113 (124)
CJ 3-4-1-26-14	396 (148)	305 (84)	244 (124)
CJ 3-4-1-2-13	335 (148)	308 (84)	465(124)
Constitution		92 (84)	99 (124)
FR1	104 (84)		128 (84)
FR2	87 (124)	136 (84)	

\* *EraKS* – Private Schools Act

\*\* *RavS* – Medicinal Products Act

## 6. Conclusion

The assumption was that FR1 and FR2 may render different results when the grammatical similarity is measured in relation to selected legal acts. Both texts of fundamental rights provide relatively equal results as shown in Tables 10 and 11, with minor exceptions. Interestingly, FR1 and FR2 predict *lex specialis* norms (*EraKS*, *RavS*), but not case law that has been adjudicated based on these norms. This may be because the grammar of case law is substantially different from legal acts that are subject to strict composition rules. Wider selection of legal acts has to be tested, including the secondary legislation.

Based on these findings, the authors have construed a hypothesis that most restrictions to fundamental rights that trigger disputes are contained in the layer of special laws and secondary legislation, but this claim needs to be verified in future research, because the initial results represent the experimental stage of the research and are largely inconclusive.

The ability to identify patterns of norms across a legal system has been researched before. For instance, Täks and others (2015) have experimented with a different quantitative methodology with the vision to offer an easy-to-search graphical representation of imperative norms. Research and development of automated solutions for legal topics date back to 1972, when the computer programme TAXMAN was written (McCarty, 1977). It helped to resolve corporate tax

issues in the reorganisation proceedings. Ravellaw<sup>7</sup> and Ross<sup>8</sup> are two state-of-the-art projects, which demonstrate possibilities of legal data retrieval and analytics as well as the application of IBM Watson artificial intelligence in legal services. Formal methods often help to identify inconsistencies or mistakes, revealing interesting and useful results (e.g., Lohk, 2015).

The results of the current research may prove useful in the fields of computational linguistics and structural analyses or data mining of legal texts. The authors acknowledge that further research is needed in order to determine the extent to which the methodology used in the current article can be fully utilised.

The research can be scaled by sampling this method on other European constitutions. Although constitutions are structurally similar, it cannot be assumed that grammatical similarity analyses with different layers of legal acts and case law in other jurisdictions provide the same results as in the Estonian example.

**Addi Rull** works as a lecturer at Tallinn Law School of the Tallinn University of Technology. He has graduated from the Stockholm University with LLM in Law and IT and LLM in European Intellectual Property Law. He is pursuing PhD in informatics at the Tallinn University of Technology. His main research interests are systematisation of legal norms, software agents, smart contracts, artificial intelligence and law.

**Tõnu Tamme** works as a lecturer in theoretical informatics at the University of Tartu. His main research interests are information management, statistical language models and mathematical logic.

**Leo Võhandu** is professor emeritus in Informatics at the Tallinn University of Technology. He defended his PhD in 1955 in applied mathematics at the University of Tartu. He has supervised 46 doctoral dissertations in mathematics, biology, medicine, economics, informatics and linguistics. Now he is mainly consulting in the field of structural study of complicated empirical systems.

---

<sup>7</sup> At [www.ravellaw.com](http://www.ravellaw.com)

<sup>8</sup> At [www.rossintelligence.com](http://www.rossintelligence.com)

## References

- Alexy, R.** (2001), 'Põhiõigused Eesti põhiseaduses,' *Juridica*, pp. 31–48.
- Balkin, J. M.** (2008), 'Abortion and Original Meaning,' in *Constitutional Commentary*, vol. 24, no. 29, pp. 291–352.
- Commented edition (2002), *Eesti Vabariigi Põhiseadus. Kommenteeritud väljaanne* (Republic of Estonia Constitution. Commented Edition), Tallinn: Juura, Õigusteabe AS.
- Constitution of the Republic of Estonia (1992), *Eesti Vabariigi Põhiseadus*, adopted on 28 June 1992, RT 1992, 26, 349; RT I 2003, 29, 174.
- David, R. & Brierley, J. E. C.** (1993), *Major Legal Systems in the World Today: Introduction to the Comparative Study of Law*, 3<sup>rd</sup> ed., Toronto: Carswell Legal Publications.
- European Convention... (1950), European Convention for the Protection of Human Rights and Fundamental Freedoms, Rome, 4.11.1950.
- Ewald, W.** (2008), 'James Wilson and the Drafting of the Constitution,' *Journal of Constitutional Law*, vol. 10, no. 5 (June), pp. 901–1009.
- GitHub (2016), *Estnltk*, Open source tools for Estonian language processing. Retrieved from <https://github.com/estnltk/estnltk> [accessed 27 Jul 2016]
- Hickman, T.** (2007), 'Proportionality: Comparative Law Lessons,' *Judicial Review*, vol. 12, pp. 31–55.
- Karstens, B.; Koolen, M.; Dari-Mattiacci, G.; Bod, R. & Ginsburg, T.** (2014), *Reference Structures of National Constitutions*, NaiL2014: 2<sup>nd</sup> International Workshop on Network Analyses in Law, 10th December, Krakow, Poland.
- Kask, O.** (2012), *Eesti Vabariigi Põhiseadus. Kommenteeritud väljaanne*, II peatükk (Republic of Estonia Constitution: Commented Edition, Ch. 2). Retrieved from <http://www.pohiseadus.ee/ptk-2/> [accessed 29 Sep 2015]
- Kerge, K.** (2010), 'Muutustest õigusdiskursuses ja –keeles,' *Õiguskeel*, no. 2010/1, pp. 1–6.
- Law of Obligations Act (2001), *Võlaõigusseadus*, adopted on 26 September 2001, RT I 2001, 81, 487; RT I, 23.03.2015, 74.
- Lohk, A.** (2015), *A System of Test Patterns to Check and Validate the Semantic Hierarchies of WordNet-type Dictionaries*, PhD thesis, supervised by L. Võhandu, Tallinn University of Technology.
- Martin, M. & Horne, A.** (2008), 'Proportionality: Principles and Pitfalls – Some Lessons from Germany,' *Judicial Review*, vol. 13, pp. 169–179.
- Maruste, R.** (1997), *Põhiseadus ja selle järelevalve*, Tallinn: Juura Õigusteabe AS.
- McCarty, L. T.** (1977), 'Reflections on TAXMAN: An Experiment in Artificial Intelligence and Legal Reasoning,' *Harvard Law Review*, vol. 90, no. 5 (March),

- pp. 837–893. <http://dx.doi.org/10.2307/1340132>
- Medicinal Products Act (2004), *Ravimiseadus*, adopted on 16 December 2004, RT I 2005, 2, 4; RT I 01.09.2015, 27.
- Moore, B.** (1984), *Privacy: Studies in Social and Cultural History*, Armonk, NY: M. E. Sharpe, Inc.
- Mosteller, F.** (1968), ‘Association and Estimation in Contingency Tables,’ *Journal of the American Statistical Association*, vol. 63, no. 321, pp. 1–28.
- Navarro, P. E. & Rodriguez, J. L.** (2015) ‘Entailed Norms and the Systematization of Law,’ in M. Araszkievicz & P. Pleszka (eds.) *Logic in the Theory and Practice of Lawmaking*, New York: Springer, pp. 97–114. [http://dx.doi.org/10.1007/978-3-319-19575-9\\_3](http://dx.doi.org/10.1007/978-3-319-19575-9_3)
- Nõlv, E.** (2016), *xLaw* [Website]. Retrieved from <http://xlaw.eu> [accessed 29 Jul 2016]
- Penal Code (2001), *Karistusseadustik*, adopted on 6 June 2001, RT I 2001, 61, 364; RT I, 23.12.2014, 16.
- Private Schools Act (1998), *Erakooliseadus*, adopted on 3 June 1998, RT I 1998, 57, 859; RT I 23.03.2015, 255.
- Reinsalu, R.** (2012), ‘Lepingukeele keerukus,’ *Õiguskeel*, no. 2012/2.
- Ruiz, B. R.** (1997), *Privacy in Telecommunications: A European and an American Approach*, The Hague, etc.: Kluwer Law International.
- Singh, R.** (2008), ‘Using Positive Obligations in Enforcing Convention Rights,’ *Judicial Review*, vol. 13, no. 1, pp. 94–100.
- Smits, J. M.** (2015), *Do Small Jurisdictions Have a More Complex Law? A Numerical Experiment in Constitutional and Private Law*, Maastricht European Private Law Institute Working Paper No. 2015/05.
- State of Emergency Act (1996), adopted on 10 January 1996, RT I 2009, 39, 262; 01.09.2015, 6), § 4 (4). § 31 (1), 5).
- Täks, E.; Rull, A.; Säär, A. & Schäfer, B.** (2015), ‘Creating CoReO, the Computer Assisted Copyright Reform Observatory,’ in M. Araszkievicz & P. Pleszka (eds.) *Logic in the Theory and Practice of Lawmaking*, New York: Springer, pp. 479–511. [http://dx.doi.org/10.1007/978-3-319-19575-9\\_18](http://dx.doi.org/10.1007/978-3-319-19575-9_18)
- Tamme, T.; Vohandu, L. & Täks, E.** (2014), ‘A Method to Compare the Complexity of Legal Acts,’ in R. Winkels & N. Lettieri (eds.) 2nd International Workshop “Network Analysis in Law” NAI2014 in conjunction with JURIX 2014: The 27th International Conference on Legal Knowledge and Information Systems. Krakow, Poland, Dec 10, pp. 105–114.
- Uibopuu, H.-J.** (2000), *Inimõiguste Rahvusvaheline Kaitse*, Salzburg & Tallinn: Juura.
- Waltl, B. & Matthes, F.** (2014), ‘Towards Measures of Complexity: Applying Structural and Linguistic Metrics to German Laws,’ in R. Hoekstra (ed.) *Legal Knowledge and Information Systems*, Frontiers in Artificial Intelligence and Applications series, vol. 271, JURIX 2014: The 27th Annual Conference, IOS Press, pp. 153–162.