24
FrameNet

Josef Ruppenhofer, Hans C. Boas and Collin F. Baker

24.1 Introduction

FrameNet is a research project that seeks to instantiate the principles of Frame Semantics as proposed by Charles J. Fillmore (1977, 1985) in the analysis of the English lexicon. The main idea is that the meanings of words are best characterized in terms of experience-based schematizations of events and objects in the speaker's world. Such schematizations relate to particular types of events and the participants and circumstances involved in them. The schematizations are referred to in Frame Semantics as "semantic frames". Individual word senses are called "lexical units" (LUs). When a lexical unit belongs to given frame, the LU is said to "evoke" that frame. Typically, senses of several distinct lemmas evoke a common schematization, that is, groups of word senses evoke the same frame. The roles associated with an event are referred to as "frame elements" (FEs). One and the same system of analysis applies to events, relations, states, and objects; the frame-evoking expressions may be single items or multi-word expressions, and they may belong to any syntactic category (Fillmore, Johnson and Petruck 2003).

For example, in (1) the adjective *different* (the "target" LU) is said to evoke the Similarity frame, which is also evoked, among others, by further adjectives such as *alike* and *disparate*; verbs such as *differ* and *resemble*; nouns such as *similarity* and *distinction*; and the prepositions *like* and *unlike*, as well as multi-words such as *spitting image.n* and *take after.v*.

(1)
The qualities that make a poet are not **different** in kind from what other men have.

The Similarity frame is characterized by relationships holding between different FEs, such as *Entity*1 and *Entity2* (which may instead be jointly presented as *Entities)*. The frame has to do with the degree to which two or more distinct entities, whether concrete or abstract, are similar to or different from each other, along some implicit or explicit *Dimension*. The *Degree* of similarity or difference may or may not be specified explicitly.

In addition to assembling frame descriptions and lists of frame-evoking LUs, frame semanticists are also concerned with documenting the syntactic realizations of *Frame Elements* (Atkins, Fillmore and Johnson 2003, Fillmore 2007).

24.2 Historical perspective

The FrameNet project is the result of the application of the insights of Charles Fillmore's theory of Frame Semantics, which he and his colleagues developed during the 1980s and 1990s (Fillmore 1982, 1985; Fillmore and Atkins 1992). Frame Semantics grew out of Fillmore's earlier work on Case Grammar (Fillmore 1968), which proposed a list of semantic roles aimed at identifying grammatically relevant facets of a verb's meaning in terms of a set of basic labels such as Agentive, Instrumental, Dative, Factitive, Locative, etc., which served to identify the role played by each of the verb's arguments in the event it denotes. The collection of sets of different semantic roles came to be known as case frames, which specify a verb's semantic valency. One of the major ideas behind these semantic roles was that they are ordered in a hierarchy depending on their realization in terms of grammatical function. This hierarchy puts the semantic role Agentive at the top, followed by Instrumental, Objective, and other semantic roles. Following the initial success of Fillmore's proposals, several problems emerged regarding the status of semantic roles, including issues with the definition and granularity of semantic roles, problems reflecting cross-role generalizations, and problems of one-to-one correspondence (for an overview, see Levin and Rappaport Hovav 2005). As a result, during the 1970s, Fillmore moved away from his initial proposal that semantic roles are primary, and should limited to a small set of roles (see, e.g., Fillmore 1975, 1976, 1977, 1978, 1979). The goal of this move was to develop a more elaborated model of the semantics of understanding, in contrast to a truth-conditional semantics, the prevalent approach to linguistic semantics at the time. More specifically, Fillmore suggested that so-called semantic frames should be regarded as primary for the description and analysis of meaning (and its syntactic relevance), and that semantic roles should be defined in terms of their semantic frames. The early 1980s saw the first fully developed version of Frame Semantics, which was articulated as a model based on the full and rich understanding required for the production of a text as well as the understanding AuQ4 of a text (Fillmore 1982) (for details, see Ziem 2008, Busse 2012, and Boas (in press)). Subsequently, Fillmore and Atkins' (1992) paper on the semantics of *risk* can be regarded as the blueprint for the architecture of FrameNet, which started in 1997 at the International Computer Science Institute in Berkeley, California, with funding from the National Science Foundation.

24.3 Core issues and topics: the FrameNet database

The core activity of FrameNet is the compilation of a lexicographic database, which currently (as of June, 2016) holds more than 1,223 frame descriptions together with lexical entries for more than 13,500 LUs and more than 202,000 manually annotated examples. The FrameNet database is made available free of charge for any purpose, and is used by hundreds of research groups worldwide.[1]

24.3.1 Lexical data

The lexicographic process starts with FN (FrameNet) staff members proposing a frame description, including what sorts of FEs the frame needs and what LUs might evoke it. At this stage, lexicographers rely both on their intuitions and on carefully constructed queries of electronic corpora. The basic criterion for delineating the boundaries of a frame is that all LUs should evoke the same type of event[2] and share the same inventory and configuration of FEs. In other words, the FEs should have the same semantic types, be of the same relative importance (which roughly equates to the traditional argument-adjunct distinction), and take the same pragmatic perspective (e.g. in active-form clauses, the verbs in the frame should take the same FE as their subject). For further discussion of these points, see Petruck et al. (2004).

24.3.2 Annotation of data

In the next step, example sentences are retrieved from electronic corpora for each target LU; trained annotators then use the FrameNet Desktop software to manually annotate roughly 10 to 20 sentences per LU. As the annotators apply labels showing which parts of the sentence fill which FE role, labels for their phrase types (PTs), and their grammatical functions (GFs) *vis-à-vis* the target word are added automatically. (The GF and PT labels can be manually corrected if necessary.) Annotations for three sentences are given in example 2:

(2)

a. [Customs officers AGENTS] **substituted** [the drugs OLD] [with another substance NEW] [as the consignment arrived TIME] . . .
b. [this measure NEW] was **substituted** [for birth weight OLD][in the regression equation PLACE] [CNI AGENT]
c. [He AGENT] **substituted** [lower price labels NEW] [for those on the goods OLD].

The goal is to arrive at an annotation set capturing every possible combination of FE, GF, and PT. For instance, for the verb *substitute* in the frame **Replacing**, the annotations in (2a) and (2b) capture different mappings of the FEs OLD and NEW to grammatical functions due to a difference in voice, whereas (2a) and (2c) illustrate the possibility to express either of the FEs OLD and NEW as an object in an active-form clause, with the other being expressed as a PP. When the annotation is completed, the valence of each LU is automatically summarized by abstracting over the annotated examples. Finally, the lexical entry is "produced" as a report from the FrameNet database. It contains a brief definition of the LU together with the frame it evokes and valence tables that exhaustively document every attested combinatorial possibility of FEs and their syntactic realizations. Depending on the LU, it may also contain information such as the semantic type of the LU and a list of annotated support or controller verbs, etc. For an in-depth description of the FN workflow, see Fillmore et al. (2003), Ruppenhofer et al. (2010), and Fillmore and Baker (2010). Baker, Fillmore and Cronin (2003) provide a technical description of the FN database.

Since 2004, FN staff have also been annotating continuous texts, thereby creating lexical entries and a continuous corpus of annotated sentences. This full-text annotation differs from FN's lexicographic research in several ways:

1 Typically, between 2 and 10 LUs are annotated per sentence, in as many different frames. By contrast, in the lexicographic work, only one LU per sentence is annotated.
2 With full texts, annotators must label whatever the text contains, regardless of syntactic complexity, ambiguity, rhetorical infelicity, etc., whereas lexicographers can choose the clearest and simplest examples of a given LU to annotate
3 Full-text annotation drives the discovery of new frames, that is, if there are no existing frames they have to be created "on the fly".

24.3.3 Frame-to-frame relations

Over time, the number of semantic frames has grown to more than 1,000. To show that frames are not simply isolated entities, FN has developed an elaborate system of frame-to-frame relations producing a network of frames in which some frames are instances of others, some are components of others, etc. The most commonly found relations include those representing generalizations (INHERITANCE, USING, PERSPECTIVE ON), complex events (SUBFRAME, PRECEDES), and "systematic" relations (CAUSATIVE OF, INCHOATIVE OF); an additional relation, called "SEE ALSO" serves as a cross-reference between frames (see Fillmore and Baker (2010) and Ruppenhofer et al. (2010) for more details on frame-to-frame relations).

24.3.4 Access to FrameNet

Although machine-readable data releases representing snapshots of the evolving FrameNet database are regularly made available for download for computational use,[3] human users mainly access FrameNet by browsing the FrameNet website.[4] The top-level page of FN's website presents several ways to access the data, which is updated daily.

The information in FN can be accessed from two major perspectives: beginning with a concept (frame) to be encoded ("onomasiologically") and looking for lemmas that express it, or beginning with a form (lemma) to be decoded ("semasiologically"), that is, finding the concepts (frames) that it evokes. The first perspective is realized by the Frame Index, which leads to a page with an alphabetized list of frames in a navigation pane on the left and a larger panel on the right in which descriptions of chosen frames are displayed. The frame descriptions contain, in order:

1 a definition;
2 a list of FEs, grouped according to grammatical prominence, with each item including a definition and one or more examples;
3 a list of frame-to-frame relations; and
4 a table of LUs that evoke the frame.

The FEs used in the example sentences of the frame and FE definitions are identified by colours. In the frame-to-frame relations section, hyperlinks provide access to related frames. Similarly, in the table of associated LUs, each item comes with two links: one to the Annotation Report for that LU and the other to the Lexical Entry Report.

The Annotation Report consists of a listing of the FEs for the associated frame and then a display of annotated instances of the LU, sub-divided into sets ('sub-corpora') according to the syntactic patterns used to extract them. These patterns usually select for particular syntactic complementation types. For example, for the noun *oath* in the Commitment frame, there are sub-corpora named *VPto* and *Sthat* which are supposed to contain the uses with infinitival and finite-clause complements, respectively.

The Lexical Entry Report contains two types of summary: First, it gives a list of the individual FEs and the ways in which they are realized in the annotated instances. Figure 24.1 shows the syntactic realizations found with the verb *give* in the **Giving** frame for the three core frame elements Donor, Theme and Recipient. Note that the notion of "realization" used by FrameNet extends to cases in which the Frame Element is not realized explicitly at all but instead "null instantiated". This can happen through the idiosyncratic licensing of a lexical unit or the licensing of a syntactic construction.

Frame Element	Number Annotated	Realization (s)
Donor	(51)	CNI.--(12) DNI.--(2) NP.Ext (36) PP[by].Dep (1)
Recipient	(51)	PP[to] .Dep (16) DNL.--(5) INI.--(2) NP.Ext (4) NP.Obj (21) NP.Dep (3)
Theme	(50)	NP.Ext (2) NP.Obj (26) NP.Dep (20) PP[of].Dep (1) INI.-- (1)

Figure 24.1 Partial syntactic realization table for *give* in the frame **Giving**

FrameNet distinguishes three types of missing elements, abbreviates DNI, INI, and CNI. DNI ("definite null instantiation") marks FEs that are unrealized but which have to be recoverable from the context. An example is the FE Recipient in a sentence such as *The prize was given by an international committee*. Figure 24.1 shows that there are five cases of DNI for the FE Recipient among the annotated instances for *give*. INI ("indefinite null instantiation") covers FEs that are merely existentially bound, an example of which is the FE Theme in *Sure, I gave to the Red Cross last year – everybody did*. CNI ("constructional null instantiation") marks all omissions licensed by a syntactic construction. A typical case is the omission of agentive FEs in the imperative construction, as in *give a generous gift to Goodwill today*.

The Lexical Entry report also contains another summary table that shows all the attested combinations in which **groups** of FEs were co-realized. Some of the attestations found with *substitute* in the **Replacing** frame are shown in Figure 24.2.

The last row in the Valence Pattern Table (Figure 24.2) shows that there is exactly one sentence in the FN database with the pattern Agent as CNI, New as External (Subject), Old in a PP*for*, Purpose in an infinitive phrase, and Time in a PP*after*, which is the sentence *After verification for the effect of the typesetting code combinations, they were substituted for more computer-recognizable codes to simplify further processing*. Both of these tables contain hyperlinked counts for the number of annotated instances exemplifying the FE (or valence pattern), which can be displayed by following the links. The two associated reports for a given lexical unit are also interlinked.

The decoding point of view is represented by the alphabetized "Lexical Unit Index". Users can simply browse this index or use a search box to retrieve LUs. There is no direct representation of lemmas per se; the different LUs (senses) connected to a given lemma simply follow each other. Unlike on the semantic side, where one can easily follow links from frame to related frame, on the form side there are no direct links between the LUs associated with the same lemma, or between LUs belonging to morphologically related lemmas. The

6 TOTAL	Agent	New	Old	Place	
(2)	CNI —	NP Ext	PPLfor] Dep	PP[in] Dep	
(1)	NP Ext	NP Obj	PP[for] Dep	AVP Dep	
(1)	NP Ext	NP Obj	PPtfor] Dep	PP[in] Dep	
(1)	NP Ext	NP Obj	PPLfor] Dep	PPLon] Dep	
(1)	NP Ext	PP[with] Dep	NP Obj	AVP Dep	
2 TOTAL	Agent	New	Old	Purpose	
(1)	CNI --	NP Ext	PPLfor] Dep	VPto Dep	
(1)	NP Ext	NP Obj	DN1 —	VPto Dep	
1 TOTAL	Agent	New	Old	Purpose	Time
(1)	CNI --	NP Ext	PPLfor] Dep	VPto Dep	PP [after] Dep

Figure 24.2 Partial valence pattern table for *substitute* in the **Replacing** frame

entry for each LU provides information on the work-flow status of the item and three links: one to the Frame description, one to the Annotation Report, and one to the Lexical Entry Report (as described earlier).

A third type of access is based on the fact that FrameNet performs some annotation of running texts, which are listed on the Full-Text annotation index. The Full-Text Annotation Report pairs a display of a document's running text with a display of the annotations of specific user-selectable items from the text. Although the annotation instances in full-text documents are also integrated into the Annotation and Lexical Entry Reports for the relevant lexical units, they can be viewed in their document context only via the Full-text Annotation Report.

A fourth kind of data access focuses on the frame relations in the lexical part of the database. Although these relations are also listed in the Frame reports, the FrameGrapher visualization tool (*https://framenet.icsi.berkeley.edu/fndrupal/FrameGrapher*) is uniquely suited to interactively exploring the topology of relations that exist in a set of related frames. For example, users can choose a focal frame on which the display is to be centred and a group of frame relation types they are interested in. In order to manage the size and layout of the resulting display, users can specify the number of levels of the (sub-)hierarchy and the number of leaf nodes to be displayed. Clicking on one of the frames in the graph displays a re-centred graph with a new focal frame; clicking on the head of the arrow representing a frame-frame relation will create an expanded graph, where the frame element-to-frame element relations that go along with the chosen frame-to-frame relationship are displayed in detail (see Fillmore and Baker 2010: 332–333 for details).

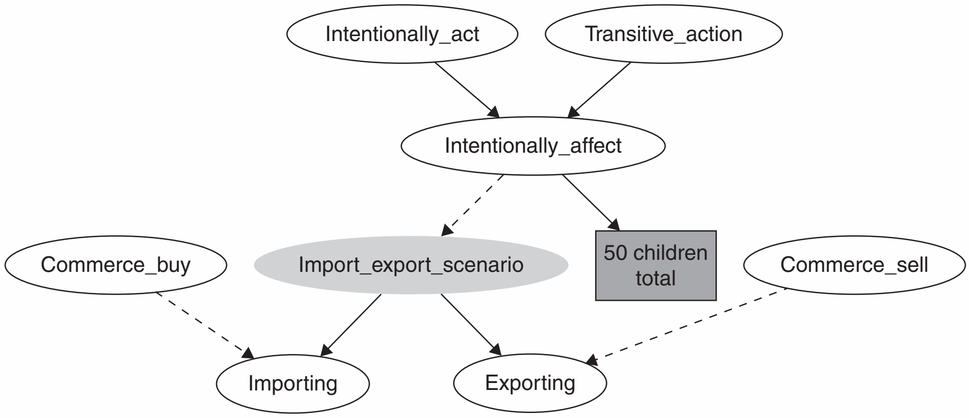

Figure 24.3 FrameGrapher display for import-export–related frames

Figure 24.3 shows a set of frames related to the idea of Importing and Exporting as portrayed in the FrameGrapher. The dashed (green) arrows indicate USING relations, the thick (red) arrows indicate INHERITANCE relations, and the thin (pink) arrows show RESPECTIVE ON relations. Thus, **Importing** and **Exporting** are different perspectives on the same event, the buyer's and the seller's respectively. The **Import export scenario** "uses" the frame **Intentionally affect**, a high-level frame with many children.

Finally, a multi- and cross-lingual view on FrameNet data is provided by the FrameSQL search tool (http://sato.fm.senshu-u.ac.jp/frameSQL/fn2_15/notes) which is developed and maintained independently by Prof. Hiroaki Sato of Senshu University, Japan. In addition to the English-language FrameNet, FrameSQL can search the frame semantic annotations of FrameNet's German, Japanese, and Spanish sister projects. Moreover, the tool integrates information from all four resources. As a result, users can see the information from the different sources aligned seamlessly, as if they were a single database. Hyperlinks among the databases and accompanying search capabilities facilitate comparison between the semantic structures of corresponding frames or lexical units across the languages (see Sato 2008).

Independently of the FN website, human users may also access the first three kinds of information (Frame, LU-specific Lexical Entry and Annotation Reports, and Full-text annotation reports) in the FN data release.[5] The download consists of a set of static interlinked XML files that are browser-viewable by means of accompanying XSL/Javascript scripts, although the FrameGrapher and FrameSQL tools are not included.

24.3.5 Efforts to generate or enhance FrameNet resources through crowdsourcing

Hong and Baker (2011) investigated frame disambiguation by non-expert annotators. The task of frame element assignment remained outside the purview of their experiments. The authors experimented with various approaches for frame assignment, the most promising being one in which sentences containing the lemma of interest were grouped by frame. This approach yielded accuracy scores between 73 percent and 92 percent. Chang et al. (2015) and extended the grouping approach of Hong und Baker (2011) to additional lemmas, and they also incorporate a mechanism that gives the annotators feedback on their work. The latter resulted in an increase of average frame assignment accuracy from 78 percent to 92 percent.

Fossati et al. (2013) also crowdsourced frame semantic annotation, including the labelling of frame elements. The reported accuracy for frame disambiguation is 90 percent and that for FE labelling 68.7 percent. (The FE figure is low because it is calculated over all FEs not merely the ones where the assumed frame was correct.) However, Fossati et al. use simplified definitions instead of the original FrameNet definitions which are often difficult to understand for laypersons due to occurrences of technical/linguistic terminology.

24.3.6 FrameNet and multilinguality

24.3.6.1 Lexicographic sister projects in other languages

The frames created by the Berkeley FrameNet for the lexicon for English have also been "recycled" for other languages. The first efforts in this direction applied frame-semantic insights to the systematic analysis of the lexicons of languages other than English (Heid 1996, Fontenelle 1997). Subsequent research demonstrated in greater detail how semantic frames derived on the basis of English data could be employed for the creation of FrameNets for other languages (Fillmore and Atkins 2000, Boas 2005, Petruck and Boas 2003). The results of this research inspired the creation of FrameNets for other languages, most notably Spanish (Subirats and Petruck 2003; Subirats 2009), Japanese (Ohara et al. 2003, Ohara 2009), German (Burchardt et al. 2009), Brazilian Portuguese (Salomão *et al*. 2013), Chinese (Liu 2011), Swedish (Borin et al. 2010), and French (Candito et al. 2014). Although the technical resources and workflows of the non-English FrameNets differ from each other, they all produce FN-style entries similar to the English original discussed earlier, including detailed information about how the semantics of a given frame are realized syntactically by different LUs evoking that frame in that language. This information allows researchers to systematically conduct contrastive and comparative research on topics in lexical semantics such as polysemy (Fillmore and Atkins 2000, Willems 2012), typological differences in profiling properties (Ohara 2009, Petruck 2009), and the interface between the lexicon and syntax (Hasegawa et al. 2010, Bouveret 2012). The multilingual data contained in the FrameNets for various languages is also relevant for cross-linguistic research on argument structure constructions, as shown by Boas (2003), Leino (2010), and Timyam and Bergen (2010), among others.

24.3.6.2 Uses in foreign language education

FrameNet data have also proven to be useful for foreign language education. To this end, an effort has been underway at the University of Texas at Austin to create an online learner's dictionary for German. The goal is to create simplified FrameNet-style lexical entries for all of the ca. 2,000 LUs in the first year online German textbook *Deutsch im Blick* (http://coerll.utexas.edu/dib/). The basic idea is to use existing English FrameNet frames for the description of those German words that beginning and intermediate learners of German have to learn during the course of their language studies. To achieve this goal, the Berkeley English FrameNet database was downloaded and installed on local servers at UT Austin and subsequently stripped of all English-specific information, leaving only the frames and the frame-to-frame relations intact. Then, a team of Germanic linguists at UT Austin went through the vocabulary lists of the textbook used for the first year of German instruction at UT in order to identify sets of relevant words evoking the same semantic frame. Using a variety of online corpora, the team of linguists extracted simple German example sentences and annotated them with frame-semantic information. Based on this information, the team then created user-friendly lexical entries to be stored in the stripped FrameNet database on local servers. The result is a set of simple to use contrastive German-English entries with notes on contrastive differences between German and English, culture-specific information, collocational information, and information about basic grammar usage. Finally, a team of web designers created an easy to use website for presenting the resulting information to learners of German. The prototype of the German Frame-based Online Lexicon (G-FOL; http://coerll.utexas.edu/frames/) currently contains nine frames with contrastive lexical entries for about 350 LUs. For details see Boas and Dux (2013) and Boas et al. (2016).

24.3.7 NLP applications

24.3.7.1 Automatic semantic role labelling

FrameNet has had a considerable impact on the field of computational linguistics. Above all, it has paved the way for the task of automatic semantic role labelling (ASRL) introduced by the seminal work of Gildea and Jurafsky (2002). Several evaluations for automatic role labelling systems have since taken place under the umbrella of the SemEval (http://aclweb.org/aclwiki/index.php?title=SemEval_Portal) series of evaluation campaigns (2004, 2007, 2010).

Because the largest amount of FrameNet data is available for English, research on ASRL systems was pioneered and is most advanced for that language, as embodied by the SEMAFOR system (Das et al. 2014). Typically, ASRL systems are language specific. This is true for SEMAFOR but also, for instance, for the Shalmaneser system, which is trained separately for German and for English (Erk and Padó 2006). Still, often such language-specific systems involve some cross-lingual transfer of knowledge from one language (typically English) to another in order to enlarge the amount of data for training the system in the target language. In a departure from this transfer approach, an attempt has recently been made to build an any-language semantic parser that uses a single model for all languages (Johannsen et al 2015).

24.3.7.2 Applications of core inference

Usually, ASRL is not an end in itself. Instead it is one of several processing steps in a pipeline that makes available to downstream processing some foundational information on “who did what to whom”. Subsequent steps then involve some kind of task-specific inferencing. The most basic inferencing makes use of information on 1) alternative ways in which FEs may be realized, 2) co-frame membership of lexical units, and 3) frame-to-frame relations. Together these types of information provide a sort of syntactic-semantic normalization of textual information that allows for the recognition of (near-)paraphrases, that is, of something like mutual entailment. Such paraphrases may involve simple syntactic variants such as active–passive or involve the more challenging recognition that two propositions involving converse predicates located in different frames such as *buy* and *sell* may be equivalent, given the frame relations and a particular filling of their FE roles. Shen and Lapata (2007) used an ASRL system in this spirit to match questions with potential answer passages in the context of the Question Answering task. However, Question Answering itself may be seen as merely one of several specific applications of an abstract textual entailment recognition task, others being Information Retrieval, Information Extraction, and Text Summarization. (Burchardt 2008) used FrameNet in his thesis on entailment recognition. Ruppenhofer and Pinkal (2011) experimented with merging frames with particular types of frame relations for the purpose of improved entailment recognition.

24.3.7.3 Applications of extended inference

Beyond the use of FrameNet for recognizing mutual entailment, research has begun to explore ways in which the resource can be used for several types of extended inference. We discuss these in what follows.

- **Overloading of semantic with opinion roles**: In recent years, FrameNet-based semantic role labelling has been used for additional inference purposes in the context of sentiment analysis. However, these new applications involve the enriching of FrameNet's current information. For instance, (Bethard et al 2004) and (Kim and Hovy 2006) used frame ASRL as a basis to assign so-called opinion roles to those sentence constituents that represent the Source or Holder of an *opinion* and its Target (or: Topic). Basically, this involves mapping frame roles to opinion roles, where appropriate, as in example 3:

  (3)

  (FFF) [Mayor Quimby Judge] **criticized** [the townspeople Evaluee] [in his annual speech Medium].

  (FFF') [Mayor Quimby Source] **criticized** [the townspeople Target] in his annual speech.

  The mappings used in the earlier cited works were created (and re-created) by specific researchers working on the sentiment analysis task. They are not currently part of FrameNet itself. Ruppenhofer and Rehbein (2012) and Ruppenhofer (2013) propose integrating this information into FrameNet, arguing that, more generally, the sentiment analysis task should be fundamentally based on semantic role labelling. They propose associating so-called opinion frames with regular frames or (groups of) LUs and mapping the roles of these two types of frames onto each other. An advantage of such an explicit representation is that multiple co-present opinions can be accurately represented. For instance, whereas the Speaker in the **Bragging** frame speaks positively about the Topic or communicates a Message with propositional content that they evaluate positively, the external reporter of the situation evoked by the frame evaluates the Speaker and/or the Topic/Message negatively.

- **Scale-related inferences**: A basic type of inference that is currently not supported by explicit information is the one that has to do with values on a scale. Knowing that *good* < *great* is needed, for instance, for the task of understanding certain dialog acts. In the exchange *A: Was it good?* B: *It was great*. B's statement can only be interpreted as an affirmative response if it is known that *good* < *great*. But in FN there is no information that directly states or lets one infer the relative ranking of the degrees of positivity conveyed by the two predicates. Of course, this same knowledge is also of great interest for sentiment analysis, where the degree of positive or negative valence expressed is an important aspect of understanding expressions of opinion.

Further, as pointed out by Fillmore and Baker (2010), the lack of information on the various types of oppositions means that FrameNet is not straightforwardly usable for paraphrase generation involving negation. Consider that in the **Compliance** frame, for example, compliance and noncompliance are binary notions, ordinarily excluding any middle ground: an act which is *compliant* with a law is *not in violation* of the law. However, when antonyms are merely contrary, as in the **Expensiveness** frame, weaker inferences are possible. For instance, if something is *cheap*, it is *not expensive*; if it is *expensive* it is *not cheap*. But the opposite inferences are not possible. Thus, something that is *not expensive*, is not necessarily *cheap*. At the moment, the differences in conceptual structure between the frames for **Compliance** and **Expensiveness** are not spelled out.

- **Event evaluativity inferences**: Reschke and Anand (2011) proposed considering the "entailments" of verbs and the attitude towards participants as the basis for calculating an external reporter's attitude towards the events. For instance, an event that results in a state of possession is evaluated positively, if both the Theme and the Recipient are evaluated positively, or if both are viewed negatively (i.e. *Schadenfreude*). The event is evaluated negatively, whenever the two participants have *different* evaluations. Reasoning like this has been shown to be useful for modelling speakers' understanding of attitudes conveyed in text. Anand and Reschke manually looked for FrameNet frames that might contain verbs bearing relevant entailments. Although for the possession entailment, some relevant frames can be retrieved by automatically exploring FrameNet's hierarchy, not all of them can. The reason is that the relevant frames need not have possession as an actual entailment. Frames in which possession post-states occur embedded under some modality are also of interest for reasoning about event evaluation. For instance, the **Offering** frame is related to **Giving** merely by the Using relationship, but the evaluative reasoning applies as if the modalized future situation were actualized, as in the **Giving** frame. If *Peter offers John the job of director*, and the speaker approves both of John and the job of director, then they should also evaluate the offering event as a whole positively. The Using relation is, however, generally too weak to be a basis for inference. Thus, a fully automatic way of extracting frames with possession entailments would miss this. Expanding FrameNet's representational capabilities to explicitly model the modalized possession would be desirable in the longer run. In the interim, an alternative approach would be to associate information about Anand and Reschke's reasoning schemes with frames and LUs, similar to what Ruppenhofer and Rehbein (2012) proposed for the fixed sentiment associated with opinion roles, and what Ruppenhofer and Brandes (2016) did for GermaNet synsets.

24.4 Looking to the future

In terms of its analytical scope, FrameNet is moving towards capturing the semantics of grammar. This means dealing both with general and abstract grammatical phenomena (negation, tense, aspect) as well as with phraseology (constructions and syntactic idioms). This expansion will enable the integration of lexical meanings and grammatical meanings and facilitate a more complete account of text understanding. Efforts to build constructional resources tied to frame semantics are underway for English (Fillmore et al. 2012), Swedish (Lyngfelt *et al.* 2012; Gruzitis et al. 2015) and Japanese (Ohara 2015) and Portuguese.[6] Throughout the FrameNet project, research on Construction Grammar has continued and been presupposed in decisions about how to represent frames. Likewise, Fillmore regarded LUs as minimal lexical constructions, and posited that many non-lexical constructions also evoke semantic frames, which serve as a partial representation of their semantics.

There are also plans for increased cooperation and coordination among the FrameNets for different languages, including (1) the creation of a database connecting the work on all of the languages, with alignments at both frame and LU levels; (2) efforts to expand existing FrameNets by semi-automatic cross-linguistic projection; and (3) closer collaboration among existing FrameNets and systematic assistance for fledgling FrameNets in new languages. This project will also move beyond the current XML distribution of the FrameNet data, making the data available in ways that are more compatible with the needs of NLP research.

24.5 Conclusion

FrameNet is a rich, yet still evolving lexical resource that seeks to apply the theory of Frame Semantics to the analysis of the English lexicon. FrameNet has attracted interest from diverse research communities, ranging from theoretical linguistics to natural language processing applications. In its original domain of lexicography, FrameNet has inspired many sister projects in other languages and is seeing new specialized uses, such as foreign language instruction. In the field of natural language processing FrameNet is used as a resource in a great number of applications, notably semantic role labelling. In addition, many efforts have been and are being undertaken to (semi-)automatically expand frame semantic resources or construct them from scratch. Such efforts, on the one hand, underscore the keen interest in using frame semantic annotations and lexicon resources, but on the other reflect the challenges and limitations of pursuing large coverage only through the manual work of linguistic experts. The next phase of frame semantic development will therefore seek to speed up the progress of frame semantic resources by combining the work of human experts with that of laypeople, while also attempting to better leverage the growing analytical powers of automatic systems.

Related topics

Lexicography and applied linguistics; lexicography and natural language processing; electronic dictionaries; user participation in the era of the Internet; information retrieval for lexicographic purposes; Wordnik; English lexicography in the Internet era.

Further reading

Boas, H. C. (ed.) (2009) *Multilingual FrameNets: Methods and Applications*. Berlin/New York: Mouton de Gruyter.

This edited volume consists of 11 chapters describing and explaining how the Berkeley FrameNet frames for English are employed to build FrameNets for other languages such as Spanish, Japanese, and German.

Fillmore, C. J. and C. F. Baker. (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.

This chapter provides a detailed overview of the FrameNet project, including its workflow, database, the structure of lexical entries, and theoretical considerations.

Fillmore, C. J., R. Lee-Goldman and R. Rhomieux. (2012). "The FrameNet Constructicon", in H. C. Boas and I. Sag (eds.), *Sign-Based Construction Grammar*. Stanford: CSLI Publications, pp. 309–372.

This chapter discusses the expansion of the FrameNet lexical database and methodology to include entries for grammatical constructions using the same format.

Ruppenhofer, J., M. Ellsworth, M. Petruck, C. Johnson, and J. Scheffczyk. (2010). *FrameNet II: Extended Theory and Practice*. http://framenet.icsi.berkeley.edu.

This technical report provides details about the policies and workflow of the FrameNet project, including frame creation, annotation of corpus examples, and creation of lexical entries in the FrameNet database.

Notes

1 See https://framenet.icsi.berkeley.edu/fndrupal/framenet_users for a list of users.
2 We discuss frames here in terms of "events" and their "participants", which are the clearest case, but semantic frames can also represent relations (such as the Similarity frame discussed in Section 24.1), states (e.g. the Dead_or_alive frame), and entities (e.g. the Clothing frame), all of which have related entities and properties which are treated as frame elements.
3 For information on the architecture and physical implementation of the FrameNet database, see Baker, Fillmore and Cronin 2003.
4 See at (http://framenet.icsi.berkeley.edu)
5 For information on FrameNet's machine-readable data distribution in XML format, please consult the FrameNet website and the Release Notes in the data distribution.
6 See at (www.ufjf.br/framenetbr-eng/projects/frames-and-constructions (2016/6/22).

References

Atkins, B.T.S., C. J. Fillmore and C. R. Johnson. (2003) "Lexicographic Relevance: Selecting Information From Corpus Evidence" ', *International Journal of Lexicography* 16(3): 251–280.

Baker, C. F., C. J. Fillmore and B. Cronin. (2003) "The Structure of the FrameNet Database", *International Journal of Lexicography* 16(3): 281–296.

Bethard, S., H. Yu, A. Thornton, V. Hatzivassiloglou and D. Jurafsky (2004) "Automatic Extraction of Opinion Propositions and Their Holders", in *Proceedings of the AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications*, pp. 20–27.

Boas, H. C. (2003) *A Constructional Approach to Resultatives*. Stanford: CSLI Publications.

Boas, H. C. (2005) "Semantic Frames as Interlingual Representations for Multilingual Lexical Databases", *International Journal of Lexicography* 18(4): 445–478.

Boas, H. C. (In press) "Computational Resources: FrameNet and Constructicon", in B. Dancyger (ed.), *The Cambridge Handbook of Cognitive Linguistics*. Cambridge: Cambridge University Press. AuQ5

Boas, H. C., and R. Dux (2013) "Semantic Frames for Foreign Language Education: Towards a German Frame-Based Dictionary", *Veredas* 17(1): 82–100.

Boas, H. C., R. Dux and A. Ziem. (2016) "Frames and Constructions in an Online Learner's Dictionary of German", in S. De Knop and G. Guilquin (eds.), *Constructional Approaches to Language Pedagogy*. Berlin: Mouton de Gruyter, pp. 303–326.

Borin, L., D. Dannélls, M. Forsberg, M. T. Gronostaj and D. Kokkinakis (2010) "The Past Meets the Present in the Swedish FrameNet++", Extended version of EURALEX 2010 presentation. https://svn.spraakdata.gu.se/sb/fnplusplus/pub/SweFN_Euralex_extended.pdf.

Bouveret, M. (2012) "GIVE Frames and Constructions in French", in M. Bouveret and D. Legallois (eds.), *Constructions in French*. Amsterdam/Philadelphia: Benjamins, pp. 99–126.

Burchardt, A. (2008) *Modeling Textual Entailment With Role-semantic Information*, PhD thesis, Universität des Saarlandes.

Burchardt, A., K. Erk, A. Frank, A. Kowalski, S. Padó and M. Pinkal (2009) "Using FrameNet for the Semantic Analysis of German: Annotation, Representation, and Automation", in H. C. Boas (ed.), *Multilingual FrameNets: Methods and Applications*. Berlin/New York: Mouton de Gruyter, pp. 209–244.

Busse, D. (2012) *Frame-Semantik: Ein Kompendium*. Berlin/New York: Mouton de Gruyter.

Candito, M., P. Amsili, L. Barque, F. Benamara, G. de Chalendar, M. Djemaa, P. Haas, R. Huyghe, Y. Mathieu, P. Muller, B. Sagot and L. Vieu (2014) "Developing a French FrameNet: Methodology and First Results", *Proceedings of LREC 2014,* pp. 1372–1379.

Chang, N., P. Paritosh, D. Huynh and C. F. Baker (2015) "Scaling Semantic Frame Annotation", in *Proceedings of The 9th Linguistic Annotation Workshop*, pp. 1–10.

Das, D., D. Chen, A.F.T. Martins, N. Schneider and N. A. Smith (2014) "Frame-Semantic Parsing", *Computational Linguistics* 40(1): 9–56.

Erk, K. and S. Pado (2006) "Shalmaneser – A Flexible Toolbox for Semantic Role Assignment", in *Proceedings of LREC 2006*, pp. 527–532.

Fillmore, C. J. (1968) "The Case for Case", in E. Bach and R. Harms (eds.), *Universals in Linguistic Theor*y. New York: Rinehart & Winston, pp. 1–90.

Fillmore, C. J. (1975) "An Alternative to Checklist Theories of Meaning", in *Proceedings of the First Annual Meeting of the Berkeley Linguistic Society (BLS)*, pp. 123–131.

Fillmore, C. J. (1976) "Frame Semantics and the Nature of Language", *Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech*, 280, pp. 20–32.

Fillmore, C. J. (1977) "Scenes-and-Frames Semantics", in A. Zampolli (ed.), *Fundamental Studies in Computer Science*. Dordrecht: North Holland Publishing, pp. 55–88.

Fillmore, C. J. (1978) "On the Organization of Semantic Information in the Lexicon", *Papers From the Parasession on the Lexicon, Chicago Linguistic Society*, pp. 148–173.

Fillmore, C. J. (1979) "Innocence: A Second Idealization for Linguistics", in *Proceedings of the Fifth Annual Meeting of the Berkeley Linguistic Society (BLS)*, pp. 63–76.

Fillmore, C. J. (1982) "Frame Semantics", in Linguistic Society of Korea (ed.), *Linguistics in the Morning Calm*. Seoul: Hanshin, pp. 111–38.

Fillmore, C. J. (1985) "Frames and the Semantics of Understanding", *Quaderni di Semantica* 6: 222–254.

Fillmore, C. J. (2007) "Valency Issues in FrameNet", in T. Herbst and K. Götz-Vetteler (eds.), *Valency: Theoretical, Descriptive, and Cognitive Issues*. Berlin/New York: Mouton de Gruyter, pp. 129–160.

Fillmore, C. J. and B.T.S. Atkins. (1992) "Toward a Frame-Based Lexicon: The Semantics of RISK and Its Neighbors", in A. Lehrer and E. Kittay (eds.), *Frames, Fields and Contrasts: New Essays in Semantic and Lexical Organization*. Hillsdale: Erlbaum, pp. 75–102.

Fillmore, C. J. and B.T.S. Atkins. (2000) "Describing Polysemy: The Case of Crawl", in Y. Ravin and C. Leacock (eds.), *Polysemy: Linguistic and Computational Approaches*. Oxford: Oxford University Press, pp. 91–110.

Fillmore, C. J. and C. F. Baker. (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., C. Johnson and M. Petruck. (2003) "Background to FrameNet", *International Journal of Lexicography* 16: 235–251.

Fillmore, C. J., R. Lee-Goldman and R. Rhomieux (2012) "The FrameNet Constructicon", in H. C. Boas and I. Sag (eds.), *Sign-Based Construction Grammar*. Stanford: CSLI Publications, pp. 309–372.

Fontenelle, T. (1997) "Using a Bilingual Dictionary to Create Semantic Networks", *International Journal of Lexicography* 10(4): 275–303.

Fossati, M., C. Giuliano and S. Tonelli (2013) "Outsourcing FrameNet to the Crowd", in *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pp. 742–747.

Gildea, D. and D. Jurafsky (2002) "Automatic Labeling of Semantic Roles", *Computational Linguistics* 28(3): 245–288.

Gruzitis, N., D. Dannélls, B. Lyngfelt and A. Ranta (2015). "Formalising the Swedish Constructicon in Grammatical Framework", in *Proceedings of ACL-IJCNLP 2015*, pp. 49–56.

Hasegawa, Y., R. Lee-Goldman, K. Ohara, S. Fuji and C. Fillmore (2010) "On Expressing Measurement and Comparison in English and Japanese", in H. C. Boas (ed.), *Contrastive Studies in Construction Grammar*. Amsterdam/Philadelphia: Benjamins, pp. 169–200.

Heid, U. (1996) "Creating a Multilingual Data Collection for Bilingual Lexicography From Parallel Monolingual Lexicons', in *Proceedings of the VIIth EURALEX International Congress*, pp. 573–559.

Hong, J. and C. F. Baker (2011) "How Good Is the Crowd at "Real" WSD?", in *Proceedings of the 5th Linguistic Annotation Workshop*, pp. 30–37.

Johannsen, A., H. Martinez Alonso and A. Soegaard (2015) "Any-Language Frame-Semantic Parsing", in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pp. 2062–2066.

Kim, S.-M. and E. Hovy (2006) "Extracting Opinions, Opinion Holders, and Topics Expressed in Online News Media Text", *Proceedings of the Workshop on Sentiment and Subjectivity in Text*, pp. 1–8.

Kipper, K., A. Korhonen, N. Ryant, and M. Palmer (2008) "A Large-Scale Classification of English Verbs", *Language Resources and Evaluation Journal* 42(1): 21–40.

Leino, J. (2010) "Results, Cases, and Constructions: Argument Structure Constructions in English and Finnish", in H. C. Boas (ed.), *Contrastive Studies in Construction Grammar*. Amsterdam/Philadelphia: Benjamins, pp. 103–136.

Levin, B. and M. Rappaport-Hovav. (2005) *Argument Realization*. Cambridge: Cambridge University Press.

Liu, K. (2011) "Research on Chinese FrameNet Construction and Application Technologies", *Journal of Chinese Information Processing* 25(6): 46–52.

Lyngfelt, B., Borin, L., Forsberg, M., Prentice, J., Rydstedt, R., SKöldberg, E., and Tingsell, S. (2012) "Adding a Constructicon to the Swedish Resource Network of Språkbanken", *Proceedings of KONVENS 2012*, pp. 452–461.

Ohara, K. (2009) "Frame-Based Contrastive Lexical Semantics in Japanese FrameNet: The Case of Risk and Kakeru", in H. C. Boas (Ed.), *Multilingual FrameNets: Methods and Applications*. Berlin/New York: Mouton de Gruyter, pp. 163–182.

Ohara, K. (2015) "Constructicon Building as a Practical Implementation of Construction Grammar and Frame Semantics: Japanese FrameNet", Poster at the 13th International Cognitive Linguistics Conference (ICLC 13) Northumbria University at Newcastle. www.northumbria.ac.uk/media/6836834/poster-abstracts-9-july.pdf.

Ohara, K. H., S. Fujii, H. Saito, S. Ishizaki, T. Ohori and R. Suzuki (2003) "The Japanese FrameNet Project: A Preliminary Report", in *Proceedings of Pacific Association for Computational Linguistics (PACLING'03),* pp. 249–254.

Petruck, M. (2009) "Typological Considerations in Constructing a Hebrew FrameNet", in H. C. Boas (Ed.), *Multilingual FrameNets: Methods and Applications*. Berlin/New York: Mouton de Gruyter, pp. 183–208.

Petruck, M. and H. C. Boas. (2003) "All in a Day's Week", *Proceedings of CIL 17*. CD-ROM. Prague: Matfyzpress.

Petruck, M., C. J. Fillmore, C. F. Baker, M. Ellsworth and J. Ruppenhofer (2004) "Reframing FrameNet Data", in *Proceedings of The 11th EURALEX International Congress*, pp. 405–416.

Ruppenhofer, J. (2013) "Anchoring Sentiment Analysis in Frame Semantics", *Veredas* 17(1): 66–81.

Ruppenhofer, J. and M. Pinkal. (2011) "Folding FrameNet – Unfolding Its Potential?", *International Journal of Computational Linguistics and Applications* 2(1–2): 25–40.

Ruppenhofer, J., M. Ellsworth, M. Petruck, C. Johnson and J. Scheffczyk (2010) *FrameNet II: Extended Theory and Practice*. http://framenet.icsi.berkeley.edu.

Ruppenhofer, J. and J. Brandes (2016) "Effect Functors for Opinion Inference", in *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, pp. 2879–2887.

Ruppenhofer, J. and I. Rehbein (2012) "Semantic Frames as an Anchor Representation for Sentiment Analysis", in *Proceedings of the 3rd Workshop in Computational Approaches to Subjectivity and Sentiment Analysis*, pp. 104–109.

Salomão, M., T. Timponi Torrent and T. Fernandes Sampaio (2013) "A linguística de corpus encontra a linguística computacional: Notícias do projeto FrameNet Brasil", *Cadernos de Estudos Linguísticos* 55(1): 7–34.

Sato, H. (2008) "New Functions of FrameSQL for Multilingual FrameNets", in *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)*, pp. 758–762.

Shen, D. and M. Lapata (2007) "Using Semantic Roles to Improve Question Answering", in *Proceedings of the Conference on Empirical Methods in Natural Language Processing and on Computational Natural Language Learning*, pp. 12–21.

Subirats, C. (2009) "Spanish FrameNet: A Frame-Semantic Analysis of the Spanish Lexicon", in H. C. Boas (ed.), *Multilingual FrameNets: Methods and Applications*. Berlin/New York: Mouton de Gruyter, pp. 135–162.

Subirats, C. and M. Petruck. (2003) Surprise: Spanish FrameNet! *Proceedings of CIL 17*, CD-ROM, Matfyzpress, Prague.

Timyam, N. and B. Bergen. (2010) "A Contrastive Study of the Caused-Motion and Ditransitive Construction in English and Thai: Semantic and Pragmatic Constraints', in H. C. Boas (eds.), *Contrastive Studies in Construction Grammar*. Amsterdam/Philadelphia: Benjamins, pp. 137–168.

Willems, D. (2012) "Verb Typology: Between Construction and Lexicon", in M. Bouveret and D. Legallois (eds.), *Constructions in French*. Amstedam/Philadelphia: Benjamins, pp. 23–48.

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