Reframing FrameNet Data

Miriam R. L. Petruck, Charles J. Fillmore, Collin F. Baker, Michael Ellsworth and Josef Ruppenhofer
International Computer Science Institute
1947 Center Street, Suite 600
Berkeley, CA 94704-1198, USA
{miriamp, fillmore, collinb, infinity, josef}@icsi.berkeley.edu
http://framenet.icsi.berkeley.edu/~framenet

Abstract

The Berkeley FrameNet Project (http://www.icsi.berkeley.edu/~framenet) is building an on-line lexical resource for contemporary English. The database provides information about the semantic and syntactic combinatorial possibilities (**valences**) of each item analyzed. This paper describes the conceptual basis for what has been called **reframing** of data in the FrameNet database and exemplifies two new frame-to-frame relations, **Causative_of** and **Inchoative_of,** the implementation of which came about as a result of reanalysis of certain frames and lexical units. The new relations are characterized with respect to a triple of frames involving the notion of attaching, and entering them into the database is demonstrated using the Frame Relations Editor. The two relations allow FrameNet to make frame-wise distinctions that capture fairly systematic semantic relationships across sets of lexical units. While the Inheritance and Subframe relations are of particular interest to the NLP research community, Causative_of and Inchoative_of may be more relevant to lexicography.

1. Introduction

The Berkeley FrameNet Project (http://www.icsi.berkeley.edu/~framenet) (Johnson, et al., 2003, Fillmore, et al., 2001) is building an on-line lexical resource for contemporary English. The database provides information about the semantic and syntactic combinatorial possibilities (**valences**) of each item analyzed. The findings are derived automatically from the manual annotation of carefully selected sentences culled from corpora[1], and can be browsed and queried through the Internet. The theoretical basis and descriptive model of the project is **Frame Semantics**, which offers an approach to the characterization and analysis of word meaning in terms of the **semantic frame**.

Users of the FrameNet database will find that it serves as both a dictionary and a thesaurus. As a dictionary, for each **lexical unit (LU)** (lemma in given sense) it provides the name of the frame that houses it, a (dictionary or FrameNet developed) definition, a valence description that summarizes the combinatorial possibilities of frame elements occurring with that LU, and sets of annotated sentences that exemplify the various syntactic patterns discovered in the corpus. The thesaurus-like nature of the FrameNet database manifests in the way that groups of lexical units are connected to frames which are also connected to other frames through various frame-to-frame relations.

The FrameNet database can be distinguished from ordinary (print) dictionaries and thesauri, as well as other lexical resources (e.g. WordNet, Fellbaum, 1998) in a number of ways.

Along with definitions, valence descriptions, and annotated example sentences, the FrameNet database provides highly specific frames and semantic roles (frame elements), as well as extremely detailed information on the various syntactic realizations of semantic roles for each lexical unit. FrameNet also includes information about relations between frames that indicate semantic relationships between collections of concepts, for example, Inheritance and Subframe.

Recently, FrameNet added two more frame-to-frame relations to its repertoire. These are the **Causative_of** relation and the **Inchoative_of** relation, the implementation of which came about as a result of reanalysis of certain frames and lexical units. This paper describes the conceptual basis for what has been called **reframing** of data in the FrameNet database, exemplifies the two new frame-to-frame relations with respect to a triple of frames involving the notion of attaching, and demonstrates how the Frame Relations Editor is used to enter the relations into the database.

The two relations allow FrameNet to make frame-wise distinctions that capture fairly systematic semantic relationships across sets of lexical units. While the Inheritance and Subframe relations are of particular interest to the NLP research community, Causative_of and Inchoative_of may be more relevant to the field of lexicography.

2. Frame Semantics

At the heart of Frame Semantics (Fillmore, 1977, 1982; Petruck, 1996) is the semantic frame, a structured schematic representation of a situation, object, or event that provides the background and motivation for the existence and everyday use of words in a language.

In Frame Semantics, a linguistic unit, here, a word (in just one of its senses) **evokes** a frame. That frame is the structure of knowledge required for the understanding and appropriate use of lexical items or phrases. The evoked frame can be very simple, for instance, Being_wet (e.g. *wet*.a, *soaked*.a, *drenched*.a), which describes a state of affairs, or it can characterize a more complex event (or set of related events), for example, Education. The frame structures the background information for words that highlight different phases, participants and props. For each frame, there is a set of frame elements (FEs) – i.e. frame-specific semantic roles, the linguistic realization of which is also recorded in terms of grammatical function and phrase type. A FrameNet lexical entry identifies the frame that underlies a single sense and lists the ways in which the FEs are realized in structures headed by the word.

For example, the word *tip* evokes a scene in which someone has paid for a service received, (typically) is satisfied with the service, and rewards monetarily the person who has provided the service. The word highlights the monetary reward given to the person who has provided the service. To illustrate, in the sentence *Fred gave the waiter a large tip*, we understand that Fred paid for the waiter's service and that the reward is understood against the background of assumptions and practices of the frame.

2.1 An Example Frame: Attaching[2]

The Attaching frame covers two situations: a scene in which somebody causes one thing to be physically connected to something else; or a scene in which somebody causes two things to be connected to each other. In the first, the frame includes an AGENT who attaches an ITEM to a GOAL by manipulating a CONNECTOR, creating an asymmetric relationship between the ITEM and the GOAL. In the second, the AGENT attaches two ITEMS to each other, where each serves as a GOAL for the other, creating a symmetric relationship between the two ITEMS. In both cases, the CONNECTOR remains to bind the two entities (either ITEM and GOAL, or two ITEMS), without creating a new entity. For example, in the sentence *They attach their canopies by a little silk pad to the plant's stem, They* fills the role of AGENT, *their canopies* instantiates the FE ITEM, *by a little silk thread* is the CONNECTOR, and *to the plant's stem* is the GOAL. Other words defined in terms of this frame are *affix*.v, *anchor*.v, *attachment*.n, *bind*.v, *cinch*.v, *detach*.v, *fetter*.v, *fuse*.v, *hook*.v, *lash*.v, *paste*.v, *pin*.v, *rivet*.v, *sew*.v, *tie*.v, *tying*.n, *untie*.v, and *weld*.v.

Syntactically, the AGENT can be expressed by a NP, as in the above example, or it may be constructionally null instantiated (e.g. *Attach the unit to the wall*, where the AGENT is not expressed because the sentence is an imperative). The ITEM is expressed as an NP, and the GOAL is expressed as a PP. The GOAL may be omitted under definite null instantiation, that is, it is understood from the discourse. In *Chuck secured the tarp with a rope*, we understand that the tarp was secured to some GOAL. Note that in this frame, if ITEM and GOAL are expressed, then ITEMS cannot be, since when two ITEMS are instantiated one serves as the GOAL for the other, as in *The robber tied Fred's feet together*, where *Fred's feet* are the realization of the FE ITEMS. Finally, the CONNECTOR is expressed as a PP, shown by the *with*-PP in *The robber tied Fred's feet together with a bungee cord*.

3. Frame-to-Frame Relations in the FrameNet Database

FrameNet represents semantic relations in a number of ways. First, there is the implied semantic similarity of grouping words together in a frame. In addition, FrameNet uses a set of semantic types that can be applied to LUs, frames, and FEs (Johnson, et al., 2003). Of particular relevance to the present work is the set of frame-to-frame relations, the more important of which we briefly discuss here.

3.1 Inheritance and Subframes

Frame inheritance is a relationship in which a child frame is a more specific elaboration of its parent frame. In such cases, all of the frame elements, other frame relations and (semantic) characteristics of the parent have equally or more specific correspondents in the child frame. Consider, for example, the Evading frame, evoked by *evade*.v, *elude*.v, and *sidestep*.v, which inherits from a more general Avoiding frame, with all of the FEs in Avoiding having correspondents in Evading. Evading is a more specific instance of Avoiding, and, for example, the EVADER role is a more specific instance of the AGENT, and PURSUER is more specific than UNDESIRABLE_SITUATION. Subframes[3] is a relationship which is used to characterize the different sequential parts of a complex event in terms of the

sequences of states of affairs and transitions between them, each of which can itself be separately described as a frame. For instance, the complex Sleep_scenario frame consists of several simpler frames, including Fall_asleep (e.g. *doze_off*.v, *fall_asleep*.v) Sleep (e.g. *doze*.v, *sleep*.v) and Waking_up (*wake*.v, *get_up*.v). The Inheritance and Subframe relations are of particular interest to the NLP research community, in part because they can make possible inferencing about the prerequisites, causes, and results of events. Subframes is also likely to be of interest to dictionary users and language learners since much coherent discourse stays on a topic for a while. To illustrate, if I tell you when I went to sleep, I might also tell you when I got up.

4. Reframing

Much of the day-to-day work of the project involves developing frame descriptions, an important part of which is determining the boundaries of individual frames. In the course of time, we have developed a better understanding of how to group words into frames. As such, we have reanalyzed many frames, a process which we call **reframing**. For the most part, the new analysis is more fine-grained than the old, and that requires first defining new destination frames, each with its own set of FEs and LUs. Generally, we reassign annotation from the LUs of a source frame to the LUs of a destination frame, or (in the easy cases) we reassign an entire LU (i.e. with all of its annotated sentences) from the source to the destination.

To illustrate, an early analysis of a set of words that concern noise were all grouped together in one frame, Make_noise (formerly called Noise). The more fine-grained analysis required defining several additional frames, such as Cause_to_make_noise, Sound_movement, and Sounds. Consider the lemma *blast*, as exemplified in the following set of sentences.

> (1) The foreman **blasted** the siren exactly at noon.
> (2) The siren **blasted** exactly at noon.
> (3) Loud rock music from the radio **blasted** through the window.
> (4) The stock car **blasted** down the track.
> (5) The **blast** of the siren woke everyone up.

In (1), an AGENT acts in a way that causes the siren to emit a blast, while in (2) the siren is construed as a sufficiently complex entity so as to emit the noise by itself. In (3), the blasting of the music is seen as moving along a path, while in (4) the verb characterizes the motion of the car. In (5) the LU names that which is perceived. After reframing, each of these uses is handled in a different frame, with the LU *blast* in (1) in the Cause_to_make_noise frame, and (2) in the Make_noise frame, the former being a causative. The LU in (3) is in a Motion_noise frame, where the noise changes location, while in (4) the LU is defined in terms of a Sound_movement frame. The noun in (5) is defined in terms of the Sounds frame, which concerns the percepts that vibrations travelling through a medium, (usually, air) produce in hearing organs.

The development of software for reframing, made it easier to reassign annotation from the LUs of the original Make_noise frame to the LUs of the various destination frames. Figure 1 shows the Frame Relation Editor in the FrameNet Desktop software with mappings of FEs between the Make_noise frame and some of the more recently created noise-related frames. The picture given here constitutes a record of the reframings that were executed by FrameNet lexicographers. Toward the top of the picture, **Source** and **Target** indicate which of the displayed frames served as the starting point for the mappings, and which were the end point(s). Note that nothing further is claimed about the frame-to-frame relations that may exist between the source frame and any one of the target frames.
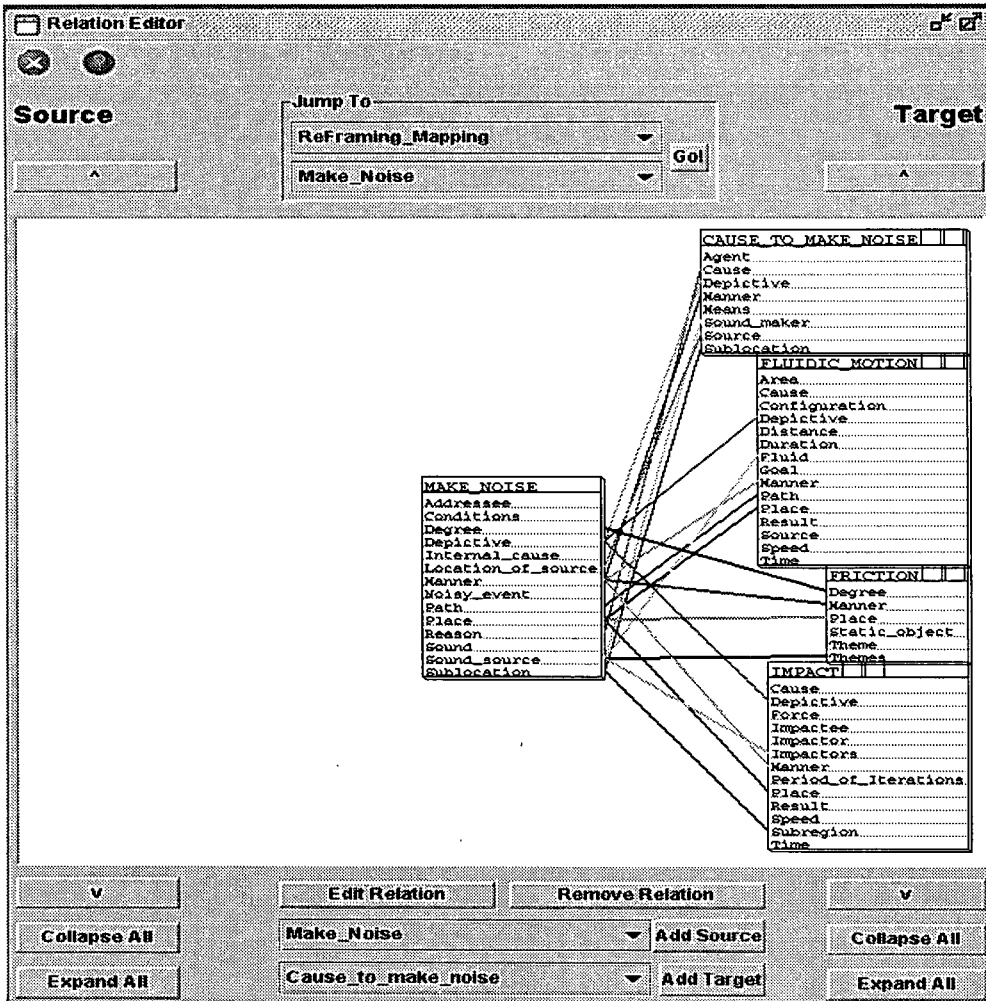

Figure 1: Reframing Mappings

4.1 Causative_of and Inchoative_of Relations

Deciding to systematically distinguish among statives, inchoatives, and causatives, so that each category of LU is in a separate frame necessitated quite a few reframings. One set of frames that illustrate these relations are Attaching, Inchoative_attaching, and Being_attached, as in the following examples.

    (6) Insects **attach** extraneous objects to themselves.   (Attaching)
    (7) Remoras **attached** to the dolphins.
    (Inchoative_attaching)
    (8) The baby remains **attached** by the cord to the mother. (Being_attached)

Recall that the Attaching frame characterizes a scene in which an AGENT attaches an ITEM to a GOAL with a CONNECTOR, or attaches two ITEMS together. The Inchoative_attaching frame captures a scene in which an ITEM comes to be attached to a GOAL with a CONNECTOR, or two ITEMS come to be attached to each other. Lexical units in this frame include the verbs *agglutinate, attach, bind, fasten, moor,* and *take hold.* The Being_attached frame describes a scene in which an ITEM is attached by a CONNECTOR to a GOAL, or two ITEMS are attached to each other. Some the LUs defined in Being_attached are the adjectives *attached, bound, fastened, lashed, moored,* and *tied.*
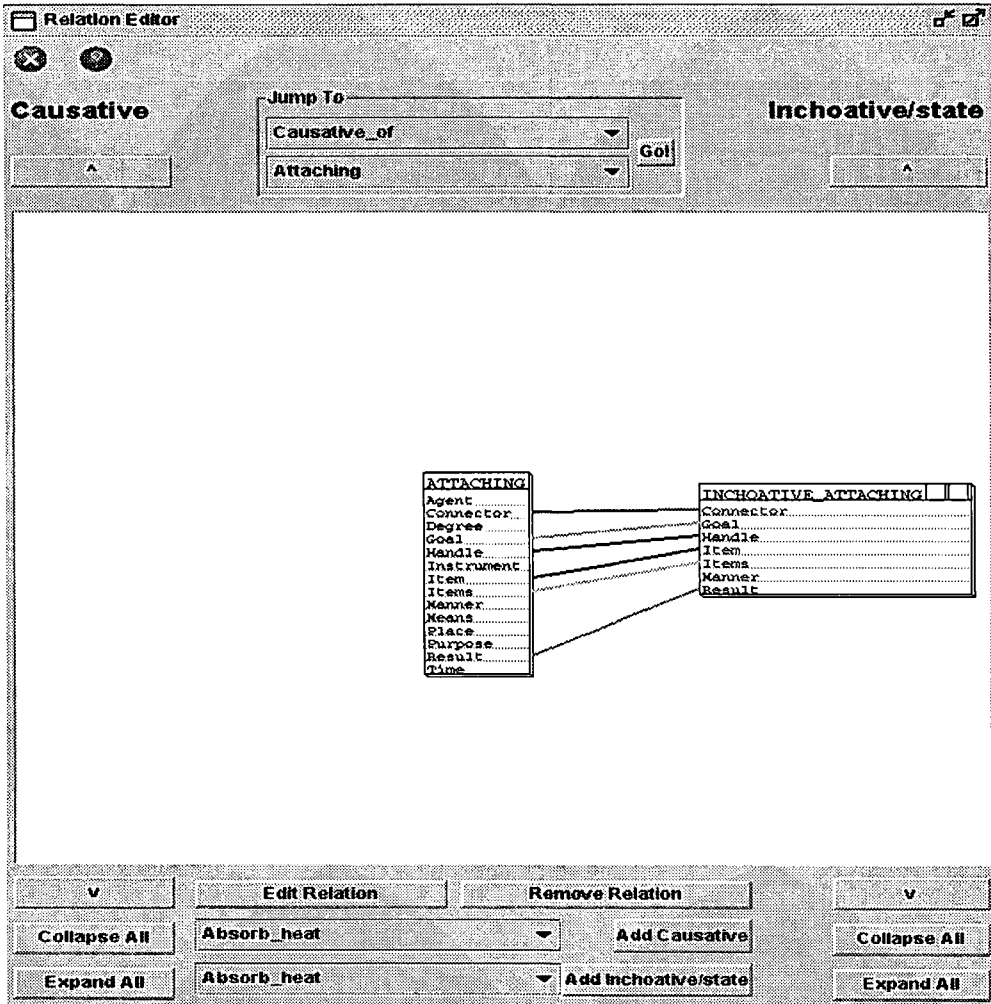
Figure 2: Causative_of Relation: Attaching and Inchoative_attaching

Figure 2 shows the Relation Editor with mappings between the FEs in Attaching and Inchoative_attaching. Such frame-to-frame relationships are added one at a time.

The Causative_of relation holds between Attaching and Inchoative_attaching, and the relationship is uni-directional.[4] This is so because some inchoative attaching events have a conceptually salient cause, while others do not. Compare the following two sentences, where in the first the cause of the attaching is conceptually salient.

(9) The ribbon is **tied** to the shirt through the hidden button hole (by the tailor).
(10) The fallen leaves got **stuck** to the bottom of the pool.

Similarly, the Inchoative_of relationship holds between Inchoative_attaching and Being_attached, and is also uni-directional. The mechanics of adding the Inchoative_of relationship is identical to that for adding the Causative_of relationship. The implementation of these relationships is on-going, although the FrameNet database already includes some triples of frames characterized by the causative-inchoative-stative distinction, as shown in Table 1. The phenomenon is quite widespread across the vocabulary, and manifests in the systematic relationship among certain transitive verbs, their intransitive counterparts, and the corresponding (stative) adjectives.[5]

Frame	LU	Example
Cause_change_of_consistency	thicken.vt	Thicken the soup with cornstarch.
Change_of_consistency	thicken.vi	Stir well as the soup thickens.
Consistency	thick.a	The soup is thick.
Cause_change_color	redden.vt	June reddened her face with rouge.
Change_of_color	redden.vi	Fred's face reddened with shame.
Color	red.a	Fred's face is red.
Killing	kill.vt	Someone killed Smedlap.
Death	die.vi	Smedlap died.
Dead_or_alive	dead.a	Smedlap is dead.
Cause_change_of_temperature	cool.vt	Jo cooled the soup in an ice bath.
Inchoative_change_of_temperature	cool.vi	The soup cooled near the window.
Temperature	cold.a	The soup is cold.
Cause_change_of_phase	liquefy.vt	Above a certain temperature, it is not possible to liquefy a gas.
Change_of_phase	liquefy.vi	Natural gas liquefies at -167°C
Phase	liquid.a	The paint is liquid and shiny.

Table 1: Triples of Frames Exhibiting Causative-Inchoative-Stative Distinction

For each triple, the Causative_of relation holds between the first and the second frame in the list; and the Inchoative_of relation holds between the second and the third frame in the list. To illustrate, the Causative_of relation holds between Killing and Death and Inchoative_of holds between Death and Dead_or_alive. The individual frames are characterized as causative, inchoative, or stative frames, respectively.[6]

5. FrameNet and Lexicography

Lexicographers writing a new entry or revising an existing one can exploit the information in the FrameNet database, some of which resulted from reanalysis and was implemented via the process of reframing.[7] First, for the kind of triples listed in Table 1, FrameNet has three separate lexical entries: the transitive verb, the intransitive verb, and the adjective, each

defined against the background of its own frame and set of frame elements. Next, for each lexical entry, there is a valence description in table format giving the semantic and syntactic combinatorial possibilities of the word, along with access to a set of annotated examples showing how each semantic role is realized. The valence tables for *fasten*.v in Attaching is given in Figure 2, for *fasten*.v in Inchoative_attaching in Figure 3, and for *fastened*.a in Being_attached in Figure 4.

Number Annotated	Patterns				
11 exx TOTAL	Agent	Connector	Goal	Item	
1 exx	-- --	-- --	PP[on] Comp	NP Ext	
1 exx	-- --	-- --	PP[to] Comp	NP Ext	
2 exx	-- --	NP Ext	PP[to] Comp	NP Ext	
1 exx	-- --	PP[with] Comp	PP[to] Comp	NP Ext	
2 exx	NP Ext	-- --	PP[to] Comp	NP Obj	
2 exx	NP Ext	NP Obj	PP[round] Comp	NP Obj	
2 exx	NP Ext	NP Obj	PP[to] Comp	NP Obj	
1 exx TOTAL	Agent	Connector	Goal	Item	Purpose
1 exx	-- --	-- --	AVP Comp	NP Ext	VPto Comp
1 exx TOTAL	Agent	Connector	Items	Means	Result
1 exx	-- --	-- --	NP Ext	AVP Comp	AVP Comp

Figure 2: Valence Table for *fasten*.v.tr. in Attaching

Notice that the transitive verb (Figure 2) differs from the intransitive one (Figure 3) in that the former has the FE AGENT (plus other related frame elements, MEANS and PURPOSE) in its valence description.

Number Annotated	Patterns		
2 exx TOTAL		Goal	Item
1 exx	NP Ext	PP[on] Comp	NP Ext
1 exx	NP Ext	PP[round] Comp	NP Ext

Figure 3: Valence Table for *fasten*.v.intr. in Inchoative_attaching

The presence of the FE AGENT also differentiates the transitive verb from the stative adjective (Figure 4). Distinguishing the intransitive verb (Figure 3) and the stative adjective (Figure 4) requires further explanation, and can be deduced, in part, from current FrameNet data. While the Inchoative_attaching frame characterizes an event, Being_attached describes a state of affairs. Thus, for example, we expect adverbs that denote speed, to occur with Inchoative_attaching verbs, but not with Being_attached adjectives. The only adverbs that occur in our data with *fastened*.a denote the strength of the connection (*strongly*, *tightly*); these are labeled with the broad category FE MANNER (Figure 4).[8]

Number Annotated	Patterns				
1 exx TOTAL	Connector	Dependent state	Goal	Item	Manner
1 exx	--	PP[on]	PP[to]	NP	AVP
	--	Comp	Comp	Ext	Comp
5 exx TOTAL	Connector	Goal	Item		
1 exx	--	PP[between]	NP		
	--	Comp	Ext		
1 exx	--	PP[on]	NP		
	--	Comp	Ext		
1 exx	PP[by]	PP[to]	NP		
	Comp	Comp	Ext		
2 exx	PP[with]	PP[to]	NP		
	Comp	Comp	Ext		
1 exx TOTAL	Dependent state	Item			
1 exx	AVP	NP			
	Comp	Ext			
1 exx TOTAL	Goal	Handle	Item		
1 exx	PP[to]	PP[at]	NP		
	Comp	Comp	Ext		
1 exx TOTAL	Goal	Item			
1 exx	PP[to]	NP			
	Comp	Ext			
1 exx TOTAL	Goal	Item	Manner		
1 exx	PP[to]	NP	AVP		
	Comp	Ext	Comp		

Figure 4: Valence Table for *fastened*.a in Being_attached

Compare the rich information in FrameNet with the entry in one ordinary print dictionary for *fasten*.v. The following is from the fourth edition of the American Heritage Dictionary.

> **fasten** *v.* –tened, –tening, – tening, –tens –*tr.* **1.** To attach firmly to something else, as by pinning or nailing. **2a.** To make fast or secure.... –*intr.* **1.** To become attached fixed or joined. **2.** To take firm hold....

We note that the transitive and intransitive forms are separate sub-entries of the headword; and although the inflectional ending for the past participle is given, there is no full entry, nor any run-on entry for the adjective *fastened*. For neither of the first senses given, is there an example sentence illustrating its use. Without expertise about grammar, ordinary users would not know how the transitive use differs from the intransitive use, or how the adjective occurs in actual sentences of the language. Having access to stative LUs is especially useful when the morphological realization is not just the participial form of the verb used for both the causative and inchoative, as in *liquefy*.v.tr, *liquefy*.v.intr., and *liquid*.a

Moreover, FrameNet captures two vocabulary-wide semantic relationships with frame-to-frame relations: Causative_of for the one-way relation between the transitive and intransitive verbs; and Inchoative_of for the one-way relation between the intransitive verb and the stative adjective.

Reframing was devised as a tool to facilitate the reassignment of annotated sentences to more semantically fine-grained frames that don't necessarily bear a frame-to-frame relationship. The results of reframing, exemplified here in distinguishing between causative, inchoative, and stative, are particularly useful for lexicography more generally, since it is also used in the service of sense discrimination. While the reframing process serves the general purpose of reorganizing existing data in FrameNet, here we have demonstrated how it has been used to implement a lexical semantic concept which has consequences for dictionary writing.

Acknowledgements

References

Atkins, Sue, Rundell, Michael, and Sato, Hiroaki 2003. The contribution of FrameNet to practical lexicography, *International Journal of Lexicography*, 16.3: 333-357.

Fellbaum, Christiane (ed.) 1998. WordNet: *An Electronic Lexical Database*. Cambridge: The MIT Press.

Fillmore, Charles J. 1977. 'Scenes-and-frames semantics' in Antonio Zampolli (ed.), *Linguistic Structures Processing*. Fundamental Studies in Computer Science, No. 59, Amsterdam: North Holland Publishing. Pp. 55-81.

Fillmore, Charles J. 1982. 'Frame semantics' in *Linguistics in the Morning Calm*, ed. by Linguistic Society of Korea, Seoul, Hanshin Publishing Co., 111-137.

Fillmore, Charles J. and Atkins, B.T.S. 1992. 'Toward a frame-based lexicon: The semantics of RISK and its neighbors' in Adrienne Lehrer and Eva Feder Kittay (eds.) *Frames, Fields and Contrasts*. Mahwah: Lawrence Erlbaum. Pp. 75-102.

Fillmore, Charles J. and Baker, Collin F. 2001. 'Frame semantics for text understanding' in Dan Moldovan, Wim Peters, Sanda Harabagiu, Louise Guthrie and Yorick Wilks (eds.) *WordNet and Other Lexical Resources*. Pittsburgh: Association for Computational Linguistics. Pp. 59-64.

Fillmore, Charles J., Petruck, Miriam R. L., Ruppenhofer, Josef and Wright, Abby 2003. FrameNet in action: The case of attaching, *International Journal of Lexicography*, 16.3: 297-332.

Johnson, Christopher R., Petruck, Miriam R. L., Baker, Collin F., Ellsworth, Michael, Ruppenhofer, Josef and Fillmore, Charles J. 2003. *FrameNet: Theory and Practice.* http:/www.icsi.berkeley.edu/~framenet/book.html.

Petruck, Miriam R. L. 1996. 'Frame semantics' in Jef Verschueren, Jan-Ola Östman, Jan Blommaert, and Chris Bulcaen (eds.) *Handbook of Pragmatics*. Philadelphia: John Benjamins. Pp. 1-13.

Pickett, Joseph R. (ed.) 2000. *The American Heritage Dictionary of the English Language.* Boston and New York: Houghton Mifflin Company.

Endnotes

[1] For most of the project, we have used the British National Corpus; recently we added newswire text from the Linguistic Data Consortium.

[2] The characterization of the Attaching frame closely follows Fillmore, et al. (2003), where the reader can also find a detailed description of the FrameNet process.

[3] In earlier work, FrameNet called this relation *composition*, but because of the potential for confusion with *semantic composition*, decided to change the name of the relation.

[4] Note that Causative_of is a more specific relation than the Subframe relation in cases of this sort. Fillmore, et al. 2003 characterized the Attaching frame as having Inchoative_attaching as a subframe, but the newly-created Causative_of relation permits a better analysis.

[5] Although *die*.v is the intransitive counterpart of *kill*.v, the two verbs are not derivationally related, as is the case with the other verbs in Table 1.

[6] This frame-wise distinction is not meant for the entire network of frames in the database.

[7] Atkins, et al. (2003) describes other ways that lexicographers can benefit from FrameNet data.

[8] The absence of an adverb denoting speed (e.g. *quickly*) with the Inchoative_attaching verb *fasten* constitutes a gap in the data.