THE LEGAL LANGUAGE AS A SEMIOTIC SYSTEM

Nicoleta Baghici,

Lecturer,
(Alecu Russo State University of Bălți, Republic of Moldova)

Abstract: *The article focuses on the principles of developing an annotated corpus for processing documents of legal discourse. Studying the content of the concept in the speakers' mind should be made based on the totality of heterogeneous media, describing it in language. On the other hand, it is necessary to note that not necessarily that all conceptual information may become linguistic expression, some of it may be stored in memory in the form of other mental representations, of non-linguistic type - in the form of images, photos diagrams, etc. The concept of Constitution in American linguistic consciousness has a special meaning, it covers practically all spheres of American society. According to the language direction this article approaches in the main study, the concept of Constitution can be characterized, for example, as an axiological one, because it is linked indissolubly, from individual basic values of a society, together with concepts such as democracy, freedom, etc., that have left a mark in the history of state formation.*

Keywords: *USA Constitution, Frame-Net, PaLinKa, frame, semantic frame, semantic role, semantic element, core, non-core, frame element.*

Rezumat: *În articol, prezentăm principiile de întocmire a unui corpus de adnotare a documentelor juridice care pun, în prim plan, diferite concepte. Cercetarea conținutului unui concept poate fi făcută în baza mediumului heterogenetic, a limbajului. Totodată, nu tot conținutul conceptului capătă exprimare glotică. O parte din acesta poate fi stocat în memorie și întrebuințat în alte reprezentări mentale, exteriorizate prin imagini, scheme, diagrame etc. În imaginarul glotic american, conceptul de 'constituție' are o semnificație aparte, deoarece vine în contact cu toate sferele vieții sociale. Conceptul dat este unul axiologic, în relație cu valorile sociale de bază, dar și cu alte concepte ca 'democrație', 'libertate' etc., care vorbesc despre faptul că statul ca formațiune are o istorie.*

Cuvinte-cheie: *constituția Statelor Unite, Frame-Net, PaLinKa, frame, frame semantic, rol semantic, element semantic, nucleu, periferie, frame element.*

Introduction

At the current stage of science development and practice of legal discourse, linguistics concerns various ways of conceptualizing the reality of truth, in particular, through semantics. In the present study, we share the views of the scientist E. Kubryakova [23, p. 555] saying that "we know about the structure of consciousness only through language that allows these structures to report and describe them in any natural language [...]. The great majority of the information needed about the world, primarily scientific and theoretical, we perceive not during our feeling activities, objectives, practice, how important they would be, but created in the course of mediated language [ibidem]. E. Coseriu's vision, though anti positive in its essence, is based on full understanding of linguistics as a science of

culture in relation to the essential universals of language (semantics, alterity, creativity) [7, p. 73]. This scientist says that speaking is a general human activity, accomplished individually by technical representatives of linguistic tradition. "Language is the common ground of language speakers' historicity, and all that is said, is said in a language which, in part, is manifested in concrete form, in speech [...] what is said to be less effective than what is expressed and understood" [*ibidem*].

The problem of coding and understanding of language units in general and concepts, in particular, lies in conceptualizing and developing mental models of reality reflected in language in general and in specific linguistic consciousness of native speakers in particular. Elaboration of research methods depends on the legal discourse, terminology researchers' ideas about the structure and interrelationships between concepts. Thus, although the differences in the definition and comprehension of concepts are observed, linguists agree with the idea that a concept has a complex structure with different degrees of difficulty.

According to E. Wuster concepts exist independently from terms and have an extra linguistic character. The concept is composed of a number of common features specific to any object class. These characters are unfolded in their turn, in other concepts that may be used to structure mentally a specialized field in order to communicate in this field. We share the view that the starting point for all activities and terminological study should be the concept having a key role in analyzing semantic relations between character: "The process of collecting of terms cannot begin from the shapes but from concepts" [4, p. 124].

The concepts are distinct structural verbal units reflecting different "pieces of reality" [24, p. 245]. Thus, for example, the typology of cognitive concepts includes: mental images, diagrams, frames, scripts, perspectives, kaleidoscopic concepts logically constructed and differ exactly through the structuring manner, representation and updating. This method of concept detecting is based on the linguistic approach, which involves determining the type of the concept based on the analysis of the dictionary definitions.

The concept is, according to the majority of the scientists, a mental representation of reality. A prime example is the method of conceptual analysis developed by Y. Stepanov [32, p. 824], which is based on studying the evolution of the content of the concept starting from the analysis of "the inner form" and ending with a modern description of the content of the concept. Thus the concept is based on any three "layers": basic, present characteristic additional or more "passive" characteristics, but which are already outdated, "historical" as well as an internal form, usually unconscious, trapped in a verbal exterior [7, p. 73].

The triple structure of the concept is also characteristic for the lingvoculturological concept [2, p. 52]. The cultural concept includes form components, concept and value. In this aspect, the signified elements of the concept are designation, description and showing the structure of definition, which is made up of real information about real or imaginary object, but the meaning of the formative concept is reduced to a consistent generalization in memory associated with a particular object, phenomenon, event, quality etc. which supports the concept of linguistic consciousness. The concept reflects valuable separate linguistic identity and the whole team of lingvoculturologists [2, p. 54]. In addressing the "classical" methodological research the study is based on verbal means of objectification, such as "the best access to description and determining the nature of the concept is the language" [32, p. 824].

In the communication process the updating means are linguistic signs because namely the word obtains the status of concept, namely acts as a linguistic sign, transmitting the content of the concept properly. The concept usually is more than a lexical unit, being represented by lexemes, phraseological combinations, phrases, sentences, and collections of texts [21, p. 390], [24, p. 245]. In addition, "the more diverse the potential symbolic expression of the concept, the older is this concept and greater is its significance of value in this linguistic community" [33, p. 313]. The study of concept content should be carried out under the totality of heterogeneous media, describing it in language. It is not necessarily that all conceptual information become linguistic expressions, some may be kept as mental representations, of the non-linguistic type - as images, photographs, diagrams etc. [32, p. 824]. For these reasons, some concepts may have a direct mental projection in mental representations, while other concepts: tactile, gustatory and olfactory sensations, have no direct linguistic projection which is a simple combination of words (e.g. "the smell of freshly baked bread") [2, p. 54].

1. Background

Continuing the investigation, in order to describe the structural subsystems of legal language as a separate judiciary semiotic system, we will refer to the analysis of frames and conceptual oppositions of legal discourse. In this study, they are considered as methods of linguistic classification phenomena and events that occur in the objective reality of legal discourse [30, p. 97]. The conceptual category is a symbolic illustration of the category of the real world (or possible world) [5, p. 145], and every act of classification of real events through language is reflected in the relevant language units [1]. Therefore, the analysis of linguistic categorization allows us to understand according to what criteria the decisions in this case are

issued - as judges systematize and rationalize phenomena and objects in their reality as a result of their professional activities.

Along with the concept, the frame is one of key-categories of cognitive linguistics. There are two main views on frame in specialized literature - frame as type of concept and as a way to represent the information about the concept. The specificity of legal text, as any text, must be sought at the function level of evocation and for understanding texts/discourses, really helpful is, for example, exploring frames involved in enunciation, in other words, the "circumstances in which they speak".

Frames mean the relationships the linguistic sign establish in a text by a semantic function of evocation which aims to sign relations with other signs, signs in other texts dealing with the relations between signs and "things", relations between signs and "knowing things and frames" [7, p. 200], [2, p. 52].

M. Minsky, one of the founders of Artificial Intelligence Laboratory at the Massachusetts Institute of Technology (MIT), in his paper A Framework for Representing Knowledge, capitalizes the notion of frame, to show that the human mind structures the flow of deeds and impressions called reality in the form of "sequences" (chunks) of thought, language, memory and perception. "These "sequences" are, actually, "micro-worlds" intimately connected with constellations knots and connections" [29, p. 112].

Therefore M. Minsky considers, "a frame is a data structure which represents a stereotypical situation, such as that of being in the living room or take part at the birthday party of a child. Each frame attaches several types of information. Some of this information is about how to use the frame. One part is about what's going to happen. Another is about what to do if these expectations are deceived" [*ibidem*].

Any frame - explains this scientist - is composed of two levels, superior (of the top) and from the bottom (basic). The higher levels of frame are fixed and represent the aspects that are always true about a supposed situation, while the inferior levels are more terminals, defined as slots of introducing different categories of data. Each terminal corresponds to a category of individual data by certain marks. First, frames are dynamic organization of data. Any frame is associated with several types of information that is working and is ordered hierarchically. Unlike the higher levels that are stable and easy to recognize, the lower levels are variable and may be able to be shared by several frames. Active relations between the constituent parts of a frame are determined by degrees of tolerance and compatibility from lower levels to higher levels. Assembling and changing the frames are fundamental operations that allow interpretation of something unknown through something known. Secondly, the frames are organized systematically and emphasize their adaptability and projective nature of human intelligence. Therefore, "a frame is a set of questions that appear in a hypothetical situation: it provides possible problems and the methods used to solve them" [29, p. 114]. M. Minsky's considerations were given

continuity in Ch. J. Fillmore's scientific works (see [11], [12], [13], [14], [15], [16], [17]). Operating with the notion of frame, Ch. J. Fillmore transposes it in the center of its own theories about the construction and understanding natural language articulated facts.

The frames contribute to the extension of the meaning of discourse. "Frames necessarily interfere in any speech activity, for there is no discourse that does not occur in a particular circumstance, that does not have a background [...]. Frames participate in almost constantly in determining the signs and often substitute verbal determinators. But their functionality is much broader: the frames direct any discourse, giving it a sense, and can even determine the truth of the statements [7, p. 73]. In terms of frame typology more taxonomies from the theory of contexts are known [7], [18], [25, p. 120].

Frames, necessarily, interfere in any speech activity (there is no legal discourse that occurs in a particular circumstance) – directing any discourse by giving it a meaning and determining even the level the truth statements. Due to their recognized importance , "it's weird how little attention has been paid to them descriptively and analytically" [8]. "So far, it is the most complex theory of the contexts" says E. Coseriu distinguishing a wide range of frames which he groups into four types: state, region, context and discourse universe.

Through situation "we must understand something much more limited and less ambiguous than is commonly understood, i.e. only spatio-temporal circumstances and relationships that are created automatically by the very fact that someone speaks (with someone about something) at a point in space and in a moment in time" [8].

Paraphrasing E. Coseriu, we could say it is strange that, such a theory of frames, has not been valued until now and also the determinations, so precisely and rigorously formulated, are not "captured" in theories and in further linguistic methods. As for determinations, C. Vilcu opines for their "conjugation" with aspects of the generative theory [33]. Many facts - in grammatical constructions, as well as in the use of vocabulary – are generally facts of speech related to this linguistics of speech.

The first sketch of integral linguistics is based on the necessary threefold of language: the language in general and/or speech (universal level), language (historical level) and discourse/text (particular level).

Linguistics of speech, linguistics of language and linguistics of discourse or text - justify their tripartite basis and the fact that three different types of content (description, significance and meaning) correspond to them.

Ch. J. Fillmore's contribution to the present linguistic development is indisputable. The theory of semantic frames is, according to its father's considerations [11, p. 123-131], [17, p. 20-32], [12, p. 175-199], [13, p.111-137], [6, p. 373-400], the result of a research program covering "an empirical semantics" as opposed to formal semantics, whose purpose is to emphasize the fact that the description of grammar and the vocabulary of the language

must be supplemented by "description of cognitive frames" and interaction through which the language user interprets its environment, builds the messages and understands the messages of others or develops their interior and creates a model of his/her world. "[12, p.175-199], [14, p.222-254]. In other words, the theory of semantic frames is a model of interpretation of the meanings of the words in relation to conceptual schemes generated through contextualizing knowledge of experience and social interaction available to the speaker.

FrameNet is an impressive database application that "highlights all the semantic and syntactic combinatory possibilities of each word, for each of the meanings, through a computerized system of annotations"[3]. In this database, are so far, more than 10,000 English lexical units, out of 6,000 which are fully annotated and organized in nearly 800 related hierarchical semantic frames and illustrated with over 135,000 sentences. Any semantic frame (SF) can be described by a set of properties, called semantic elements (SE), which can be core and non-core. Thus, for example the semantic frame of the verb ESTABLISH is disclosed through a definition Created_entity and Creator as core elements and other non-core semantic elements which vary from case to case, from verb to verb:

establish.v

Frame Element	Core Type
Components	Extra-Thematic
Co_participant	Extra-Thematic
Created_entity	Core
Creator	Core
Depictive	Extra-Thematic
Instrument	Peripheral
Manner	Peripheral
Means	Peripheral
Place	Peripheral
Purpose	Peripheral
Role	Extra-Thematic
Time	Peripheral

Figure 1: *Example of semantic frame*

The presence or the absence of core and non-core elements reflect the dynamic of distribution of the units that make up the statement. The degree of participation of elements setting up the semantic frame differs from statement to statement. As exemplified by the passage from the Constitution of the United States, where we see the verb ESTABLISH:

Figure 2: *Examples of the verb ESTABLISH from USA Constitution*

Another way of representing the relationship between the constituents of a sentence is to include in square brackets the lexical units which illustrate various elements of the frame, adding the role fulfilled: [Creator WE THE PEOPLE OF THE UNITED STATES], [Purpose IN ORDER TO FORM A MORE PERFECT UNION], [Verb ESTABLISH] [created _entity JUSTICE]..., [Verb ESTABLISH] [created_entity THIS CONSTITUTION FOR THE UNITED STATES OF AMERICA].

The semantic richness and the significance of many lexical units is illustrated by the network of links between frames. For example, the English verb. to establish [to institute (as a law) permanently by enactment or agreement [19], [rom.to determine, establish, to found, to install (in a position , to turn (religion) inst. leg, to prove, confirm (a fact), is indexed on FrameNet as a single frame: ESTABLISH.v (Intentionally_create) Finished_Initial, as a verb and as a noun it has two meanings (Businesses and Intentionally_create) and Created - as a statute the lexical unit LU Establish (2015).

The verb ESTABLISH in the network of FrameNet, expresses its semantic-syntactic identity by attracting actants [20, p. 43-54], which are realized in the form of some arguments (semantic complements). From the perspective of this semantics, the verb's (ESTABLISH) valences determine the association the core roles (Creator and Created_entity with syntactic functions which can be analyzed on the surface of the statement:

(1) subject [Creator WE THE PEOPLE OF THE UNITED STATES], object [Purpose IN ORDER TO FORM A MORE PERFECT UNION], [Verb ESTABLISH] [created _entity JUSTICE]..., [Verb ESTABLISH] [created _entity   THIS CONSTITUTION FOR THE UNITED STATES OF AMERICA];

(2) ... Verb To ESTABLISH direct object [Creator AN UNIFORM RULE OF NATURALIZATION, and direct object [created _entity UNIFORM LAWS ON THE SUBJECT OF BANKRUTPCIES THROUGHOUT THE

UNITED STATES...Verb TO ESTABLISH direct object [created _entity POSTOFFICES AND POSTROADS...;

(3) ... Creator COURTS [indirect object] AS THE CONGRESS may from time to time ... Verb ESTABLISH [object].

The semantic conception of Ch. J. Fillmore was adopted and nuanced by us using the model of FrameNet, adapting the main text of the Constitution to its basic model. In Romania attempts to build a similar database as FrameNet are recent. One of the first steps has resulted in trying to translate , the statements from FrameNet from English into Romanian [27].

For Moldova this project is new. In Romanian language, the assumed significance of Ch. J. Fillmore's scientific vision is minor. The conception of this great contemporary linguist was much more welcomed and valued by Romanian specialists in computational linguistics. The literature analysis allowed the identification of basic characteristics of linguistic frame, which makes this study, the most appropriate method of describing the structure of specialized discourse.

The main factor of choosing the theme of judicial decisions, was the universality, guaranteed by a unique terminosystem of understanding law for nonprofessionals in the field. The frequency of litigation occurred on this issue, in accordance with certain modifications was taken into account. Each of these sections of constitutional law is governed by the main document - the appropriate change of the Constitution, which provides the opportunity to explore without special legal training. Detailing of the regulatory of each civil liberty by a specific amendment is guaranteed, independently of others, so that each can be assigned a unique system.

2. Main Focus of the article: Issues, Controversies, Problems

Information creates challenges for many areas, including terminology. The exponential growth of the number of specialized documents with new terms is beyond human cognitive ability. A possible solution for this problem is to employ automatic or semi-automatic proceedings to allow individuals and/or small groups to build qualitative terminology. Personalized tools for annotating discourse techniques turned out to be quite safe as they provide the processing of terminology. These techniques consider terms as independent lexical units that meet certain criteria, when the terms are integral parts of a coherent system.

Annotating legal discourse phenomenon is a very difficult task which cannot be achieved without appropriate tools for annotation. Major issues are temporal resources, financial and human to ensure an (almost) perfect corpus. What happens if we have short, limited human resources and material possibilities? A good solution is to use existing linguistic resources, built with considerable effort for a given language, and importing them into a new language. In this article, we will militate for this idea by providing and building the resources of the Romanian semantic role (for Moldova) starting with the frames defined in FrameNet.

The lexicographic project FrameNet contains detailed information about predicational English words (verbs, nouns adjectives, etc.). The basic unit is the semantic frame a structure such as a scenario of inferences, bound by linguistic convention of meanings of lexical units, defined as an event type or status A framework is a set of semantic elements SE, semantic roles SR and a set of lexical units that participate in their updating LU [16, p. 613].

A lexical unit represents a word for which combinatorial properties are applied. Semantic elements constitute the target word valences and can be therefore mandatory for the achievement of the lexical-semantic verb. Semantic elements are classified in core elements, non-core elements and peripherals. Core SE usually corresponds to direct arguments of a verb and ensures direct semantic correctness of the statement, while non-core SE the verb's (ESTABLISH) modifiers represent completing the statement with additional information. Below are examples of core Semantic elements:

Semantic Roles (SR);
Core Semantic Element (CSE): Agent, Evaluee, Reason;
Peripheral Semantic Element (PSE): Depicted Degree, Tool, Manner, Means;
Place, Purpose, Reason Action, Result, Time;
Lexical units (LU);
Noun: punishment_act (Rewards and Punishments) - punishment_penalty (Rewards_and_Punishments) verb: PUNISH
[Each House] AGENT… [punish] VERB [its members] EVALUEE [for disorderly Behavior] REASON…
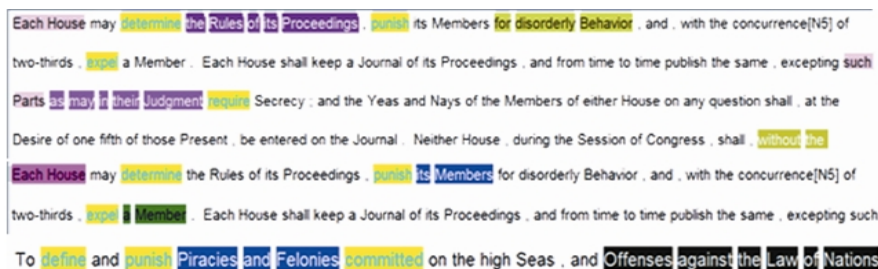[punish] VERB [Piracies and Felonies] EVALUEE…:

Figure 3: *Annotated exemples of the verb PUNISH from USA Constit*ution

These examples include the verb punish which has the following semantic roles: <agent>, <evaluator>, <reason> or <verb>, <evaluator>. The goal is to enrich RoFrameNet resource (the Romanian FrameNet [27] by creating a FrameNet for Moldova that would include semantic frames of verbs in legal discourse. In FrameNet the relations are established between the frames, not words. Therefore, lexical relations such as antonymy and synonymy are not taken into account. In the case of complex frameworks such as Criminal_process, every sequence of events or states is described as a single frame, linked to the complex frame through the relations of other sub-frame and other subframes through previous relations.

Frame CRIMINAL_PROCESS is divided into four sub-frames temporally successful: ARREST, ARRAIGNMENT, TRIAL, and SENTENCING. Frame "ARRAIGNMENT" (is divided into three subframes: NOTIFICATION_OF_ CHARGES, ENTER-ING_A_PLEA and BAIL_ DECISION. Frame TRIAL has also three sub-frames: COURT_EXAMINATION, JURY_ DELIBERATION and VERDICT. Figure 4 and Figure 5 represent the CRIMINAL_PROCESS frame and relationships between frames:

Figure 4: *Frame CRIMINAL_PROCESS and its relati*ons

Figure 5: *Frame CRIMINAL_PROCESS and its relations (USA)*

The criminal process in Moldova represents different steps in comparison with the American criminal process. First, according to Moldova Criminal Procedure Code, there are different procedures for judging an accused charged with murder. The jury's procedure is based on American criminal proceedings.

In Moldovan court, the jury has not been implemented since the Criminal Procedure Code does not provide for this procedure, however it is necessary to emphasize its importance to implement the adversarial principle in criminal proceedings. We believe that in the nearest future Moldova will establish the jury as a characteristic tool typical for a classical adversarial criminal trial. Unlike the American criminal process, Moldovan criminal proceedings are not initiated in Moldova when the suspect is arrested, but only when the prosecutor starts criminal proceedings, sets in motion the criminal proceedings, sends to trial, the prosecution argues in front of court, sums up a conclusion of the sentencing exercise the appealing procedures etc.

Figure 6: *Crime Scenario*

The ramifications from the criminal process - setting in motion the criminal proceedings, the sending to court, supporting the prosecution before court, drawing conclusions of conviction, setting up a recourse action etc. and following the judge's decision.
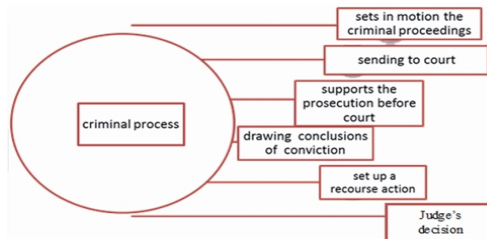

Figure 7: *CRIMINAL_PROCESS in Moldova*

The criminal trial frame for Moldova is still developing. Considering annotation problems with the legal corpus and the tags created for the American system, we have decided to create a database of lexical semantic and the tags to provide legal corpus (apprehension, arrest, search, detention of goods). Each of these frames can have well-defined sub-branches.

The methodology used for developing the Moldovan CRIMINAL_PROCESS frame differs from the method used by FrameNet. Thus, while FrameNet methodology is based on linguistic achievement, the development of the Moldovan CRIMINAL_PROCESS is based on identifying major legal concepts and the stages of criminal proceedings. The methodologies applied at the creation of the frames could explain the difference of granularity of the CRIMINAL_PROCESS within FrameNet and the CRIMINAL_PROCESS frame in Moldova.

The frame NOTIFICATION_OF_CHARGES perfectly illustrates the differences between the two legal systems. Within FrameNet, NOTIFICATION_OF_CHARGES is the first step to build a more general frame, called the arraignment. The arraignment session is a step in the American criminal law that does not exist in the Moldovan criminal process and is followed by ENTERING_A_PLEA frame, then BAIL_DECISION.

We conclude that using semantic FrameNet for social fields, including law, requires special consideration, and perhaps some adjustments. Although Romanian lexical units have equivalents in English, the legal scenario evoked by the English lexical unit is different from the legal

scenario evoked by lexical units in Moldova. The evoked scenarios are changing because legal systems are changing too.

FrameNet was the basis for automatic extraction of syntactico-semantic frames. These frames visualize the connection between meaning and syntactic structure by which it is shown. As a basis for semantic marking - the semantic frames are used - conceptual structures that represent events, objects and properties. Each frame includes a set of elements (frame elements). Each frame is assigned with a number of words-lexicon units that evoke the meaning of the given frame.

The legal terms can be mapped from annotated corpora. A "semantic annotation" presents a more accurate description of the knowledge contained in the text and its legal semantics. A semantic annotation should be well defined, easily understood by experts in the field and not be ambiguous. To comply with these requirements, a semantic annotation must rely on a formal model of the domain.

The texts which are annotated with appropriate legal terms in the legal field, the concepts that describe this area will improve the process of extracting information from texts and documents with legal content for the Romanian language. The texts will contribute to more qualitative results by disambiguation automatic translations of legal terms.

We plead for the idea that corpora should be constantly updated, the correctness of data, is one of many prerequisites that can elucidate the persistent problems in common language: examples from the press: "This is an area where, by acts of commission and omission on the part of... and laity over many decades, we have, I confess, acted shamefully" [22, p. 18]. "Cautionary tales of Baba Yaga acted to deter children from wandering... It acts as a place for refugees to flee from war, a source of income for..." [31]. ACT (V) - to carry out an action/to appear or seem to be/[ROLE]/ [PERFORMER]:

, declaring what Officer shall then act as President .

such Officer shall act accordingly ,

Figure 8: *Example of the verb "act"*

ABRIDGE (V) - to reduce the length of (a written work) by condensing or rewriting/to curtail; diminish [AGENT]/[ATTRIBUTE]/[CAUSE]/ [DIFFERENCE]/ [ITEM]:

any law which shall abridge the privileges or immunities of citizens of the United States ;

citizens of the United States , or in any way abridged ,

Figure 9: *Example of the verb "abridge"*

Using in common language examples from the press: "Chris Matthews: Republicans raped 26th Amendment And "Abridge" Youth Voting Rights In

NC [26]. We should be aware that polysemantism of the terms may be another instigation to build legal corpus. Legal Terms in the Romanian language in criminal law, criminal procedure, can evoke different scenarios, and different words can evoke the same scenario.

The word accuse has some lexical units (JUDGMENT_ COMMUNICATION)(NOTIFICATION_OF_CHARGES)(JUDGMENT) - as a more general word than lexical unit charge. While accuse (to accuse) evokes the CRIMINAL_INVESTIGATION, lexical unit accuse (to charge) evokes the CHARGING. Same with the terms abridge or act.

Legal and penal areas include many examples of polysemantic words. The word to testify, which means to give evidence, evokes the CRIMININAL_INVESTIGATION frame, the PROBATORY_HEARING frame, and the COURT_EXAMINATION frame. The word to evidence evokes the same frames.

Although the verbs to testify and to give evidence can be considered as synonyms in certain contexts, they present a variation of meaning. The word to testify is generally linked to the witness of a crime when he/she decides to testify voluntarily, while the word to give evidence is used in contexts where the person is required by an authority to witness.

The difference of meaning could be shaped in a legal lexicon based on the creation of two different frames: WITNESS AND ACCUSED. The frame WITNESS will be evoked by the verb to testify and the element ACCUSED will be evoked by the verb to evidence.

The annotator should analyze the context in which the lexical unit is introduced in order to choose the appropriate frame. In this election the annotator must use his/her intuition about language. Generally, the combination of words in a sentence helps to identify the significance of lexical units.

Relying on the analysis of terminological resources for the nomination of mental model elements of reality in the legal aspect, we can make the following conclusion: within different frames modeled in legal discourse, there are invariant features and components for each frame within which there are enough differences and personalized traits. Indeed, as it has also been noted by researchers, the type of such mental representation frame is determined by the specific coded representation.

4. Solutions and Recommendations

Annotation of legal discourse cannot be accomplished without the aid of specific tools. In case study we proceed to analyze the concept of USA Constitution using software PaLinKa. Recently, the need to produce reusable corpora led to an increasing use of XML coding in annotation.

As a result, annotation cannot be applied using simple text editors. In addition, annotating discourse is usually complicated requiring specialized

tools. In this section we will present the most important features of a speech annotation tool.

PaLinKa, the instrument presented in this paper meets all these requirements and is suitable for annotating legal discourse. The Constitution document was divided into nine XML files that open with another program Notepad ++ is quite simple to install on a PC:

```
[MARKER]
NAME:CRIME
BGCOLOR:10,17,2
FGCOLOR:248,248,250
ATTR:ID=# ;defines an unique id for each tag
PREFIX:ACT
PUT_IN_TREE:1
HOT_KEY:F6

[MARKER]
NAME:MESSAGE
BGCOLOR:243,109,132
FGCOLOR:248,248,250
ATTR:ID=# ;defines an unique id for each tag
PREFIX:ACT
PUT_IN_TREE:1
HOT_KEY:F6

[MARKER]
NAME:ACT
BGCOLOR:163,243,109
FGCOLOR:8,17,2
ATTR:ID=# ;defines an unique id for each tag
PREFIX:ACT
PUT_IN_TREE:1
HOT_KEY:F6
```

Figure 10: *Part of the preferences file used for annotation*

These are some preferences that were used for reference annotation words. NAME: - the tag name can be chosen by the annotator, depending on the SR (semantic roles and roles it wishes to highlight, ACT, CRIME, MESSAGE. As it can be seen, the main program does not display XML tags so that the text can be easily read. In order to identify the tags presented in the text, we specified a base color to show the text to be annotated and marked explicitly in this case, they were displayed and labeled four files of tags that are not repeated (over 140 tags) nor by name, neither by their chromatic color.
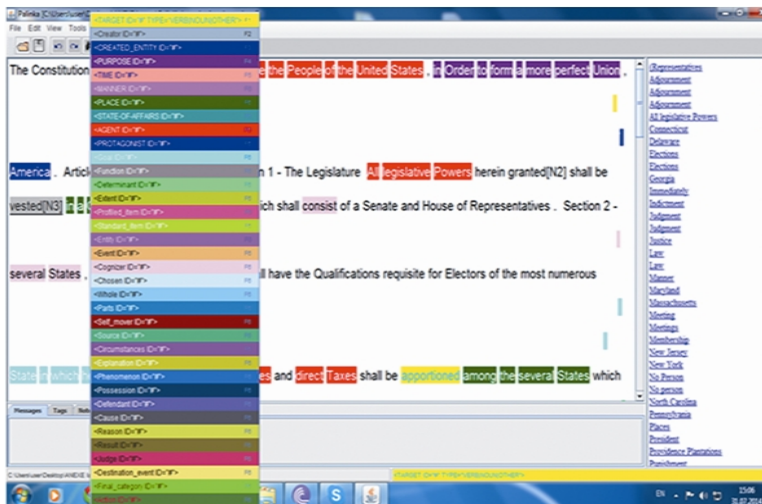
Coreferential annotation is notoriously a time consuming task and work. We marked manually the coreferent links between the entities from the text. Typically, each entity receives a unique ID and a link between the two entities that was scored using these IDs. These IDs are managed automatically by the program PaLinKa. Some links refer to more than one entity. This fact can also be codified.

The coreferential chains can be identified quickly using the entity tree on the right side of the screen (see Figure 10) or by highlighting them. Each frame contains several lexical units LU; some of these also include other LU. Due to this wealth of tags, one of the advantages of using the program PaLinKa is notorious that it hides XML tags using colors for each label. In addition, it is possible to adapt the program to mark the beginning and the end of each particular labeling using a new character. This feature has also proved to be useful for the annotation taken. It is possible to observe this in Figure 10, where each tag is marked by square brackets. Each tag (SR) has a special color because we examined only the verbs of the US Constitution, they have the yellow color . For instance, the LU CHOOSE (v) has the following SR COGNIZER/POSSIBILITIES/CHOSEN, and the frame is CHOOSING. Cognizers make decisions for the CHOSEN (be it an item or a course of action), from a set of POSSIBILITIES. COGNIZER may have an INTENDED_PURPOSE for CHOSEN:

Figure 12: *Examples of LU and SR marked in the text*

CHOSEN identifies the entity or the course of action, which is selected from POSSIBILITIES. CHOSEN chooses the COGNIZER from all POSSIBILITIES. COGNIZER makes a choice between a set of POSSIBILITIES. The POSSIBILITIES are usually expressed through oblique

phrases indicating the alternative or the alternatives, or a dependent clause (choosing of doing or choosing not to do so), usually driven by "or" or "if".

Schematically it will be represented so:

№	Subiect	Verb	CHOSEN	COGNIZER	POSSIBILITIES
1.	The Members	chosen	every second year	by the people of the several States	
2.		choose	their other Officers...a President pro tempore	The Senate	
3.		choose	their Speaker and other Officers	The House of Representatives	
4.		choose	a President	The House of Representatives	whenever the right of choice shall have devolved upon them
5.		choose	a Vice President	Senate	whenever the right of choice shall have devolved upon them

Table 1: *Syntactical-semantic relationship of legal frames*

The verb choose, as shown in the table above, requires compulsory CHOSEN and COGNIZER (core SE), but optional POSSIBILITIES (non-core SE). The verbs with related meanings such as pick, select or opt, have the same meaning, but in a different order. Each of these verbs indicates or refers to various aspects of the framework. The verb choose, focuses on the COGNIZER and CHOSEN having the POSSIBILITIES in the background or other explanations. The idea is that knowing the meaning of any of these verbs requires knowing what takes place in a legal environment, in this case knowing the US Constitution content, and knowing the meaning of any verb means that, in some sense, knowing the meaning of all the verbs. The knowledge and the frame structured experience provide background and motivation for the categories represented by words. The words that are linguistic material evoke the frame (in the mind of the speaker/listener); the interpreter (a speech or texts in which are the words) refer to the frame. A full description of these verbs must also include information on grammatical

properties and various syntactic patterns in which they occur. What elements or aspects of the frame may be realized as subject of the verb as object, if it exists, and it will be its form? What frame elements or its aspects can be seen as subject or object? Which of these items are optional and which are mandatory? For example in the sentence The Members chosen every second year by the people of several States, The Members - subject, verb chosen, every second year - CHOSEN, by the people of several States - COGNIZER. Defining words regarding frames and prototypes provide a useful approach to the limited issue for linguistic categories. In order to illustrate this approach, it is rather defined as a background frame than in terms referring to all unusual circumstances in which the word could be used. The fact that the verb choose may arise in contexts that do not fit the prototype suggests that the speakers are willing to extend the frame of the word to create a new one.

Another concept applied is the one of perspective. In the example The Senate choose ... their other Officers of the President pro tempor, evokes the frame Choosing, while mentioning all the frame, report similar to the legal environment, priority would be given to the one who chooses - COGNIZER. Similarly, the phrase their other Officers ... the President pro tempore was chosen by the Senate is the report of the legal environment from the perspective of the one who chooses (COGNIZER).

These are semantic roles, the frame of the verb to choose - Choosing. We come with combinatorial properties of a set of basic English vocabulary such as categorization of frames, in this case, the frames of the US Constitution.

4. Future Research Directions

According to the linguistic direction that we deal with in the main study, the concept of the Constitution can be characterized, for example, as an axiological one because it is linked, inseparably, to the individual core values of a society, along with such concepts as democracy freedom etc., that have left a mark in the history of state formation.

Analysing the definitions of the term Constitution from legal dictionaries revealed its constituent components and their concepts: The analysis is based on the logical and linguistic analysis of the definitions of the Constitution determining the meanings of terms. For example, the definitions allow us to compare the characteristics of the concepts that appear in definitions of related terms. In this case - in some definitions of a term, as a document, as a system of state governance, as a source of power, as power limitation according to the law, that document may be modified in accordance with the applicable procedures.

According to those stated above, there were identified three lexical-semantic paradigms of the concept Constitution, explaining its structure.

The US Constitution as a lexical-semantic paradigm outlined in texts through representative lexemes - (e.g. basic means of representation a cognitive concepts in language - text, definition, amendment, provision, rule, clause, article, title, to write/rewrite, to cite, to contain, Federal etc.). The last two paradigms are more closely related than the first one, the conceptual component approaches them - updating the powers of the Constitution as a special type of document within the society, due to its projection on society.

The metaphorical analysis of the concept Constitution demonstrated the use of metaphors on a large scale (from 149 cases of the use of the term Constitution in 250 thousand of uses, were found 38 cases, or 26% of total employment, of the metaphor) the vast majority are verbal metaphors representing one model - personification.

In the text of court decisions, the Constitution is conceptualized as having independence from other authorities reasonably equipped with power and being fair. This type of personification leads to consideration by the judges of the Constitution as a full member of the judicial process, which defines the powers of others, such as local and federal authorities. Lexical-semantic paradigms define the direction that will continue the conceptualization, developing and implementing the concept.

A relation between the constituents of a sentence was established, lexical units which illustrate various elements of the frame, with the addition of the fulfilled role. The semantic richness and the open significance of many lexical units are illustrated by the network of links between frames adapting the main text of Constitution to filimorian model.

FrameNet's use to annotate the corpus of the Romanian language, for example it requires a deeper thinking about the equivalence of the lexical units of the English and Romanian. The legal corpus faces a dual challenge: (1) the equivalence of the lexical units (2) the equivalence of the legal concepts. After identifying the lexical unit that evokes the frame to receive annotation, it is necessary to find an equivalent for that lexical unit in English and check which setting is evoked by the English lexical unit in FrameNet. At this stage of annotations, the annotator can refer to his/her polyglot knowledge, or use a bilingual dictionary. The second challenge of the annotator while using the tags in FrameNet is the unmatchability of frames in legal systems.

Research results focus to a great extent on reporting the novelty elements (consisting of words or new meanings) with reference to the annotation theory and the role of semantic frames as lexical units. The results can be used to supplement general explanatory dictionaries of the Romanian language and the development of specialized databases, as well as create special university courses for students, the foundation and completion of

undergraduate courses, suggesting new research directions, such as polysemy, synonymy and antonymy terminology, terminological metaphor, the establishment of legal terminology etc.

A recommendation would be to engage automatic or semi-automatic proceedings to allow individuals and/or small groups to effectively build high quality terminology of their own resources which closely reflect their individual goals and viewpoints. Nonetheless, theoretical and practical results of the research take on a special dimension in the accession process of the Republic of Moldova to the European Union.

Conclusions

The attempt of annotating specialized texts is relevant in terminology as a term may be used in various statements in which the speaker associates additional meanings. Identifying a certain sense in the case of a polysemantic term is achieved contextually, the selection being determined by the semantics of lexical units with which it is associated in communication. The movement of terms depends on their availability and possible opportunities for semantic extensions. Some terms depend on their degree of specialization and work only in the field where they were launched, while the others are moving to other areas of specialization. The low degree of specialization of terms allows their usual language by adding additional meanings. The annotation theory was developed and demonstrated in relatively simple texts, the purposes of using it in a wide range of legal documents, such as specialized dictionaries, lexicographical dictionaries, glossaries, vocabularies, lexicons. Terms processed through these filters result in a lexical-semantic national corpus for creating a specialized FrameNet. Addressing terminology, on what we based our research, imposes a new direction in science deducted from adopting the semantic frame theory from the computational perspective of the project FrameNet, analyzing the terminology of the American legal discourse and treating legal terminology as lexical units through semantic frames roles of PaLinKa software.

References
1. Adam Jean-Michel, Lingvistica textuală. Introducere în analiza textuală a discursurilor. Traducere de Corina Iftimia, prefață de Rodica Nagy. Iași: Institutul European, 2008.
2. Бабушкин А. П. Концепты разных типов в лексике и фразеологии и методика их выявления. В: Методологические проблемы когнитивной лингвистики. Волго¬град, 2001, с. 52-57.
3. Backer Collin F., Fillmore C., Lowe J. The Berkeley FrameNet Project. In: Proceedings of the COLING-ACL. Montreal, 1998.
4. Baghici N., Conceptul de strategie și tactică în discursul juridic. In: Studia Universitatis. Seria „Științe Umanistice". Chișinău, USM, 2012, №4 (54), p. 124-127.

5. Болдырев Н. Н., Панасенко Ж. А. Когнитивные аспекты функциональной поликате¬горизации английского глагола. В: Когнитивные аспекты языковой категоризации. Рязань, 2000, с. 145-150.

6. Cognitive Linguistics: Basic Readings. Mouton de Gruyter, Berlin, New York, pp. 373-400. https://arkitecturadellenguaje.files.wordpress.com/2013/11/cognitive-linguistics-basics-readings-dirkgeeraerts.pdf [visited 10.02.2016].

7. Coşeriu E. Omul şi limbajul său. Studii de filozofie a limbajului, teorie a limbii şi lingvistică generală, antologie. Argument şi note de Dorel Fînaru. Iaşi: Editura Universităţii „Alexandru Ioan Cuza", 2009, p. 73-221.

8. Coseriu E. Language Theory and General Linguistics. Five studies. Bucureşti: Editura Enciclopedică, 2004.

9. Establish. An Encyclopedia Britannica Company. http://www.merriam-webster.com/dictionary/establish?show=0&t=1406728245 [visited 10.02.2016].

10. Fillmore Ch. J. An Alternative to Checklist Theories of Meaning. In: Cogen C., Thompson H., Thurgood G., Whistler K., Wright J. (eds.) Proceedings of the First Annual Meeting of the Berkeley Linguistics Society. Berkeley, 1975, pp. 123-131.

11. Fillmore Ch. J. The Case for Case Reopened, Syntax and Semantics, Grammatical Relations. Vol. 8. Nr. 1, 1977. New York, San Francisco, London: Academic Press. Republished in: Fillmore Ch. J. Form and Meaning in Language. Vol. I: Papers on Semantic Roles. Stanford: CSLI Publications, 2003, pp. 175-199.

12. Fillmore Ch. J. Frame Semantics. In: The Linguistics Society of Korea (ed.), Linguistics in the Morning Calm. Seoul: Hanshin Publishing Company, 2006, pp. 111-137.

13. Fillmore Ch. J. Frames and the Semantics of Understanding. In: Quaderni di semantica. Vol. VI. Nr. 2, December, 1985, pp. 222-254.

14. Fillmore Ch. J. Form and Meaning in Language. In: Papers on Semantic Roles. Vol. I. Stanford: CSLI Publications, 2003, pp. 175-199.

15. Fillmore Ch. J. Frame Semantics. In: K. Brown (editor-in-chief). Encyclopedia of Language and Linguistics. Second edition. Oxford: Elsevier, 2006, pp. 613-620.

16. Fillmore Ch. J. Frame Semantics and the Nature of Language. In: Annals of the New York Academy of Sciences. Vol. 280: Origins and Evolution of Language and Speech. October, 1976, pp. 20-32.

17. FrameNet. https://framenet.icsi.berkeley.edu/fndrupal.

18. Guţu A. Tipologia raporturilor structurale în arborescenţele terminologiei tehnico-ştiinţifice franceze. În: Lecturi filologice. Nr. 1, 2006, p. 43-54.

19. Карасик В.И. Языковой круг: личность, концепты, дискурс. Москва: Гнозис, 2004, с. 390.

20. Кириленко Н. П. Функционирование научно-технических терминов в газетных текстах. Автореферат дис. ...канд. филол. наук. Москва, 1984, с. 18.

21. Кубрякова Е. С. Части речи в когнитивном аспекте. Москва, 1997, с. 435.

22.   Kubryakovan E. S., Demyankov V. 3., Pankratz Y. G., Luzin L.G. Brief Dictionary of Cognitive Terms. Moscow: Moscow State University Press, 1996. http://www.twirpx.com/file/383435/ [visited 10.02.2016].

23.   Maingueneau D. Analiza textelor de comunicare. Traducere de Mariana Şovea. Cuvânt înainte de Oana Popârda. Iaşi: Institutul European, 2007.

24.   Levinson Stephen C. Pragmatics. Cambridge: Cambridge University Press, 1989.

25.   Matthews Ch. Republicans Violate. 26th Amendment and "Abridge" Youth Voting Rights in NC. http://datab.us/2vxf7hZVi00#Chris%20Matthews:%20Republicans%20Violate%2026th%20Amendment%20And%20%22Abridge%22%20Youth%20Voting%20Rights%20In%20NC [visited 10.02.2016].

26.   META-Share.   http://metashare.infoiasi.ro/repository/search/?q=&selected_facets=languageNameFilter_exact%3Aenglish [visited in july 2014].

27.   Meyer I., Mackintosh K. When Terms Move into Our Everyday Lives: An Overview of De-terminologization. In: Terminology. Vol. 6. Nr. 1, 2000, pp. 111-138.

28.   Minsky M. A Framework for Representing Knowledge. In: Haugeland J. (ed.). Mind Design II. 2nd edition. Cambridge, Massachusetts, London: The MIT Press, 1997.

29.   Moscal D. Viziuni structurale asupra lexicului înainte de teoria câmpurilor lexicale elaborată de Jost Trier. In: Philologica Jassyensia. Anul VII. Nr. 1 (13), 2011, p. 97–105.

30.   Sama H. A Compelling and Haunting Document that Slips through any Attempt at Categorization. http://screen.artshub.com.au/news-article/reviews/film/sama-hugo-/the-vanquishing-of-witch-baba-yaga-245187 [visited 10.02.2016].

31.   Степанов Ю.С. Константы: Словарь русской культуры. Москва: Школа «Языки русской культуры», 1997. 824 с.

32.   Стернин И. А. Язык и национальное сознание. Воронеж, гос. ун-т. Филол. фак. Каф. общ. языкознания и стилистики, Межрегион, центр коммуникатив. исслед. Воронеж: Истоки, 2002. 313 с.

33.   Vîlcu, C.-D. Orizontul problematic al integralismuluiю Vol. 1: Integralism şi fenomenologie. Cluj-Napoca: Editura Argonaut & Scriptor, 2010. http://www.romling.unituebingen.de/energeia/zeitschrift/2013/pdf/Rezension_Vilcu.pdf.