Constructing parallel lexicon fragments based on English FrameNet entries: Semantic and syntactic issues

Hans C. Boas
The University of Texas at Austin
Department of Germanic Studies and Department of Linguistics
1 University Station, C3300, Austin, TX 78712-0304, U.S.A.
E-mail: hcb@mail.utexas.edu

Abstract

This paper investigates how semantic frames from FrameNet can be re-used for constructing FrameNets for other languages. Section one provides a brief overview of Frame Semantics (Fillmore 1982). Section 2 introduces the main structuring principles of the Berkeley FrameNet project. Section three presents a typology of FrameNets for different languages, highlighting a number of important issues surrounding the universal applicability of semantic frames. Section four shows that while it is often possible to re-use semantic frames across languages in a principled way it is not always straightforward because of systematic syntactic differences in how lexical units express the semantics of frames. Section five summarizes the issues discussed in this paper.

Keywords: Computational Lexicography, FrameNet, Frame Semantics, Syntax

1. Frame Semantics

Research in Frame Semantics (Fillmore 1982, 1985) is empirical, cognitive, and ethnographic in nature. It seeks to describe and analyze what users of a language understand about what is communicated by their language (Fillmore & Baker 2010). Central to this line of research is the notion of *semantic frame*, which provides the basis for the organization of the lexicon, thereby linking individual word senses, relationships between the senses of polysemous words, and relationships among semantically related words. In this conception of the lexicon, there is a network of hierarchically organized and intersecting frames through which semantic relationships between collections of concepts are identified (Petruck et al. 2004). A frame is any system of concepts related in such a way that to understand any one concept it is necessary to understand the entire system; introducing any one concept results in all of them becoming available. In Frame Semantics, word meanings are thus characterized in terms of experience-based schematizations of the speaker's world, i.e. frames. It is held that understanding any element in a frame requires access to an understanding of the whole structure (Petruck & Boas 2003).[1] The following section shows how the concept of semantic frame has been used to structure the lexicon of English for the purpose of creating a lexical database.

2. The Berkeley FrameNet Project

The Berkeley FrameNet Project (Lowe et al. 1997, Baker et al. 1998, Fillmore et al. 2003a, Ruppenhofer et al. 2010) is building a lexical database that aims to provide, for a significant portion of the vocabulary of contemporary English, a body of semantically and syntactically annotated sentences from which reliable information can be reported on the valences or combinatorial possibilities of each item targeted for analysis (Fillmore & Baker 2001). The method of inquiry is to find groups of words whose frame structures can be described together, by virtue of their sharing common schematic backgrounds and patterns of expressions that can combine with them to form larger phrases or sentences. In the typical case, words that share a frame can be used in paraphrases of each other. The general purposes of the project are both to provide

[1] See Petruck (1996), Ziem (2008), and Fillmore & Baker (2010) on how different theories employ the notion of “frame.”

reliable descriptions of the syntactic and semantic combinatorial properties of each word in the lexicon, and to assemble information about alternative ways of expressing concepts in the same conceptual domain (Fillmore & Baker 2010).

To illustrate, consider the sentence *Joe stole the watch from Michael*. The verb *steal* is said to evoke the `Theft` frame, which is also evoked by a number of semantically related verbs such as *snatch*, *shoplift*, *pinch*, *filch*, and *thieve*, among others, as well as nouns such as *thief* and *stealer*. [2] The `Theft` frame represents a scenario with different Frame Elements (FEs) that can be regarded as instances of more general semantic roles such as AGENT, PATIENT, INSTRUMENT, etc. More precisely, the `Theft` frame describes situations in which a PERPETRATOR (the person or other agent that takes the GOODS away) takes GOODS (anything that can be taken away) that belong to a VICTIM (the person (or other sentient being or group) that owns the GOODS before they are taken away by the PERPETRATOR). Sometimes more specific information is given about the SOURCE (the initial location of the GOODS before they change location). [3] The necessary background information to interpret *steal* and other semantically related verbs as evoking the `Theft` frame also requires an understanding of illegal activities, property ownership, taking things, and a great deal more (see Boas 2005b, Bertoldi et al. 2010, Dux 2011).

Based on the frame concept, FrameNet researchers follow a lexical analysis process that typically consists of the following steps according to Fillmore & Baker (2010:321-322): (1) Characterizing the frames, i.e. the situation types for which the language has provided special expressive means; (2) Describing and naming the Frame Elements (FEs), i.e. the aspects and components of individual frames that are likely to be mentioned in the phrases and sentences that are instances of those frames; (3) Selecting lexical units (LUs) that belong to the frame, i.e. words from all parts of speech that evoke and depend on the conceptual background associated with the individual frames; (4) Creating annotations of sentences sampled from a very large corpus showing the ways in which individual LUs in the frame allow frame-relevant information to be linguistically presented; (5) Automatically generating lexical entries, and the valence descriptions contained in them, that summarize observations derivable from them (see also Atkins et al. 2003, Fillmore & Petruck 2003, Fillmore et al. 2003b, Ruppenhofer et al. 2010).

The results of this work-flow are stored in FrameNet (http://framenet.icsi.berkeley.edu), an online lexical database (Baker et al. 2003) currently containing information about more than 1,000 frames and more than 10,000 LUs.[4] Users can access FrameNet data in a variety of ways. The most prominent methods include searching for individual frames or specific LUs.

Valence Patterns:

These frame elements occur in the following syntactic patterns:

Number Annotated	Patterns				
1 TOTAL	Frequency	Goods	Perpetrator	Victim	
(1)	AVP Dep	NP Obj	NP Ext	PP[from] Dep	
1 TOTAL	Goods	Instrument	Time		
(1)	NP Obj	NP Ext	PP[after] Dep		
1 TOTAL	Goods	Manner	Perpetrator		
(1)	NP Obj	AVP Dep	NP Ext		
1 TOTAL	Goods	Means	Perpetrator	Place	Victim
(1)	NP Ext	PP[in] Dep	CNI --	PP[at] Dep	INI --
1 TOTAL	Goods	Means	Perpetrator	Source	
(1)	NP Obj	PPing[by] Dep	NP Ext	INI --	
1 TOTAL	Goods	Means	Perpetrator	Victim	
(1)	NP Obj	PPing[by] Dep	NP Ext	INI --	

Figure 1: Partial valence table for *steal.v* in the `Theft` frame

Each entry for a LU in FrameNet consists of the following parts: (1) A description of the frame together with definitions of the relevant FEs, annotated examples sentences illustrating the relevant FEs in context, and a list of other LUs evoking the same frame; (2) An

[2] Names of frames are in courier font. Names of Frame Elements (FEs) are in small caps font.

[3] Besides so-called core Frame Elements, there are also peripheral Frame Elements that describe more general aspects of a situation, such as MEANS (e.g. *by trickery*), TIME (e.g. t*wo days ago*), MANNER (e.g. *quietly*), or PLACE (e.g. *in the city*).

[4] For differences between FrameNet and other lexical databases such as WordNet see Boas (2005a/2005b/2009).

annotation report displaying all the annotated corpus sentences for a given LU; (3) A lexical entry report which summarizes the syntactic realization of the FEs and the valence patterns of the LU in two separate tables (see Fillmore et al. 2003b, Fillmore 2007).

Figure 1 above illustrates an excerpt from the valence patterns in the lexical report of *steal* in the `Theft` frame. The column on the far left lists the number of annotated example sentences (in the annotation report) illustrating the individual valence patterns. The rows represent so-called frame element configurations together with their syntactic realizations in terms of phrase type and grammatical function. For example, the third frame element configuration from the top lists the FEs GOODS, MANNER, and PERPETRATOR. The GOODS are realized syntactically as a NP Object, the MANNER as a dependent ADVP, and the PERPETRATOR as an external NP. Such systematic valence tables allow researchers to gain a better understanding of how the semantics of frames are realized syntactically.[5]

3. FrameNets for other languages

3.1. Similarities and differences

Following the success of the Berkeley FrameNet for English, a number of FrameNets for other languages were developed over the past ten years. Based on ideas outlined in Heid (1996), Fontenelle (1997), and Boas (2001/2002/2005a), researchers aimed to create parallel FrameNets by re-using frames constructed by the Berkeley FrameNet project for English. While FrameNets for other languages aim to re-use English FrameNet frames to the greatest extent possible, they differ in a number of important points from the original FrameNet (see Boas 2009).

For example, projects such as SALSA (Burchardt et al. 2009) aim to create full-text annotation of an entire German corpus instead of finding isolated corpus sentences to identify lexicographically relevant information as is the case with the Berkeley FrameNet and Spanish FrameNet (Subirats 2009). FrameNets for other languages also differ in what types of resources they use as data pools. That is, besides exploiting a monolingual corpus as is the case with Japanese FrameNet (Ohara 2009) or Hebrew FrameNet (Petruck 2009), projects such as French FrameNet (Pitel 2009) or BiFrameNet (Fung and Chen 2004) also employ multilingual corpora and other existing lexical resources. Another difference concerns the tools used for data extraction and annotation. While the Japanese and Spanish FrameNets adopted the Berkeley FrameNet software (Baker et al. 2003) with slight modifications, other projects such as SALSA developed their own tools to conduct semi-automatic annotation on top of existing syntactic annotations found in the TIGER corpus, or they integrate off-the shelf software as is the case with French FrameNet or Hebrew FrameNet. FrameNets for other languages also differ in the methodology used to produce parallel lexicon fragments. While German FrameNet (Boas 2002) and Japanese FrameNet (Ohara 2009) rely on manual annotations, French FrameNet and BiFrameNet use semi-automatic and automatic approaches to create parallel lexicon fragments for French and Chinese. Finally, FrameNets for other languages also differ in their semantic domains and the goals they pursue. While most non-English FrameNets aim to create databases with broad coverage, other projects focus on specific lexical domains such as football (a.k.a. soccer) language (Schmidt 2009) or the language of criminal justice (Bertoldi et al. 2010). Finally, while the data from almost all non-English FrameNets are intended to be used by a variety of audiences, Multi FrameNet 6 is intended to support vocabulary acquisition in the foreign language classroom (see Atzler 2011).

3.2. Re-using (English) semantic frames

To exemplify how English FrameNet frames can be re-used for the creation of parallel lexicon fragments consider Boas' (2005a) discussion of the English verb *answer* evoking the `Communication_Response` frame and its counterpart *responder* in Spanish FrameNet. The basic idea is that since the two verbs are translation equivalents they should evoke the same

[5]For details about the different phrase types and grammatical functions, including the different types of null instantiation (CNI, DNI, and INI) (Fillmore 1986), see Fillmore et al. 2003b, Boas 2009, Fillmore & Baker 2010, and Ruppenhofer et al. 2010.

[6]http://www.coerll.utexas.edu/coerll/taxonomy/term/627

semantic frame, which should in turn be used as a common structuring device for combining the respective English and Spanish lexicon fragments. Since the MySQL databases representing each of the non-English FrameNets are similar in structure to the English MySQL database in that they share the same type of conceptual backbone (i.e., the semantic frames and frame relations), this step involves determining which English LUs are equivalent to corresponding non-English LUs.

However, before creating parallel lexicon fragments for Spanish and linking them to their English counterparts via their semantic frame it is necessary to first conduct a detailed comparison of the individual LUs and how they realize the semantics of the frame. To begin, consider the different ways in which the FEs of the `Communication_response` frame are realized with *answer*.

FE Name	Syntactic Realization
SPEAKER	NP.Ext, PP_*by*_Comp, CNI
MESSAGE	INI, NP.Obj, PP_*with*.Comp, QUO.Comp, Sfin.Comp
ADDRESSEE	DNI
DEPICTIVE	PP_*with*.Comp
MANNER	AVP.Comp, PPing_*without*.Comp
MEANS	PPing_*by*.Comp
MEDIUM	PP_*by*.Comp, PP_*in*.Comp, PP_*over*.Comp
TRIGGER	NP.Ext, DNI, NP.Obj, Swh.Comp

Table 1: Partial realization table for the verb *answer* (Boas 2005a)

Table 1 shows that that there is a significant amount of variation in how FEs of the `Communication_ Response` frame are realized with answer. For example, the FE DEPICTIVE has only one option for its syntactic realization, i.e. a PP complement headed by *with*. Other FEs such as SPEAKER and MANNER exhibit more flexibility in how the FEs of the frame are realized syntactically while yet another set of FEs such as MESSAGE and TRIGGER exhibit the highest degree of syntactic variation. Now that we know the full range of how the FEs of the `Communication_Response` frame are realized syntactically with *answer* we can take the next step towards creating a parallel lexical entry for its Spanish counterpart *responder*.

This step involves the use of bilingual dictionaries and parallel corpora in order to identify possible Spanish translation equivalents of *answer*. While this procedure may seem trivial, it is a rather lengthy and complicated process because it is necessary to consider the full range of valence patterns (the combination of FEs and their syntactic realizations) of the English LU *answer* listed in FrameNet. It lists a total of 22 different frame element configurations, totaling 32 different combinations in which these sequences may be realized syntactically. As the full valence table for *answer* is rather long we focus on only one out of the 22 frame element configurations, namely that of SPEAKER (Sp), MESSAGE (M), TRIGGER (Tr), and ADDRESSEE (A) in Table 2.

	Sp	M	Tr	A
a.	NP.Ext	NP.Obj	DNI	DNI
b.	NP.Ext	PP_*with*.Comp	DNI	DNI
c.	NP.Ext	QUO.Comp	DNI	DNI
d.	NP.Ext	Sfin.Comp	DNI	DNI

Table 2: Excerpt from the Valence Table for *answer* (Boas 2005a)

As Table 2 shows, the frame element configuration exhibits a certain amount of variation in how the FEs are realized syntactically: All four valence patterns have the FE SPEAKER realized as an external noun phrase, and the FEs TRIGGER and ADDRESSEE not realized overtly at the syntactic level, but null instantiated as Definite Null Instantiation (DNI). In other words, in sentences such as *He answered with another question* the FEs TRIGGER and ADDRESSEE are understood in context although they are not realized syntactically.

With the English-specific information about *answer* and the more general frame information in place we are now in a position to search for the corresponding frame element configuration of its Spanish translation equivalent *responder*. Taking a look at the lexical entry

of *responder* in Spanish FrameNet we see that the variation of syntactic realizations of FEs is similar to that of *answer* in Table 1.

FE Name	Syntactic Realizations
SPEAKER	NP.Ext, NP.Dobj, CNI, PP_por.COMP
MESSAGE	AVP.AObj, DNI, QUO.DObj, queSind.DObj, queSind.Ext
ADDRESSEE	NP.Ext, NP.IObj, PP_a.IObj, DNI, INI
DEPICTIVE	AJP.Comp
MANNER	AVP.AObj, PP_de.AObj
MEANS	VPndo.AObj
MEDIUM	PP_en.AObj
TRIGGER	PP_a.PObj, PP_de.PObj, DNI

Table 3: Partial Realization Table for the verb *responder* (Boas 2005a)

Spanish FrameNet also offers a valence table that includes for *responder* a total of 23 different frame element configurations. Among these, we find a combination of FEs and their syntactic realization that is comparable in structure to that of its English counterpart in Table 2 above.

	Sp	M	Tr	A
a.	NP.Ext	QUO.DObj	DNI	DNI
b.	NP.Ext	QueSind.DObj	DNI	DNI

Table 4: Excerpt from the Valence Table for *responder* (Boas 2005a)

Comparing Tables 2 and 4 we see that *answer* and *responder* exhibit comparable valence combinations with the FEs SPEAKER and MESSAGE realized syntactically while the FEs TRIGGER and ADDRESSEE are not realized syntactically, but are instead implicitly understood (they are definite null instantiations). With a Spanish counterpart in place it now becomes possible to link the Spanish set of frame element configurations in Table 4 with its English counterpart in Table 2 via the `Communication_Response` frame as the following Figure illustrates.

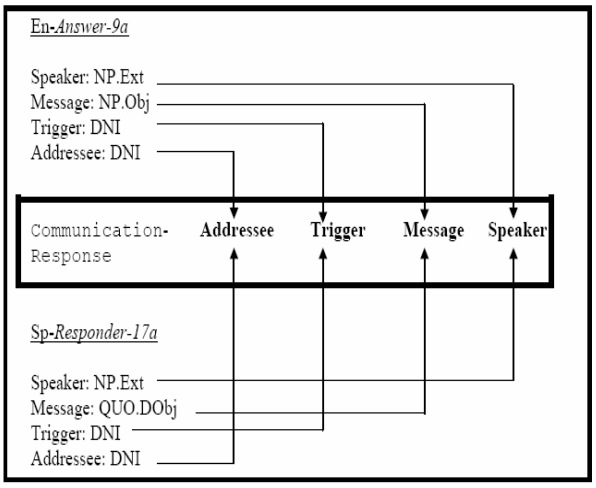

Figure 2: Linking partial English and Spanish lexicon fragments via semantic frames (Boas 2005a)

Figure 5 shows how the lexicon fragments of *answer* and *responder* are linked via the `Communication_Response` frame. The 'a' index points to the respective first lines in the valence tables of the two LUs (cf. Tables 2 and 4) and identifies the two syntactic frames as being translation equivalents of each other. At the top of Figure 2 we see the verb *answer* with one of its 22 frame element configurations, i.e. SPEAKER, TRIGGER, MESSAGE, and ADDRESSEE. Figure 2 shows for this configuration one possible set of syntactic realizations of these FEs, that given in row (a) in Table 2 above. The 9a designation following *answer* indicates that this lexicon fragment is the ninth configuration of FEs out of a total of 22 frame element configurations listed in the complete realization table. Of the ninth frame element configuration 'a' indicates that it is the first of a list of various possible syntactic realizations of these FEs (there are a total of four, cf. Table 2 above). As already pointed out, the FE SPEAKER is realized syntactically as an external NP, MESSAGE as an object NP, and both TRIGGER and ADDRESSEE are null instantiated. The bottom of Figure 2 shows *responder* with the first of the 17 frame element configurations (recall that there are a total of 23). For one of these configurations, we see one subset of syntactic realizations of these FEs, namely the first row catalogued by Spanish FrameNet for this configuration (see row (a) in Table 3).

The two parallel lexicon fragments at the top and the

bottom of Figure 2 are linked by indexing their specific semantic and syntactic configurations as equivalents within the `Communication_Response` frame. This linking is indicated by the arrows pointing from the top and the bottom of the partial lexical entries to the mid-section in Figure 2, which symbolizes the `Communication_Response` frame at the conceptual level, i.e. without any language-specific specifications. Note that this procedure does not automatically link the entire lexical entries of *answer* and *responder* to each other. Establishing such a correspondence link connects only the relevant frame element configurations and their syntactic realizations in Tables 2 and 4 via the common semantic frame, because they can be regarded as translation equivalents.

Although linking the two lexicon fragments this way results in a systematic way of creating parallel lexicon fragments based on semantic frames (which serve as interlingual representations), it is not yet possible to automatically create or connect such parallel lexicon fragments. This means that one must carefully compare each individual part of the valence table of a LU in the source language with each individual part of the valence table of a LU in the target language. This step is extremely time intensive because it involves a detailed comparison of bilingual dictionaries as well as electronic corpora to ensure matching translation equivalents. Recall that Figure 2 represents only a very small set of the full lexical entries of *answer* and *responder*. The procedure outlined above will have to be repeated for each of the 32 different valence patterns of *answer* – and its (possible) Spanish equivalents. The following section addresses a number of other issues that need to be considered carefully when creating parallel lexicon fragments based on semantic frames.

4. Cross-linguistic problems

Creating parallel lexicon entries for existing English FrameNet entries and linking them to their English counterparts raises a number of important issues, most of which require careful (manual) linguistic analysis. While some of these issues apply to the creation of parallel entries across the board, others differ depending on the individual languages or the semantic frame. The following subsections, based on Boas (to appear), briefly address some of the most important issues, which all have direct bearing on how the semantics of a frame are realized syntactically across different languages.

4.1. Polysemy and profiling differences

While translation equivalents evoking the same frame are typically taken to describe the same types of scenes, they sometimes differ in how they profile FEs. For example, Boas (2002) discusses differences in how *announce* and various German translation equivalents evoke the `Communication_Statement` frame. When *announce* occurs with the syntactic frame [NP.Ext _ NP.Obj] to realize the SPEAKER and MESSAGE FEs as in *They announced the birth of their child*, German offers a range of different translation equivalents, including *bekanntgeben*, *bekanntmachen*, *ankündigen*, or *anzeig-en*. Each of these German LUs comes with its own specific syntactic frames that express the semantics of the `Communication_ Statement` frame. When *announce* is used to describe situations in which a message is communicated via a medium such as a loudspeaker (e.g. *Joe announced the arrival of the pizza over the intercom*), German offers *ansagen* and *durchsagen* as more specific translation equivalents of *announce* besides the more general *ankündigen*. Thus, by providing different LUs German offers the option of profiling particular FEs of the `Communication_Statement` frame, thereby allowing for the representation of subtle meaning differences of the frame and the perspective given of a situation (see Ohara 2009 on similar profiling differences between English and Japanese LUs evoking the `Risk` frame).

4.2. Differences in lexicalization patterns

Languages differ in how the lexicalize particular types of concepts (see Talmy 1985), which may directly influence how the semantics of a particular frame are realized syntactically. For example, in a comparative study of English, Spanish, Japanese, and German motion verbs in *The Hound of the Baskervilles* (and its translations), Ellsworth et al. (2006) find that there are a number of differences in how the various concepts of motion are associated with different types of semantic frames. More specifically, they show that English *return* (cf. *The wagonette was paid off and ordered to return to Coombe Tracey forthwith, while we started to walk to*

Merripit House) and Spanish *regresar* both evoke the `Return` frame, whereas the corresponding German *zurückschicken* evokes the `Sending` frame. These differences demonstrate that although the concept of motion is incorporated into indirect causation, the frames expressing indirect causation may vary from language to language (see Burchardt et al. 2009 for a discussion of more fine-grained distinctions between verbs evoking the same frame in English and German).

4.3 Polysemy and translation equivalents

Finding proper translation equivalents is typically a difficult task because one has to consider issues surrounding polysemy (Fillmore & Atkins 2000, Boas 2002), zero translations (Salkie 2002, Boas 2005a, Schmidt 2009), and contextual and stylistic factors (Altenberg & Granger 2002, Hasegawa et al. 2010), among others. To illustrate, consider Bertoldi's (2010) discussion of contrastive legal terminology in English and Brazilian Portuguese. Based on the English `Criminal Process` frame (see FrameNet), Bertoldi finds that while there are some straightforward translation equivalents of English LUs in Portuguese, others involve a detailed analysis of the relevant polysemy patterns.

Consider Figure 3, which compares English and Portuguese LUs in the `Notification_of_charges` frame. The first problem discussed by Bertoldi (2010) addresses the fact that although there are corresponding Portuguese LUs such as *denunciar*, they do not evoke the same semantic frame as the English LUs, but rather a frame that could best be characterized as evoking the `Accusation` frame. The second problem is that six Portuguese translation equivalents of the English LUs evoking only the `Notification_of_charges` frame, i.e. *acusar*, *acusação*, *denunciar*, *denuncia*, *pronunciar*, and *pronuncia*, potentially evoke three different frames.

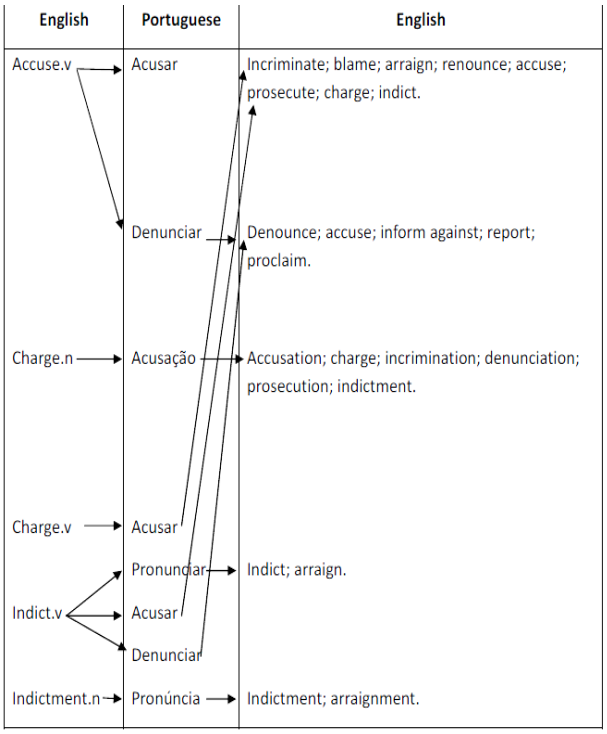

Figure 3: English LUs from the `Notification_of_Charges` frame and their Portuguese translation equivalents (Bertoldi 2010: 6)

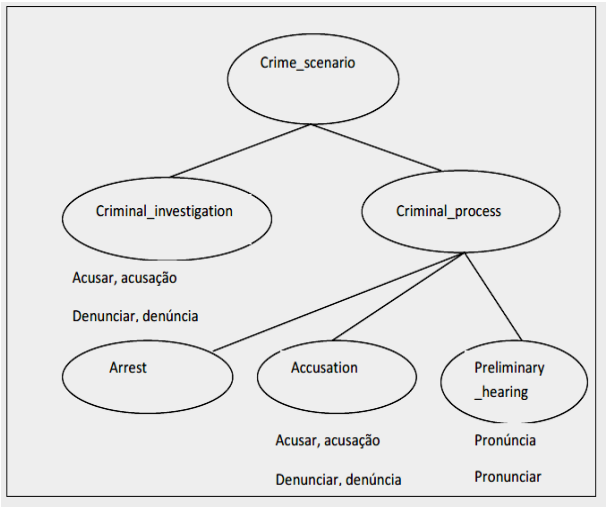

Figure 4: LUs evoking multiple frames in the Portuguese `Crime_scenario` frame (Bertoldi 2010: 7)

This leads Bertoldi to claim that the LUs *acusar*, *acusação*, *denunciar*, and *denuncia* may *evoke* two different `Criminal_Process` sub-frames, besides other general language, non-legal specific frames, as is illustrated by Figure 4. Bertolid's analysis shows that finding translation equivalents is not always an easy task

and that one needs to pay close attention to different polysemy networks across languages, which may sometimes be influenced by systematic differences such as differences between legal systems.

4.4 Universal frames?

Claims about the universality of certain linguistic features are abundant in the literature. When it comes to semantic frames the question is whether frames derived on the basis of English are applicable to the description and analysis of other languages (and vice versa). While a number of studies on motion verbs (Fillmore & Atkins 2000, Boas 2002, Burchardt et al. 2009, Ohara 2009) and communication verbs (Boas 2005a, Subirats 2009), among other semantic domains, suggest that there are frames that can be re-used for the description and analysis of other languages, there also seem to be culture-specific frames that may not be re-usable without significant modification.

One set of examples comes from the English `Personal_Relationship` frame, whose semantics appears to be quite culture-specific. Atzler (2011) shows that concepts such as dating (*to date*) seem to be quite specific to Anglo culture and may not be directly applicable to the description of similar activities in German. Another, perhaps more extreme example, is the term *sugar daddy*, which has no exact counterpart in German, but instead requires a lengthy paraphrase in German to render the concept of this particular type of relationship in German.

A second example comes from the intransitive Finnish verb *saunoa* (literally 'to sauna'), which has no direct English counterpart because it very culture-specific, and in effect evokes a particular type of frame. To this end, Leino (2010:131) claims that this verb (and correspondingly the Finnish `Sauna` frame) “expresses a situation in which the referent of the subject goes to the sauna, is in the sauna, participates in the sauna event, or something of the like.” Dealing with such culture-specific frames thus requires quite lengthy paraphrases to arrive at an approximation of the semantics of the frame in English.

5. Conclusions and outlook

This paper has outlined some of the basic steps underlying the creation of parallel lexicon fragments. Employing semantic frames for this purpose is still a work in progress, but the successful compilation of several FrameNets for languages other than English is a good indication that this methodology should be pursued further.

Clearly, the problems outlined in the previous section need to be solved. The first problem, polysemy and profiling differences, is perhaps the most daunting one. Decades of linguistic research into these issues (see, e.g. Leacock & Ravin 2000, Altenberg & Granger 2002) seem to suggest that there is no easy solution that could be implemented to arrive at an automatic way of analyzing, comparing, and classifying different polysemy and lexicalization patterns across languages. This means that for now these issues need to be addressed manually, in the form of careful linguistic analysis, in the near future.

The same can be said about the problems surrounding lexicalization patterns, zero translations, and the universality of frames. Without a detailed catalogue of linguistic analyses of these phenomena in different languages, and a comparison across language pairs, any efforts regarding the effective linking of parallel lexicon fragments, whether on the basis of semantic frames or not, will undoubtedly hit many roadblocks.

6. Acknowledgements

Work on this paper was supported by a fellowship for experienced researchers from the Alexander von Humboldt Foundation, as well as by Title VI grant #P229A100014 (Center for Open Educational Resources and Language Learning) to the University of Texas at Austin.

7. References

Altenberg, B., Granger, S. (2002): Recent trends in cross-linguistic studies. In B. Altenberg & S. Granger (Eds.), Lexis in Contrast. Amsterdam/Philadelphia: John Benjamins, pp. 3-50.

Atkins, B.T.S., Fillmore, C.J., Johnson, C.R. (2003): Lexicographic relevance: Selecting information from corpus evidence. International Journal of Lexicography, 16(3), pp. 251-280.

Atzler, J. (2011): Twist in the line: Frame Semantics as a vocabulary teaching and learning tool. Doctoral Dissertation, The University of Texas at Austin.

Baker, C.F., Fillmore, C.J., Lowe, J.B. (1998): The Berkeley FrameNet Project. In COLING-ACL '98: Proceedings of the Conference, pp. 86-90.

Baker, C.F., Fillmore, C.J., Cronin, B. (2003): The Structure of the FrameNet Database. In International Journal of Lexicography, 16(3), pp. 281-296.

Bertoldi, A. (2010): When translation equivalents do not find meaning equivalence: a contrastive study of the frame `Criminal_Process`. Manuscript. UT Austin.

Bertoldi, A., Chishman, R., Boas, H.C. (2010): Verbs of judgment in English and Portuguese: What contrastive analysis can say about Frame Semantics. Calidoscopio, 8 (3), pp. 210-225.

Boas, H.C. (2001): Frame Semantics as a framework for describing polysemy and syntactic structures of English and and German motion verbs in contrastive computational lexicography. In Proceedings of Corpus Linguistics 2001, pp. 64-73.

Boas, H.C. (2002): Bilingual FrameNet dictionaries for machine translation. In: Proceedings of the Third International Conference on Language Resources and Evaluation, Vol. IV, pp. 1364-1371.

Boas, H.C. (2005a): Semantic frames as interlingual representations for multilingual lexical databases. International Journal of Lexicography, 18(4), pp. 445-478.

Boas, H.C. (2005b): From theory to practice: Frame Semantics and the design of FrameNet. In S. Langer & D. Schnorbusch (Eds.), Semantik im Lexikon. Tübingen: Narr, pp. 129-160.

Boas, H.C. (2009): Recent trends in multilingual lexicography. In H.C. Boas (Ed.), Multilingual FrameNets in Computational Lexicography: Methods and Applications. Berlin/New York: Mouton de Gruyter, pp. 1-36.

Boas, H.C. (to appear): Frame Semantics and Translation. In I. Antunano and A. Lopez (Eds.), Translation in Cognitive Linguistics. Berlin/New York: Mouton de Gruyter.

Burchardt, A., Erk, K., Frank, A., Kowalski, A., Pado, S., & Pinkal, M. (2009): Using FrameNet for the semantic analysis of German: annotation, representation, and automation. In H.C. Boas (Ed.), Multilingual FrameNets in Computational Lexicography: Methods and Applications. Berlin/New York: Mouton de Gruyter, pp. 209-244.

Dux, R. (2011): A frame-semantic analysis of five English verbs evoking the Theft frame. M.A. Report. The University of Texas at Austin.

Ellsworth, M, Ohara, K., Subirats, C., & Schmidt, T. (2006): Frame-semantic analysis of motion scenarios in English, German, Spanish, and Japanese. Presentation given at the 4th International Conference on Construction Grammar (ICCG-4), Tokyo.

Fillmore, C.J. (1982): Frame Semantics. In Linguistic Society of Korea (Ed.), Linguistics in the Morning Calm. Seoul: Hanshin, pp. 111-138.

Fillmore, C.J. (1985): Frames and the semantics of understanding. Quadernie di Semantica, 6, pp. 222-254.

Fillmore, C.J. (2006): Pragmatically controlled zero anaphora. BLS, 12, pp. 95-107.

Fillmore, C.J. (2007): Valency issues in FrameNet. In: T. Herbst & K. Götz-Vetteler (Eds.), Valency: theoretical, descriptive, and cognitive issues. Berlin/New York: Mouton de Gruyter, pp. 129-160.

Fillmore, C.J., Atkins, B.T.S. (2000): Describing polysemy: The case of 'crawl'. In Y. Ravin and C. Laecock (Eds.), Polysemy. Oxford: Oxford University Press, pp. 91-110.

Fillmore, C.J., Baker, C.F. (2010): A frames approach to semantic analysis. In: B. Heine and H. Narrog (Eds.), The Oxford Handbook of Linguistic Analysis. Oxford: Oxford University Press, pp. 313-340.

Fillmore, C.J., Petruck, M.R.L. (2003): FrameNet Glossary. International Journal of Lexicography, 16(3), pp. 359-361.

Fillmore, C.J, Johnson, C.R., Petruck, M.R.L. (2003a): Background to FrameNet. International Journal of Lexicography, 16(3), pp. 235-251.

Fillmore, C.J., Petruck, M.R.L., Ruppenhofer, J., Wright, A. (2003b): FrameNet in action: The case of Attaching. International Journal of Lexicography, 16(3), pp. 297-333.

Fontenelle, T. (1997): Using a bilingual dictionary to create semantic networks. International Journal of Lexicography, 10(4), pp. 275-303.

Fung, P., Chen, B. (2004): BiFrameNet: Bilingual frame semantics resource construction by cross-lingual induction. In Proceedings of COLING 2004.

Hasegawa, Y., Lee-Goldman, R., Ohara, K., Fujii, S.,

Fillmore, C.J. (2010): On expressing measurement and comparison in English and Japanese. In H.C. Boas (Ed.), Contrastive Studies in Construction Grammar. Amsterdam/Philadelphia: John Benjamins, pp. 169-200.

Heid, U. (1996): Creating a multilingual data collection for bilingual lexicography from parallel monolingual lexicons. In Procedings of the VIIth EURALEX International Congress, pp. 559-573.

Leino, J. (2010): Results, cases, and constructions: Argument structure constructions in English and Finnish. In H.C. Boas (Ed.), Contrasive Studies in Construction Grammar. Amsterdam/Philadelphia: John Benjamins, pp. 103-136.

Lowe, J.B., Baker, C.F., Fillmore, C.J. (1997): A frame-semantic approach to semantic annotation. In Proceedings of the SIGLEX Workshop on Tagging Text with Lexical Semantics: Why, What, and How? Held April 4-5, in Washington, D.C., in conjunction with ANLP-97.

Ohara, K. (2009): Frame-based contrastive lexical semantics in Japanese FrameNet: The case of risk and kakeru. In H.C. Boas (Ed.), Multilingual FrameNets in Computational Lexicography: Methods and Applications. Berlin/New York: Mouton de Gruyter, pp. 163-182.

Petruck, M.R.L. (1996): Frame Semantics. In J. Verschueren, J.-O. Östman, J. Blommaert, C. Bulcaen (Eds.), Handbook of Pragmatics. Amsterdam/Philadelphia: John Benjamins, pp. 1-13.

Petruck, M.R.L. (2009): Typological considerations in constructing a Hebrew FrameNet. In H.C. Boas (Ed.), Multilingual FrameNets in Computational Lexicography: Methods and Applications. Berlin/New York: Mouton de Gruyter, pp. 183-208.

Petruck, M.R.L, Boas, H.C. (2003): All in a day's week. In . Hajicova, A. Kotesovcova, J. Mirovsky (Eds.), Proceedings of CIL 17. CD-ROM. Prague: Matfyzpress.

Petruck, M.R.L., Fillmore, C.J., Baker, C.F., Ellsworth, M., Ruppenhofer, J. (2004): Reframing FrameNet data. In Proceedings of the 11th EURALEX International Conference, pp. 405-416.

Pitel, G. (2009): Cross-lingual labeling of semantic predicates and roles: A low-resource method based on bilingual L(atent) S(emantic) A(nalysis). In H.C. Boas (Ed.), Multilingual FrameNets in Computational Lexicography: Methods and Applications. Berlin/New York: Mouton de Gruyter, pp. 245-286.

Ruppenhofer, J., Ellsworth, M., Petruck, M.R.L., Johnson, C., Scheffczyk, J. (2010): FrameNet II: Extended theory and practice. Available at http://framenet.icsi.berkeley.edu

Salkie, R. (2002): Two types of translation equivalence. In B. Altenberg and S. Granger (Eds.), Lexis in Contrast. Amsterdam/Philadelphia: John Benjamins, pp. 51-72.

Schmidt, T. (2009): The Kicktionary – A multilingual lexical resource of football language. In H.C. Boas (Ed.), Multilingual FrameNets in Computational Lexicography: Methods and Applications. Berlin/New York: Mouton de Gruyter, pp. 101-134.

Subirats, C. (2009): Spanish FrameNet: A frame-semantic analysis of the Spanish lexicon. In H.C. Boas (Ed.), Multilingual FrameNets in Computational Lexicography: Methods and Applications. Berlin/New York: Mouton de Gruyter, pp. 135-162.

Talmy, L. (1985): Lexicalization patterns: semantic structures in lexical forms. In T. Shopen (Ed.), Language Typology and Syntactic Description. Cambridge: Cambridge University Press, pp. 57-149.

Ziem, A. (2008): Frames und sprachliches Wissen. Berlin/New York: Mouton de Gruyter.