

Linguistic Issues in Language Technology – LiLT
Submitted, January 2012

**Frame Semantics and legal corpora
annotation**

Theoretical and applied challenges

Anderson Bertoldi
Rove Chishman

Published by CSLI Publications

Frame Semantics and legal corpora annotation

Theoretical and applied challenges

ANDERSON BERTOLDI, *Universidade do Vale do Rio dos Sinos*

ROVE CHISHMAN, *Universidade do Vale do Rio dos Sinos*

1 Introduction

This paper presents a theoretical discussion about the use of Frame Semantics as corpora annotation paradigm. Frame Semantics (Fillmore, 1982, 1985) is a cognitive linguistics theory that explains the complex system of cognitive relations that a speaker must know in order to understand an utterance. Understanding that something was bought requires an understanding that something was sold. Understanding a commercial scenario requires the identification of the different participants of the commercial event, such as the buyer, the seller, and the good. Frame Semantics does not distinguish between linguistic knowledge and world knowledge. In other words, the understanding of the words to buy and to sell implies the cultural knowledge related to a commercial event: someone needs to have money to buy a good and a seller expects to receive money for a good.

FrameNet (Fillmore et al., 2003) is a lexical database that organizes the meaning of the words according to the Frame Semantics principles. FrameNet frames have been used for developing lexical databases and annotated corpora for different languages. Spanish FrameNet (Subirats, 2009), Japanese FrameNet (Ohara, 2009), and FrameNet

Brasil (Salomão, Maria Margarida M., 2009) are just some examples of FrameNets created for languages other than English. Kicktionary (Schmidt, 2009) is an example of a specialized multilingual frame-based lexicon for the soccer language. Boas (Boas, 2005) suggests the use of semantic frames as interlingual representation for multilingual lexicons. SALSA project (Burchardt et al., 2009) applies the FrameNet semantic tags to manually annotating a German corpus and automatically developing a frame-based lexicon of German. Other works have applied FrameNet frames to automatic development of lexicons using automatic transfer of corpus annotation in parallel corpora (Padó and Lapata, 2005), (Padó and Lapata, 2009), (Tonelli and Pianta, 2008) and for legal ontology improvement (Agnoloni et al., 2009).

The objective of this paper is to evaluate the applicability of Frame Semantics theory and FrameNet paradigm for the semantic annotation of legal texts. The work presented in this paper is an initial step in the construction of a treebank for the Brazilian legal language. A semantic annotated treebank of the Brazilian legal language could be used for machine training, especially to summarize legal decisions.

This semantic annotated corpus to be composed by legal decisions and laws is part of a larger project that researches how linguistic information could be used to improve legal information management and legal information retrieval in the Brazilian courts¹. This corpus is being planned to be representative of the entire legal production of Brazilian courts and Brazilian legislative houses, such as the laws published by the Brazilian Senate and the judicial decisions of the Federal Courts. These legal texts will compose a corpus of the Brazilian legal language and will be annotated with semantic frames. The semantic annotation of the corpus will follow the running text methodology. While the FrameNet lexicographic approach chooses sample sentences to be annotate, in the running text methodology all the sentences of the corpus should receive some kind of annotation. Therefore, one plans to annotate all the frame-evoking lexical units of the corpus and the major parts of the sentences that represent the frame elements of the frame-evoking lexical units.

In order to discuss the use of the Frame Semantics theory and FrameNet paradigm for corpus annotation, the remaining sections of this paper are organized as follows: Section 2 presents Frame Semantics theory. Section 3 presents the organization of the FrameNet lexical

¹This work was developed in the scope of the Semantic Technologies and Legal Information Retrieval Systems Project, supported by CAPES-CNJ (Brazil) under the rubric N°. 020/2010/CAPES/CNJ and coordinated by Prof. Dr. Rove Luiza de Oliveira Chishman.

database according to the principles of Frame Semantics. Section 4 discusses the use of Frame Semantics and FrameNet paradigm in the legal domain. It is in this section that the limits of FrameNet in multilingual applications in the legal domain will be approached. The limitations discussed in this paper are lexical equivalence, frame correspondence, frame element equivalence, and polysemy. In the Section 5, the future directions of this work are presented.

2 Frame Semantics

Frame Semantics is a linguistic theory founded by Fillmore (Fillmore, 1982, 1985) based on the concept of **frame** (Minsky, 1974, Goffman, 1974). In the beginning of the theory, Fillmore (Fillmore, 1975) makes a distinction between **scene** and **frame**. Scene is considered as a large concept including visual scene, interpersonal transactions, standard scenarios defined by culture, institutional structures, body image, human beliefs, actions, experiences or imaginings. On the other hand, frame is considered as a system of linguistic choices made by speakers including collection of words and choices of grammatical rules that can be associated with a scene.

Fillmore (Fillmore, 1977) associates Frame Semantics to Case Grammar (Fillmore, 1968) and relates semantic roles to scenes and frames in order to explain how speakers understand the meaning of the words. In his paper *The case for case*, Fillmore (Fillmore, 1968) suggests a case frame compound by semantic roles such as agentive, dative and instrumental. Frame case (Fillmore, 1968) motivated most of the studies on semantic roles. It is in this paper (Fillmore, 1977) that Fillmore first illustrates the relation between frame and semantic roles (called frame elements in FrameNet). Fillmore uses the commercial transaction scenario to explain that the verbs to buy, to sell and to cost represent different perspectives of the same event. For example, a seller gives a good in exchange for money and a buyer gives money in exchange for a good. An event such as a commercial transaction represents the change of possession of two goods: the money passes from the buyer to the seller and the good passes from the seller to the buyer. The analysis of the commercial transaction may be considered an attempt to describe frame elements for a semantic frame. These frame elements substitute the case frame proposal (Fillmore, 1968).

The initial distinction between scene as a cognitive structure and frame as a linguistic structure was abandoned in Fillmore's later works (Fillmore, 1982, 1985). According to Fillmore (Padó and Lapata, 2009), "by the term 'frame' I have in mind any system of concepts related in

such a way that to understand any one of them you have to understand the whole structure in which it fits (...)" (p.111). In Frame Semantics, words can 'evoke' an entire world knowledge that is stored in the speaker's mind through a cognitive structure call 'frame'. Fillmore (Fillmore, 1985) states that "a frame is evoked by the text if some linguistic form or pattern is conventionally associated with the frame in question" (p.232).

Fillmore and Atkins (Fillmore and Atkins, 1992) presented the first linguistic analysis based on Frame Semantics and made reference to a future on-line dictionary based on frames. Analyzing words that express risk, such as risk, danger and hazard, Fillmore and Atkins presented a proposal of eleven categories to describe the participants of the Risk frame: chance, harm, victim, valued object, risky situation, deed, actor, intended gain, purpose, beneficiary and motivation. These semantic roles are still not called frame elements.

3 FrameNet

FrameNet is a lexical database that describes word meanings based on the principles of Frame Semantics. According to Fillmore et al. 2003, "the central idea of Frame Semantics is that word meanings must be described in relation to semantic frames (...)" (p.235). Therefore, the frame and the lexical unit are the units of lexical analysis in FrameNet. A lexical unit is the combination of a word form with a meaning. Every new meaning of a word represents a new lexical unit. It is the lexical unit that evokes the frame, not the word. FrameNet does not describe relations between words and lexical relations, only between frames and conceptual relations.

In FrameNet, each lexical unit is related to a semantic frame, that is, each lexical unit evokes a frame. The semantic frame describes a certain situation and the participants of that situation that are likely to be mentioned in the sentences (Fillmore and Atkins, 1992). The valence properties of a lexical unit are expressed as the entities that can participate in the frame evoked by the lexical unit (Fillmore et al., 2003). The semantic roles played by these entities are called frame elements in FrameNet. SALSA project (Burchardt et al., 2009) uses frame elements to annotate corpora of German language. The methodology adopted in SALSA project is running text, that is, all the sentences in a corpus receive some annotation. In FrameNet, annotation is lexicographic, that is, only selected sample sentences are annotated with frame elements.

According to Fillmore and Baker (Fillmore, 2010), the method of lexical analysis in FrameNet follows five steps:

Characterizing the frames: A situation for which the language has provided lexical units is described, i.e., arresting a suspect.

Describing and naming frame elements: After characterizing a frame, the frame elements are named, i.e., Authorities, Charges, Offence, Suspect.

Selecting lexical units: the frame-evoking lexical units are identified, i.e., apprehend.v, apprehension.n, arrest.n, arrest.v, book.v, bust.n, bust.v, collar.v, cop.v, nab.v, summons.v.

Creating annotations of sample sentences: Sample sentences collected from BNC are annotated, i.e., Are [*you*_{AUTHORITIES}] arresting [*me*_{SUSPECT}] [for the murder of Topaz Brown?_{OFFENCE}].

Automatically generating lexical entries: Annotated sample sentences are converted to lexical entries. The lexical entries contain the definition of the lexical unit, the syntactic realizations of each frame element and the valence patterns.

FrameNet makes a differentiation between ‘core’ frame elements and ‘peripheral’ frame elements. In general, frame elements that are necessarily realized are ‘core’. ‘Peripheral’ frame elements represent more general information such as time, manner, place, and purpose.

4 Legal Corpora Annotation

This section presents the challenges to be faced when using not only the FrameNet semantic tags but also Frame Semantics as a general linguistic theory for legal corpora annotation. The first distinction that has to be made here is between automatic and manual semantic annotation. Padó and Lapata (Padó and Lapata, 2005) and ão Padó (Padó and Lapata, 2009) use a system of automatic transfer of FrameNet tags from English corpus to other languages corpora. The discussion presented in this paper shows that automatic approaches may not consider mismatches between frames. The observations made in this work take into account manual annotation of corpora, but these observations are useful even for automatic corpus annotation based on the FrameNet semantic tags.

4.1 Lexical Equivalence

Using FrameNet to annotate corpora of languages other than English requires thinking about the equivalence of the lexical units between languages. Legal corpora will face a double challenge: (i) the equivalence of lexical units and (ii) the equivalence of the legal concepts. After identifying the frame-evoking lexical unit to receive annotation, it is necessary to find an equivalent for that lexical unit in English and to

verify which frame is evoked by that English lexical unit in FrameNet. In this step of the annotation, the annotator may refer to his/her own knowledge of the languages in question, or use a bilingual dictionary.

Lexical equivalence is one of the critical points for legal corpora annotation because the legal events in Brazil and the USA are not always correspondent to each other. Considering the verb *acusar* in Portuguese, an annotator will very easily identify to accuse as an English equivalent for *acusar*. In the next step the annotator can go to the Framenet on-line database to search which semantic frame is evoked by the lexical unit to accuse. The annotator will see that this lexical unit evokes three different frames, Judgment_communication, Judgment, and Notification_of_charges. Analyzing the three frames, the annotator will perceive that only Notification_of_charges is related to the legal domain. That is a simple case only to illustrate the manual work of finding equivalents in English for lexical units of a Brazilian legal corpus. After identifying the English equivalent and the evoked frame, the annotator has to verify whether the legal event and the frame elements described by that frame are correspondent to the Brazilian legal event. That topic will be discussed in Sections 4.2 and 4.3.

When considering the verb *denunciar*, there is a more complex case. The annotator may be in doubt about the English equivalent for the verb *denunciar*. The most obvious English equivalent for *denunciar* is to denounce. Notwithstanding, when the annotator searches the FrameNet database for the lexical unit to denounce, he/she will see that FrameNet has only Judgment_communication as a frame evoked by the lexical unit to denounce. Since the frame Judgment_communication is not a frame that describes a legal event, the annotator will have to find another equivalent for to denounce and search the FrameNet database again. In this case, he/she can use a bilingual dictionary. The possible result is a lexical entry like in figure 1, extracted from (de Noronha Goyos Jr, 1992):

Portuguese	English
<i>denunciar</i>	to denounce; accuse; inform against; report; proclaim.

FIGURE 1 Bilingual legal dictionary entry for *denunciar*

In this case, the equivalents in English will make the work of the annotator more difficult. Firstly, the verb to denounce does not evoke a legal scenario in FrameNet. Secondly, the verb to accuse evokes a

legal frame, the Notification_of_charges frame, but the legal scenario described by the FrameNet Notification_of_charges frame does not fit to describe the legal scenario evoked by the Portuguese verb *denunciar*. Another possible equivalent for *denunciar* could be inform against, but FrameNet does not describe complex units extensively. If the annotator searches the FrameNet database for inform, the frames that he/she will find are not related to the legal domain: Reporting and Telling.

4.2 Frame Correspondence

The second challenge that an annotator will face while using the FrameNet tags is frame correspondence between legal systems. Returning to the preceding examples, the lexical units *acusar* and *denunciar* have the same equivalent in English, to accuse. In the FrameNet database, the verb to accuse evokes the frame Notification_of_charges, but verbs *acusar* and *denunciar* are polysemic and may evoke more than one frame in Brazilian Portuguese. Frame correspondence involves legal event correspondence and frame relations correspondence. In this point of the argument, it is necessary to describe the FrameNet organization of the frame Criminal_process to understand why frame relations correspondence is so difficult in legal domain.

The frame Notification_of_charges is part of a larger frame called Criminal_process. In the FrameNet terminology, Criminal_process frame is a non-lexical frame. The function of non-lexical frames is to connect semantically related frames. Non-lexical frames do not present frame-evoking lexical units. They represent complex events divided in more specific frames. Criminal_process frame describes the different steps of a criminal process according to the American legal system.

In FrameNet, relations are established between frames, not words. Therefore, lexical relations, such as antonymy and synonymy, are not considered. In case of complex frames, like Criminal_process, each sequence of events or states is described as a single frame, related to the complex frame through Subframe relations and to the other subframes through Precedes relations.

Criminal_process frame is divided into four subframes temporally succeeded: Arrest, Arraignment, Trial, and Sentencing. Arraignment frame is divided into three subframes: Notification_of_charges, Entering_a_plea, and Bail_decision. Trial frame also presents three subframes: Court_examination, Jury_deliberation and Verdict. Figure 2 shows Criminal_process frame and relations among frames.

The Brazilian criminal process presents steps different from the American criminal process. Firstly, according to the Brazilian Code of Criminal Procedures, there are different procedures to try a defendant

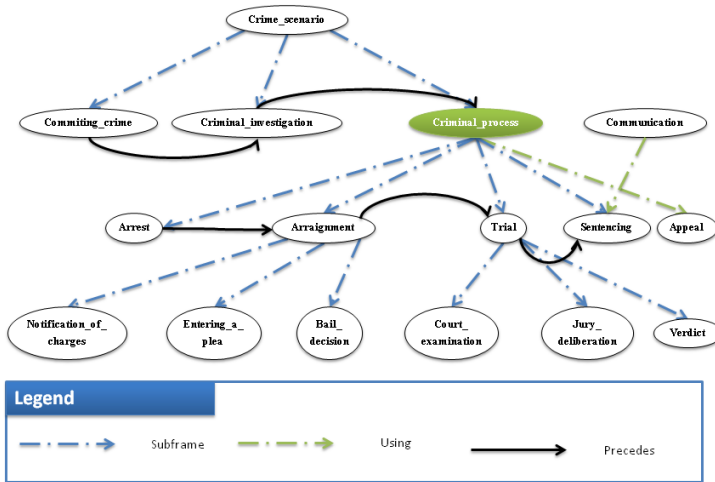


FIGURE 2 Criminal_process frame and frame relations

accused of a crime. Secondly, this study will present only one of these procedures, a special procedure called the jury. The special procedure of the jury is based on the American criminal process. In Brazil the special procedure of the jury is used only in cases of intentional crime against life, like murders.

Unlike the American criminal process, the Brazilian criminal process does not start when the suspect is arrested but when the prosecution charges the suspect. The judge issues a writ requesting the accused to present a (written) response against the accusation. In the writ, the accused is informed of the charges against him. This step of the Brazilian criminal process could be considered equivalent to the FrameNet Notification_of_charges and Entering_a_plea frames. The prosecution can present a (written) rebuttal to the defense of the accused. The judge can request further investigation to clarify any doubt and, only after all of these steps, the accused goes to the court for a preliminary hearing. After this preliminary hearing, if the accused is considered a defendant, he/she will face the jury.

A Brazilian Criminal_process frame is under development (Figure 3). Considering the problems of annotating a Brazilian legal corpus with

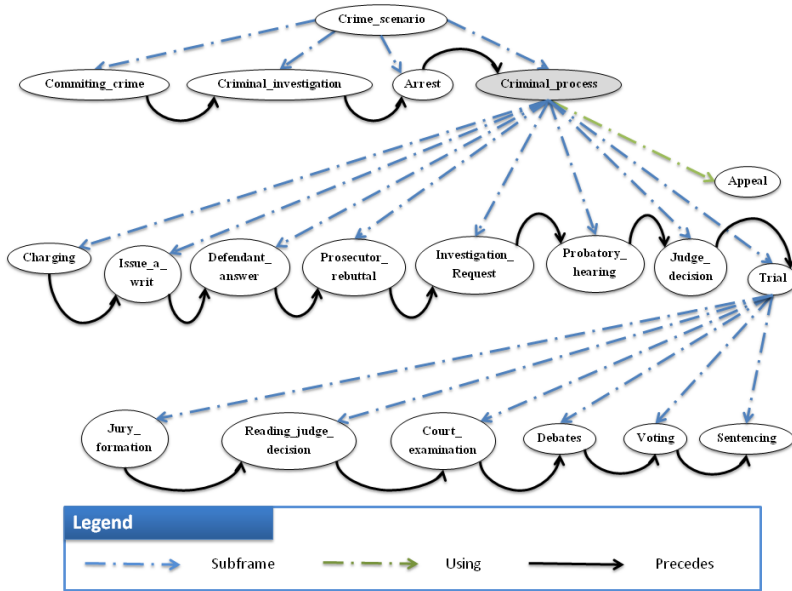


FIGURE 3 Brazilian Criminal_process frame and frame relations

the tags created for the American system, one decided to create a frame-based lexical database to provide semantic tags for legal corpora annotation of Brazilian legal documents. The Brazilian Criminal_process frame is divided into eight subframes: Charging, Issue_a_writ, Defendant_answer, Prosecutor_rebuttal, Investigation_request, Probatory_hearing, Judge_decision, and Trial. Trial frame is divided into six subframes: Jury_formation, Reading_judge_decision, Court_examination, Debates, Voting, and Sentencing.

The methodology used to develop the Brazilian Criminal_process frame differs from the methodology applied by FrameNet. While FrameNet methodology is grounded in the linguistic realization, the development of the Brazilian Criminal_process frame was grounded in the identification of the major legal concepts and steps of a criminal process. The methodologies applied to the creation of both frame could explain the difference of granularity of the FrameNet Criminal_process frame and the Brazilian Criminal_process frame. Another important observation is that FrameNet does not present a terminological orientation.

The lexical equivalence and legal event correspondence are essential

to use the FrameNet semantic tags. Although the lexical unit *acusar* has to accuse as an equivalent in English, the frame evoked by *acusar* is Charging in the Brazilian Criminal_process frame, while the lexical unit to accuse evokes the frame Notification_of_charges in FrameNet. Even though to accuse may be considered as translation equivalent of the Portuguese lexical units *acusar* and *denunciar*, the scenario evoked by the lexical unit in English is different of the scenario evoked by its translation equivalents *acusar* e *denunciar*.

Considering that an annotator finds in a dictionary to accuse as equivalent of *acusar*, he/she goes to the FrameNet lexical database and searches for that lexical unit. The annotator will have, as a result, the frames Judgment, Judgment_communication and Notification_of_charges. The annotator obviously chooses the frame Notification_of_charges because he/she is annotating a legal corpus. Is the frame Notification_of_charges appropriate to describe the scenario evoked by the lexical unit *acusar* in Portuguese?

The frame Notification_of_charges illustrates very well the differences between two legal systems. In FrameNet, Notification_of_charges is the first step of a more general frame, called Arraignment. The arraignment session is a step in the American criminal process that does not find a correspondence in the Brazilian criminal process. The second step of the frame Arraignment is Entering_a_plea and the third is Bail_decision. All of these three frames are related by Precedes relations. Therefore, if the annotator tags *acusar* as evoking the frame Notification_of_charges, he/she is enriching a Brazilian Portuguese text with (social-oriented) semantic information of another legal system. If the legal scenario changes, what say about the participants?

If the annotator uses the FrameNet Criminal_process frame to annotate Brazilian legal texts, he/she will not find the appropriate frame elements. As the frame Notification_of_charges is only one step of a larger frame, there is an entire inheritance of frame elements of the frame Arraignment that do not find correspondence in the Brazilian legal system. Even the relations between FrameNet frames will not be useful to represent the Brazilian scenario of a criminal process. The use of the FrameNet semantic tags in social-oriented areas, such as Law, requires a special attention, and maybe some adjustments. Although the Portuguese lexical units have equivalents in English, the legal scenario evoked by the English lexical unit is different from the legal scenario evoked by the Portuguese lexical unit. The scenarios evoked change because the legal systems change.

4.3 Frame Element Correspondence

Once a legal event does not present correspondence, the semantic frame, that is, the schematic representation of the legal event, will differ from one legal system to another. When the legal scenario changes, the participants of that event may also change. According to FrameNet, in the frame *Notification_of_charges*, “the judge or other court officer (the *Arraign_Authority*) informs the Accused of the Charges against him/her”.² The core frame elements in this frame are *Arraign_Authority*, *Accused*, and *Charges*.

Let’s consider that an annotator is using FrameNet tags to annotate a Brazilian legal corpus and he/she has to designate a semantic frame for each frame-evoking lexical unit and a frame element for each major part of the sentence. He/she finds in the corpus a sentence whose frame-evoking lexical unit is *acusar*. Then he/she goes to the FrameNet on-line database to search for the appropriate frame. Since the annotator knows that the English equivalent to the lexical unit *acusar* is to accuse, he/she searches for the verb to accuse in the FrameNet database and finds out that to accuse evokes the frame *Notification_of_charges*. The annotator will perceive immediately that the frame element *Arraign_Authority* does not properly fit the Brazilian criminal process scenario. What should the annotator do? Should he/she use the frame element *Arraign_Authority* in spite of the differences in the legal scenario or create a new frame to describe the Brazilian legal scenario evoked by the lexical unit *acusar*?

In FrameNet the lexical unit to accuse evokes part of a complex scenario of an arraignment. In the Brazilian *Criminal_process* frame, the lexical unit *acusar* evokes a non-complex scenario of Charging. In the Charging frame a prosecutor, the *Prosecution_Authority*, charges an Accused for an offense. The core frame elements of the Charging frame are *Prosecution_Authority*, *Accused* and *Charges*. The non-correspondence between the legal events in the American and Brazilian criminal procedures causes that mismatch of the frame elements. The annotator will have to decide between using the FrameNet tags and ignoring the mismatches or creating a new frame each time he/she perceives a non-correspondence of the legal events and a mismatch between the role of a participant in the American legal system and the role of a participant in the Brazilian legal system.

²FrameNet description for the frame *Notification_of_charges*.

4.4 Polysemy and synonymy in the legal domain

Polysemy and synonymy in legal corpora may be another challenge. Portuguese words in legal and criminal domains can evoke different scenarios and different words can evoke the same scenario. The word *acusar* has at least two lexical units related to the legal domain: *acusar* as a more general lexical unit (to accuse) and *acusar* as a more terminological lexical unit (to charge). While the lexical unit *acusar* (to accuse) evokes the Criminal_investigation frame, the lexical unit *acusar* (to charge) evokes the Charging frame. The same polysemous context occurs with the word *denunciar*. The lexical unit *denunciar* (to denounce) evokes the Criminal_investigation frame and the lexical unit *denunciar* (to charge) evokes the Charging frame. The words *acusar* and *denunciar* are synonymous in both contexts: Criminal_investigation and Charging.

Legal and criminal domains are full of examples of polysemous words. The word *depor* (to testify) evokes the Criminal_investigation frame, the Probatory_hearing frame, and the Court_examination frame. The word *testemunhar* (to testify) evokes the same frames the word *depor* evokes. Although the words *depor* and *testemunhar* may be considered as synonymous in some contexts, they present a variation of meaning. The word *testemunhar* is generally related to the witness of a crime when he/she decides to testify voluntarily, while the word *depor* is used in contexts when the person is required by an authority to testify. The meaning difference between *depor* and *testemunhar* could be modeled in a frame-based legal lexicon creating two different frame elements: WITNESS and ACCUSED. The frame element WITNESS will be evoked by the verb *testemunhar* and the frame element ACCUSED will be evoked by the verb *depor*.

The annotator has to analyze the context in which the lexical unit is inserted in order to choose the appropriated frame. In this choice the annotator needs to use his/her intuition about the language. Generally, the combination of words in a sentence helps in the identification of the meaning of the lexical units.

5 Final Remarks and Future Directions

This paper analyzed the use of the Frame Semantics paradigm and the FrameNet semantic tags for legal corpora annotation. The motivation for this study is the intention of creating a treebank of the Brazilian legal language to be used for natural language processing. The assumption of this work is that Frame Semantics provides a theoretical support for semantic annotation without ignoring the syntax of the sentences.

Therefore, the importance of Frame Semantics for corpora annotation stands in the possibility to map the syntactic and semantic valence of the lexical units at the same time. Using only one approach it is possible to enrich syntactic chunks with syntactic and semantic information.

The evaluation of the FrameNet tags to annotate legal corpora in Portuguese deserves some remarks. Firstly, the more complex the event described by a frame is, the more difficult it is to use for semantic annotation in another languages. This is the case of the frame *Criminal_process* and its subframes. Secondly, non-complex frames, like *Law* and *Legality*, tend to be more universal. Yet, how to face the lack of correspondence of semantic frames between two different legal systems and languages?

The difficulties observed in applying FrameNet semantic frames for legal corpus annotation in Portuguese lead to the following conclusions: First of all, FrameNet is not a terminological database. The second, differences of legal systems make difficult to apply to Brazilian Portuguese the semantic tags created for the English lexical units, and based in the American legal system. Such conclusions point to a methodology to be used in annotation projects that work with legal language. In order to annotate legal corpora it is firstly necessary to create the legal frames specific for each legal system. Some FrameNet frames may be very similar in different languages and can be used as a starting point to the development of a new frame-based legal lexicon. On the other hand, legal events that are very specific of a country have to be described as a new frame.

In order to create a treebank of the Brazilian legal language, compound by Brazilian laws and Brazilian court decisions, two methodological steps will be taken. The first is the development of a frame-based legal lexicon to provide frames and frame elements for semantic annotation of the Brazilian legal texts. The second is the annotation of legal corpora with the frames and frame elements especially created for the Brazilian legal system.

A frame-based lexicon of the Brazilian legal language is now under development. The motivation for such a lexical database is to provide an appropriate description of the knowledge about the Brazilian legal system and legal procedures. The assumption is that a lexical database developed especially for the Brazilian legal system may avoid some of the semantic annotation problems listed in this paper. Frame Semantics was the linguistic theory chosen for this enterprise for two reasons. The first, the theory permits to describe the roles of the participants of the predication and to make the mapping between syntax and semantics. The second, a frame-based lexicon of the Brazilian legal language could

be connected to other linguistic resources based on frames, such as the Brazilian FrameNet.

A frame-based lexicon for the Brazilian legal language will probably make the annotation easier when it avoids the problems of lexical equivalence, discussed in Section 4.1, frame correspondence, discussed in Section 4.2, and frame element correspondence, discussed in Section 4.3. Notwithstanding, the annotator will continue to face the challenge of the polysemy of the language.

Acknowledgments

The authors acknowledge the support offered by the Brazilian Federal Agencies CAPES and CNPq.

References

- Agnoloni, Tommaso, Meritxell Fernández Barrera, Maria-Teresa Sagri, Daniela Tiscornia, and Giulia Venturi. 2009. When a framenet-style knowledge description meets an ontological characterization of fundamental legal concepts. In P. Casanovas, U. Pagallo, G. Sartor, and G. Ajani, eds., *AICOL Workshops*, vol. 6237 of *Lecture Notes in Computer Science*, pages 93–112. Springer. ISBN 978-3-642-16523-8.
- Boas, Hans C. 2005. Semantic frames as interlingual representations for multilingual lexical databases. *International Journal of Lexicography*. 18(4):445–478.
- Burchardt, Aljoscha, Katrin Erk, Anette Frank, Andrea Kowalski, Sebastian Padó, and Manfred Pinkal. 2009. *Using FrameNet for the semantic analysis of German: annotation, representation, and automation.*, pages 209–244. Berlin/New York: Mouton de Gruyter.
- de Noronha Goyos Jr, Durval. 1992. *Noronha's Legal Dictionary – Noronha Dicionário Jurídico: English-Portuguese, Portuguese-English – Inglês-Português, Português-Inglês*. São paulo: Observador Legal, 1st edn.
- Fillmore, Baker C., Charles J. 2010. *A frames approach to semantic analysis.*, pages 313–339. Oxford: Oxford University Press.
- Fillmore, Charles J. 1968. *The case for case*. New York: Holt, Rinehart & Winston.
- Fillmore, Charles J. 1975. An alternative to checklist theories of meaning. In *Proceedings of the first annual meeting of the Berkeley Linguistics Society*, pages 123–131. Berkley: [16] Berkeley Linguistics Society.
- Fillmore, Charles J. 1977. *Scenes-and-frames semantics*. No. 59 in *Fundamental Studies in Computer Science*. North Holland Publishing.
- Fillmore, Charles J. 1982. *Frame semantics*, pages 111–137. Seoul, South Korea: Hanshin Publishing Co.
- Fillmore, Charles J. 1985. Frames and the semantics of understanding. *Quaderni di Semantica* 6(2):222–254.

- Fillmore, Charles J. and B. T. S. Atkins. 1992. *Towards a frame-based lexicon: The semantics of RISK and its neighbors*, pages 75–102. Hillsdale: Lawrence Erlbaum Associates.
- Fillmore, Charles J., Christopher R. Johnson, and Miriam R.L. Petruck. 2003. Background to FrameNet. *International Journal of Lexicography* 16.3:235–250.
- Goffman, Erving. 1974. *Frame Analysis*. New York: Harper.
- Minsky, Marvin. 1974. A framework for representing knowledge. Tech. rep., Cambridge, MA, USA.
- Ohara, Kyoko Hirose. 2009. *Frame-based contrastive lexical semantics in Japanese FrameNet: The case of risk and kakeru*, pages 163–182. Berlin/New York: Mouton de Gruyter.
- Padó, Sebastian and Mirella Lapata. 2005. Cross-linguistic projection of role-semantic information. In *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing, HLT '05*, pages 859–866. Stroudsburg, PA, USA: Association for Computational Linguistics.
- Padó, Sebastian and Mirella Lapata. 2009. Cross-lingual annotation projection of semantic roles. *J. Artif. Int. Res.* 36:307–340.
- Salomão, Maria Margarida M. 2009. Framenet brasil: um trabalho em progresso. *Caleidoscópio* 7(3):171–182.
- Schmidt, Thomas. 2009. *A multilingual lexical resource of football language*, pages 102–132. Berlin/New York: Mouton de Gruyter.
- Subirats, Carlos. 2009. *Spanish FrameNet: A frame-semantic analysis of the Spanish lexicon*, pages 135–162. Berlin/New York: Mouton de Gruyter.
- Tonelli, Sara and Emanuele Pianta. 2008. Frame information transfer from English to Italian. In B. M. J. M. J. O. S. P. D. T. Nicoletta Calzolari (Conference Chair), Khalid Choukri, ed., *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*. Marrakech, Morocco: European Language Resources Association (ELRA). ISBN 2-9517408-4-0. <http://www.lrec-conf.org/proceedings/lrec2008/>.