

FRAMENET IN ACTION: THE CASE OF ATTACHING

Charles J. Fillmore* (fillmore@icsi.berkeley.edu),
Miriam R.L. Petruck* (miriamp@icsi.berkeley.edu),
Josef Ruppenhofer* (josef@icsi.berkeley.edu),
Abby Wright** (acwright@socrates.berkeley.edu)

* International Computer Science Institute and University of California, Berkeley,
California, USA

** International Computer Science Institute, Berkeley and University of
Pennsylvania, USA

I. Introduction

In this paper we provide a step-by-step description of the full process within FrameNet of carrying the treatment of a lexical unit from beginning to end. Following Cruse (1986), we use the term *lexical unit*, abbreviated LU, to designate a ‘word’ taken in one of its senses.¹ The process begins with a description of the conceptual structure, or *frame*, that the LU belongs to, and this description requires (1) characterizing schematically the kind of entity or situation represented by the frame, (2) choosing mnemonics for labeling the entities or components of the frame, and (3) constructing a working list of words that appear to belong to the frame, where membership in the same frame will mean that the phrases that contain the LUs will all permit comparable semantic analyses. In this article we characterize a frame of *Attaching* and suggest a list of words that all have to do with somebody causing one thing to be physically connected to something else, or causing two things to be connected to each other. After briefly surveying a small number of frame-related words, we show the steps toward constructing a complete description of the verb *tie* as it is used in this frame. During this process we encounter uses of *tie* that have different meanings, and we suggest how they might be related to other frames.

The FrameNet working procedure differs in important ways from familiar kinds of lexicography. Instead of concentrating on one word (lemma) at a time and exploring all of that word’s senses (in our terms, all of the associated lexical units) before going on to a new word, FrameNet analyzes lexical units one frame at a time. Because of this, FrameNet work does not lead to the systematic exploration of patterned structures of polysemy. The paper in this volume by

Atkins, Rundell and Sato, however, shows something of how polysemy structures can be studied from FrameNet results.

Our description in this paper follows steps that the reader can relate to the workflow diagram shown as Figure 1.

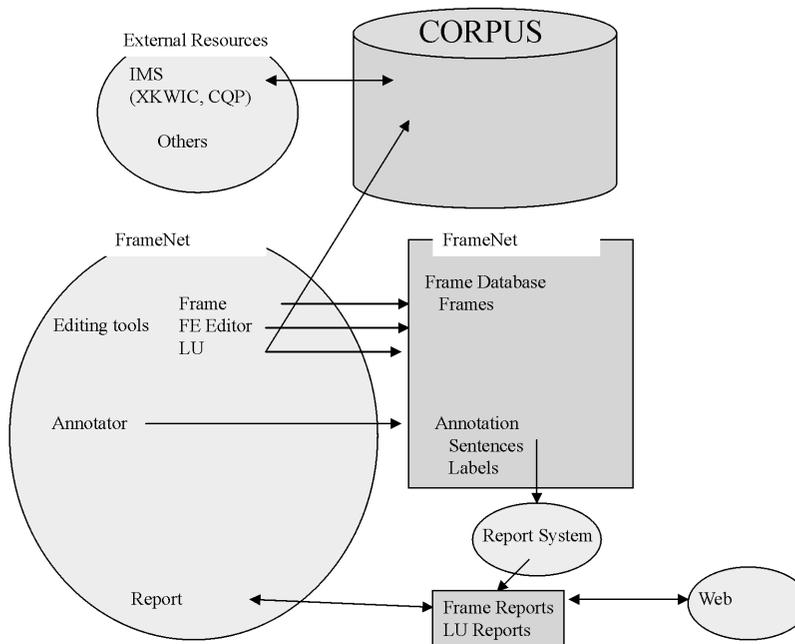


Figure 1: FrameNet Workflow Diagram

Figure 1 illustrates the most important connections among the sub-processes of FrameNet activities. Using External Resources (upper left corner) such as the IMS Corpus Workbench², which is linked to our corpus³, the FrameNet lexicographer scans key-word-in-context attestations of the word being treated, and notices the syntactic and collocational contexts that are most likely to select the intended sense. This information is used to make sure that examples of the lexical unit we want to analyze will be well represented in the sentences we extract from the corpus.

Most of the lexicographic work is carried out using the FrameNet Desktop (see Figure 5 below), a suite of tools for (a) entering and editing frame descriptions and lists of lexical units, (b) extracting sentences from the corpus, (c) annotating selected sentences with frame-relevant labels, and (d) organizing and displaying the results. Describing the frames and choosing *frame element* (FE) names to be used in the annotation is done through the Frame Editor and the FE Editor. Through the LU Editor the lexicographer uses observations about the syntactic and collocational contexts of a given LU to set parameters for creating subcorpora of sentences likely to contain instances of the word in that particular sense.

The collected sentences are made available to the Annotation software, and the annotators choose representative instances of each LU, following selection and labeling principles discussed in section 7 below. The annotated sentences, organized according to relevant syntactic/semantic contexts of use for the LU, and labeled with respect to the manner in which frame-relevant features are expressed in them, are added to the Annotation Database. Various automatic processes (the Report System of Figure 1) operating on the resulting annotations produce corpus-based formal descriptions of the syntactic and semantic combinatory properties of the LU, these being of different types for different kinds of words. A multiplicity of reporting and viewing tools, both FrameNet DeskTop-internal and web-based, make it possible to display the results of the analysis and the annotation in a variety of ways. The FrameNet data is distributed in the form of text files with XML⁴ markup; this enables researchers in various fields, such as Natural Language Processing (NLP), to download the data and utilize it in systems they are building.

2. Choosing a Frame and Developing a Wordlist

In sections 2 and 3 we develop an *Attaching* frame and consider how it can be delimited semantically from other frames. Although, for the sake of narrative expediency, our *Attaching* description will portray this process as a series of discrete steps occurring in a particular order, the reader should know (and will no doubt discover) that the process can loop back at various points. At any point in the process we may choose to modify an earlier decision on the basis of corpus evidence and then proceed anew from there. With this caveat, we now turn to our descriptive task.

The lexicographer begins work on a new frame by giving an informal description of the type of situation or happening that the frame represents and creating a list of words with senses that he or she thinks can be explained with reference to the frame. This is the ‘armchair linguistics’ part of our work: we appeal to native speaker intuition; we consult paper and electronic dictionaries and thesauri; and occasionally we make forays into the lexical semantics literature. The frame that we have decided to walk through here has to do with situations in which somebody attaches (or affixes or joins) one thing to another thing, using some kind of connector.

We create a list of verbs that strike us as belonging to this frame: *append, attach, connect, fasten, join, link, secure, bind, chain, clip, glue, lace, lash, nail, paste, pin, screw, staple, stick, tack, tie, tether, weld, yoke*, etc. We note that some of these verbs imply the use of specific kinds of connectors, such as the noun-derived verbs *glue* and *chain*, while others, such as *connect* and *join*, are more schematic; but all the verbs in the list presuppose the use of some connecting device or means and thus fit our initial characterization of the frame.

With the exception of one sense of the noun *attachment* (the one that refers to the event rather than to an accessory part), and of gerundial nouns based on

Attaching verbs, e.g. *appending*, *attaching*, *joining*, etc., there are few nouns that refer to Attaching actions and we will leave them out of this discussion. However, there exist a number of English nouns that are the names of tools or products that exist for the purpose of attaching things to each other, or that can easily be put to such service. Though these nouns obviously have a special affinity to the Attaching frame, we will compile a separate wordlist for them and put them in a separate Connectors frame, where they will receive a kind of annotation that pays attention to their use in phrases that identify connectors in Attaching scenes. The list of such nouns will include *belt*, *binder*, *cable*, *chain*, *clip*, *clothes pin*, *connector*, *duct tape*, *fastener*, *glue*, *hair clip*, *hook*, *link*, *nail*, *paper clip*, *paste*, *ribbon*, *rope*, *screw*, *shoe lace*, *staple*, *string*, *tack*, *tape*, *tether*, *thread*, *thumb tack*, *twist*, *wire*, *zipper* and many others. In section 7.2 we discuss the distinct kind of annotation we provide for such nouns, and we will show the results for the noun *ribbon*.

3. Choosing a Target LU and Exploring its Use

Using our corpus we extract sentences containing the verb *tie*, one of the central members of the frame, and explore these both to check our understanding of the syntax and semantics of the words in the Attaching frame and to decide on contextual clues for the uses of this lemma in the attaching sense. Figure 2 shows a partial display of the search results for the verb *tie* in the familiar Keyword In Context format.

KWIC list:	
1	is may have to be cooked separately . <Tie> the beef and lamb into compact rolls or squares so that they retain
2	one clasped in both feet , his waist <Tied> down with leather to prevent any unnecessary movement of bum on t
3	the fate of South <pb n=94> Vietnam , <tying> their fortunes to those of the Diem government . Diem proved a f
4	the selvages have been secured) by <tying> a knot on each pair of warp strands in a horizontal direction as
5	o horribly bent captives , their feet <Tied> by a cord and their heads looking as if they are propped upright
6	e sense . Somehow I know it has to be <Tied> in with the leadership election , but I simply do n't know how .
7	if sleeping refreshed it . ` You 've <Tied> up your hair again , " he said casually . ` It is more practical
8	-western edge . Ben jumped ashore and <Tied> the rowboat up to the small , wooden jetty , then turned to give
9	reaned into the house . Sikes quickly <Tied> the rope around the chimney , then began to tie the other end ar
10	this : a pedlar called at a mill and <Tied> up his donkey outside . When he emerged he found that his donkey
11	alton Street . She did n't want to be r <Tied> down by a full-time job , so she decided to become a free-lance t
12	and collect him . Says he 's had to <tie> the guy up , gag him and lock him in the coal cellar to keep him d
13	he centre of each curtain with ribbon <Tied> into bows . Bay windows present the same problem as corner window
14	rs followed the machine , picking and <tying> up bundles of corn . The girls were perhaps more interesting to
15	d shouting . A bloody-mouthed mastiff <Tied> by a chain to a lintel of a door snarled and barked . Somewhere n
16	way the battery pack sits in a pouch <Tied> around your waist and the connecting wire goes straight up your b
17	the lake . " he said . ` See ? We 'll <tie> up on the island and stretch our legs a bit . You 'll feel better
18	a prepare for teachers are as closely <Tied> in with the demands of the curriculum as possible , especially w
19	owever . Monsieur Lemarchand wants to <tie> up all our affairs with some speed . " Glyn just glossed over the
20	country ? Many have all their capital <Tied> up in their homes and have to use the meagre resources of the bas
21	to the headboard with sticky tape and <Tied> the driving wheel to Pete 's cardboard box with string . Mum help
22	ay , Great-Aunt Alicia had decided to <tie> Sara and Matthew Preston together in joint ownership of a house ,
23	ad you ca n't put your ego in a box , <tie> a rock to it and throw it out into the lake , " she said sweetly .
24	ary treatment . McPhee was cleared of <tying> the dog 's paws together . He was unavailable for comment at his

Figure 2: KWIC Results for *tie*

After noting some (relatively) straightforward examples, we briefly address the following issues: the difference between asymmetric and symmetric uses (tying one thing to another versus tying two things together); the mention or non-mention of connectors; and participation of the word in phrasal verbs.

3.1. Observations

Sentences [15] and [21] from Figure 2, the relevant parts of which are reproduced below, exemplify the basic *Attaching* scenario: a smaller object is attached to a larger object that it would not normally be connected to. Typically the attachment to the larger object prevents autonomous movement of the smaller object by holding it in place. Note that the modifying clause in sentence [15], which is in passive voice, represents a state that follows from some agent having performed the attaching.

[15] A bloody-mouthed mastiff **tied** by a chain to a lintel of a door snarled and barked.

[21] He ... **tied** the driving wheel to Pete's cardboard box with string.

Sentence [3], reproduced below, is a metaphorical use of the basic *Attaching* sense. It is special because the phrase *the fortunes of the Diem government* is probably construed as being metaphorically in motion rather than just static and fixed. Thus, the fortunes of the Americans will follow those of the Diem government, instead of simply being attached to them.

[3] Henceforward the Americans became involved in the fate of South Vietnam, **tying** their fortunes to those of the Diem government.

3.1.1. *Symmetric and asymmetric uses.* The initial description of the *Attaching* frame assumed an asymmetric relation, with a smaller object connected to a larger, more fixed object. In examples [5] and [24], however, no such asymmetry is implied: each object (foot or paw) is of equal status and has both the roles of attached object and attached-to object.

[5] ... two horribly bent captives, their feet **tied** by a cord and their heads looking ...

[24] McPhee was cleared of **tying** the dog's paws together.

We must decide on the relationship between the symmetric and asymmetric uses: do both belong to a single *Attaching* frame or should they be treated in separate frames? If all verbs allowed both symmetric and asymmetric uses, we could comfortably treat both uses as belonging to a single *Attaching* frame. Or, if there are verbs that are exclusively associated with asymmetric attaching as well as verbs that are exclusively associated with symmetric attaching, we could easily recognize two separate frames. What we find is that while all the verbs in our list exhibit the asymmetric use, there are some that have only this use. In particular, *attach*, *append*, and *secure*, do not allow symmetric uses, as is shown in the ungrammatical examples (1)–(3).⁵

(1) *I **appended** the letters (together/to each other).

(2) *I **attached** the letter and the photo (together/to each other).

(3) *I **secured** the cables (together/to each other).

Based on these facts, we include all our candidate words in a single *Attaching* frame, showing the behavioral differences with annotation labels, as will be discussed below.

3.1.2. *Connectors*. As noted previously, we do not always expect the connector to be expressed overtly when the verb's meaning entails the use of a certain type of connector. Example [24], shown above, illustrates the point: the dog's paws are understood to be attached to each other by a rope-like entity. Now consider examples [9] and [16] which also lack overt expression of a connector.

[9] Sikes quickly **tied** the rope around the chimney.

[16] ... the battery pack sits in a pouch **tied** around your waist ...

These sentences characterize slightly different situations from that in [24] where the connector is simply not mentioned. While distinct in level of detail, both [9] and [16] can be understood as instances of an attached object that has its own connector: the rope, whose ends form a knot to secure it, and the pouch, which undoubtedly comes with ties of some sort. The decision to treat sentences like [9] as examples belonging to our *Attaching* frame finds support in the fact that there are no lexical items dedicated to describing attaching a connector to one of the objects in some larger *Attaching* scene. The same entity can instantiate both the connector element and the object being connected. It is part of our world knowledge that the shapes of certain objects allow parts of them to function as connectors.

3.1.3. *The Treatment of Phrasal Verbs*. Many examples in our corpus consist of combinations of *tie* with various particles and prepositions, including *up*, *down*, *in*, and *into*. As we will see, some of these illustrate the sense of the single word *tie* that we have targeted ([2], [7]), others illustrate the phrasal verbs, *tie up* and *tie down*, which we treat as separate lexical units, but still in the *Attaching* frame ([8], [10], [17]); and still others need to be treated as belonging to different frames ([1], [12], [13], [14], [19], [20]).

[1] **Tie** the beef and lamb **into** compact rolls so that they retain their shape

...

[2] ... his waist **tied down** with leather to prevent any movement ...

[4] ... the selvages have been secured by **tying** a knot on each pair of warp strands ...

[6] Somehow I know it has to be **tied in** with the leadership election ...

[7] You've **tied up** your hair again.

[8] Ben jumped ashore and **tied** the rowboat **up** to the small, wooden jetty.

[10] A pedlar called at a mill and **tied up** his donkey outside.

[11] She didn't want to be **tied down** by a full-time job ...

[12] Says he's had to **tie** the guy **up**, gag him and lock him in the coal cellar.

[13] ... and draw in the center of each curtain with ribbon **tied into** bows.

[14] The harvesters followed the machine, picking and **tying up** bundles of corn.

- [17] We'll **tie up** on the island and stretch our legs a bit.
 [18] ... for teachers are as closely **tied in** with the demands of the curriculum ...
 [19] Monsieur Lemarchand wants to **tie up** all our affairs with some speed.
 [20] Many have all their capital **tied up** in their homes ...

Some of the particles in the phrasal verbs indicate the resulting position of the attached item, as in [2] and [7], and these can be seen as exemplifying our sense of *tie*: by a tying action, something that is likely to come down is made to stay 'up' or something with an inclination to get up is made to stay 'down'. In other cases, the phrasal verb does not signal a resulting position on a vertical scale. For example, *tie up* in [8] and [10] indicates a kind of completive meaning, still within the *Attaching* frame, but belonging to a different LU.

The particle *up* in sentence [12] also exemplifies the completive sense, but primarily reports an act of *Immobilization*, where an attaching act is the means of accomplishing this resultant state. In other sentences ([14], [19], [20]) we also find different LUs that would not be in our *Attaching* frame. In sentence [14], the bundles of corn may be created by tying together individual corn-stalks, and thus this sense of *tie* is mainly about creation. The figurative sense of *tie up* with a completive meaning in [19] belongs to the *Activity_finish* frame that also contains verbs like *complete* and *wrap up*, as well as idioms such as *tie up loose ends*, *wrap things up*. Because [20] specifically implies a lack of access to a resource, we conclude that it does not exemplify our sense of *tie* in the *Attaching* frame. We see *tie down* in [11] as somewhat similar in meaning to *tie up* in [20], although being *tied down* is something to be resented.

The phrasal verbs *tie in* and *tie into* are yet different LUs in other frames. Sentences [6] and [18] have a cognitive meaning and would belong to a frame dealing with associations, perhaps at a quite general level. We treat [1] and [13], which refer to a shape or configuration resulting from tying one or more objects together, as examples of a separate sense of *tie* in a frame that deals with transformations. Finally, in sentence [4], *tie* seems to exhibit yet another, slightly different sense: there is no notion of transformation but just of creation.

Although more could be said about the senses of *tie* and their inter-relationships, we believe that the analysis is sufficient for our immediate concern, delimiting the boundaries of the *Attaching* frame.

3.2. What the dictionary says

Having completed our initial analysis of the uses of *tie*, we compare our findings with those of the *Concise Oxford Dictionary of the English Language* (Pearsall 1999, henceforth COD) whose entry for *tie* as a verb is given in Figure 3.

tie ● v. (**tying**) **1** attach or fasten with string, cord, etc. > form into a knot or bow. > (**tie someone up**) restrict someone's movement by binding their arms or legs or binding them to something. > (**tie something up**) bring something to a satisfactory conclusion. **2** restrict; limit. > (**tie someone down**) restrict someone to a particular situation or place. > (usu. **be tied up**) informal occupy someone to the exclusion of other activity. > (**tie something up**) invest or reserve capital so that it is not immediately available for use. **3** connect; link. > (**tie in**) be or cause to be in harmony with something. > hold together by a crosspiece or tie. > Music unite (written notes) by a tie. > Music perform (two notes) as one unbroken note. **4** achieve the same score or ranking as another competitor. **5 (tie into)** N. Amer. informal attack or get to work on vigorously.

Figure 3: COD entry for tie.v

While we are pleased that sense 1 of the COD entry is the one we identified in our *Attaching* frame, we must note some important differences between our observations and the dictionary's. To begin with, the dictionary has several senses that we did not see in our 24-sentence corpus. We know independently that COD senses 4 and 5 are represented in the BNC and we would have found them eventually; but two of the subsenses of 3, “hold together by a crosspiece or tie” and “unite (written notes) by a tie”, are rather technical and we might not have encountered them. We note too that we have found uses that are not covered in the COD entry, including the resultative use as in *tie the meat into rolls*. Other important differences exist in the structuring of senses and subsenses. For instance, the COD lists the creation use of the verb (e.g. *tie a knot*) as a subsense of the central *Attaching* sense, while we include it in *Knot_Creation*, a frame that is exemplified in sentence [4] in Figure 2. Similarly, the COD also has the physical restraint and the completion senses as subsenses of the central sense, rather than as separate senses on a par with the *Attaching* sense. Furthermore, among the metaphorical *limit* senses, being tied up is distinguished for capital and people, where we might collapse them under the heading of resource. Finally, the dictionary does not provide an indication of sense-to-sense relations, whereas we at least have the ability to do so via frame-to-frame relations (see section 5).

4. Defining Frames and Frame Elements

Now that we have a clearer sense of the spectrum of senses in the verb *tie*, we want to characterize more precisely the frame in which “our” sense of *tie* occurs. In the previous section we made decisions regarding both the limits of our frame and of the targeted sense of *tie*: we decided that *tie in with* does not fit the frame,

and that while *tie up (the donkey)* is a part of our frame, it exemplifies a different LU, the phrasal verb *tie up*. The next step is to incorporate these decisions regarding our sense of *tie* into a *frame definition*, hoping that what we come up with will not require revision as we examine other words in our initial wordlist.

A FrameNet frame definition is a schematic presentation of a situation type that underlies the meaning of a word (or of the members of sets of words) along with named participant roles or aspects of the situation, which we call frame elements. Of lexicographic interest is the manner in which the frame elements are given linguistic expression in sentences containing our LU. Frame elements can often be seen as instances of such broader semantic roles as AGENT, UNDERGOER, EXPERIENCER, etc., but we define them in frame-specific ways. There are various reasons for this, the most important being that we are able to give precise definitions to the frame elements for a particular group of words without needing to decide first how they can be made to fit into a small predetermined inventory of semantic roles.

To begin we might first consider some canonical examples, not limiting ourselves to the verb *tie*. Consider the following sentences.

- (4) They took her home and **tied** her to the bed with string saved from their Christmas packages.
- (5) Lise retrieves them to **tie** his hands together.
- (6) Amelia rolled up Nina's torn sleeve and **tied** the tourniquet tight around her upper arm.
- (7) The calves, when only a few days old, are **tethered** by the neck in a stall about 60 cms wide and 150 cms long, until they are too large to turn.
- (8) A garland was **stapled** to the underside of the mantelpiece.

From these examples and our informal characterization of the frame as 'having to do with situations in which somebody attaches (or affixes or joins) one thing to another thing, using some kind of connector', we begin to characterize the core participants in our frame, which will correspond to our *core frame elements*.

First we notice the presence of an AGENT, which, for this frame, is the person who brings about the attaching of two items. In sentences (4)–(6), above, the AGENT is instantiated by the phrases *they*, *Lise*, *Amelia*; the AGENT is omitted in (7) and (8). Secondly we have an ITEM which, in the asymmetric case, is the smaller object affixed to the larger, more stable GOAL. In the above examples, the ITEM is realized as *her*, *the tourniquet*, *the calves*, and *a garland*. Note that in example (5) *his hands* is not exactly the same as the other ITEM phrases: this is the symmetrical version of *Attaching* as described in Section 3. The two things joined together cannot be separated as ITEM and GOAL; for such cases the naming convention we follow is to use the plural of the frame-element name, here 'ITEMS'. In example (5), then, the phrase *his hands* is said to instantiate the frame element ITEMS.

The GOAL is realized in the above sentences in a variety of prepositional

phrases, for example, *to the bed* and *around her upper arm*.⁶ In example (4), we see a frame element instantiated by the phrase *with string saved from their Christmas packages*, which we will call CONNECTOR, since this constituent identifies the physical object used to join the ITEM and the GOAL. As mentioned in Section 2, many verbs in this frame incorporate their CONNECTOR, for example *glue* and *staple*. Lastly, the phrase *by the neck*, as seen in example (7), represents what we will call the HANDLE, the part of the ITEM that allows the CONNECTOR to hold it in place.

Besides these *core* frame elements, there are others, which we identified as *peripheral* from the point of view of the *Attaching* frame, since they show up in all frames of agentive action. In particular, we have expressions of PLACE, TIME, MEANS and MANNER. There is an additional kind of frame element that we need to recognize, and for now categorize as peripheral: those that become available when the words we are studying appear in certain special constructions. In example (5), we see the frame element RESULT in the word *together*: the result of tying the hands is that the hands are now together.

With these frame elements in mind, we define our frame, making careful note of the following entailments and presuppositions. The AGENT is a volitional actor, to whom intentions and motives can be ascribed; the ITEM must be attached to a larger or more stable GOAL, whether or not the latter is mentioned; the attaching must be achieved with the help of a CONNECTOR, independently of whether such an entity is identified or incorporated into the verb's meaning; and the ITEM and GOAL once connected are not seen as constituting a new entity, as would be the case in a *Fusion* frame. The final frame description for *Attaching* is presented in Figure 4.⁷

Definition:

An **Agent** attaches an **Item** to a **Goal**, generally by manipulating a **Connector**, such that the **Item** is fixed to the **Goal**. Alternatively, an **Agent** uses a **Connector** to attach two **Items** to each other, where each serves as a **Goal** for the other. In either case, the **Connector**, if present, remains to bind the **Item** or **Items**. This asymmetric or symmetric relationship profiles the parts and has them remain as parts, without creating a new entity.

The robbers **TIED** Paul to his chair with a rope.

The robbers **TIED** his ankles together.

The **Goal** may be expressed solely with a preposition:

Patsy **TIED** her skirt on.

In this sentence the **Goal** is understood as being "on her body".

FEs:

Core: The **Agent** manipulates an object. **Agent** typically refers to animate beings, but is also extended to mechanical and other inanimate objects. It is most frequently expressed as the external argument of verbs.

Agent [Agent]

The robber **TIED** Harry to the chair.

Connector [Conn]

This FE identifies the entity or substance that forms the bond that maintains the **Item** or **Items** in a fixed position; it is usually expressed in a "with"-PP.

The robber **TIED** Harry to the chair with a rope.

Goal [Goal]

Goal identifies the location to which an **Item** is attached.

The robber **TIED** Harry to the chair.

Item [Item]

This FE identifies the **Item** that the **Agent** attaches to the **Goal**.

The robber **TIED** Harry to the chair.

Items [Items]

This FE identifies the plural **Items** that are tied together, resulting in their mutual attachment.

The robber **TIED** Harry's ankles together.

Figure 4: Frame Description for Attaching

In the next section we describe the software which allows us to enter this information into the database.

5. Entering the Attaching Frame into the Database

The frame and frame element definitions that we developed are entered into our database via various editors that are accessed from the integrated FrameNet DeskTop, shown in Figure 5 and Figure 6. The working screen for the FrameNet DeskTop has a *navigation frame*⁸ at the left and a *content space* on the right. The content space holds the editors that appear when database objects are accessed from the navigation frame. The navigation frame contains an expanding list that gives direct access to the main kinds of database objects: frames, frame elements, and lexical units. Any of the objects in the list can be expanded into yet another list, bottoming out at the example sentences. The frame expands to lists of frame elements and lexical units. Lexical units in turn expand to information about the lemma including its part of speech, and to a list of subcorpora. The lemma expands to a lexeme along with a list of word forms. The subcorpora finally expand to individual sentences. Section 6 will show in detail how new lexical units are added.

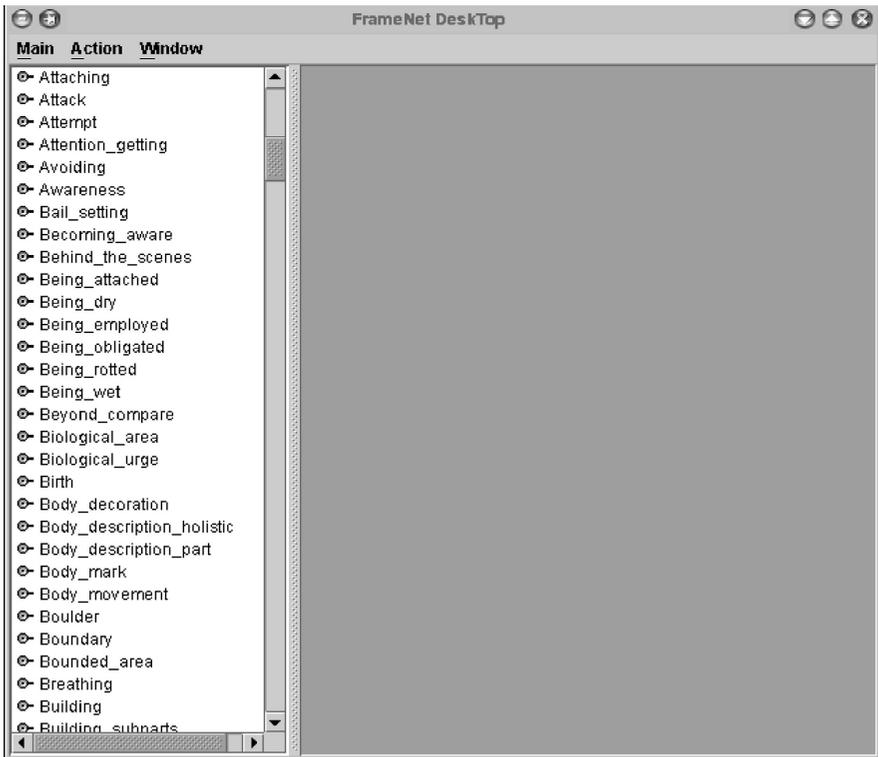


Figure 5: Initial view of the FrameNet Desk Top

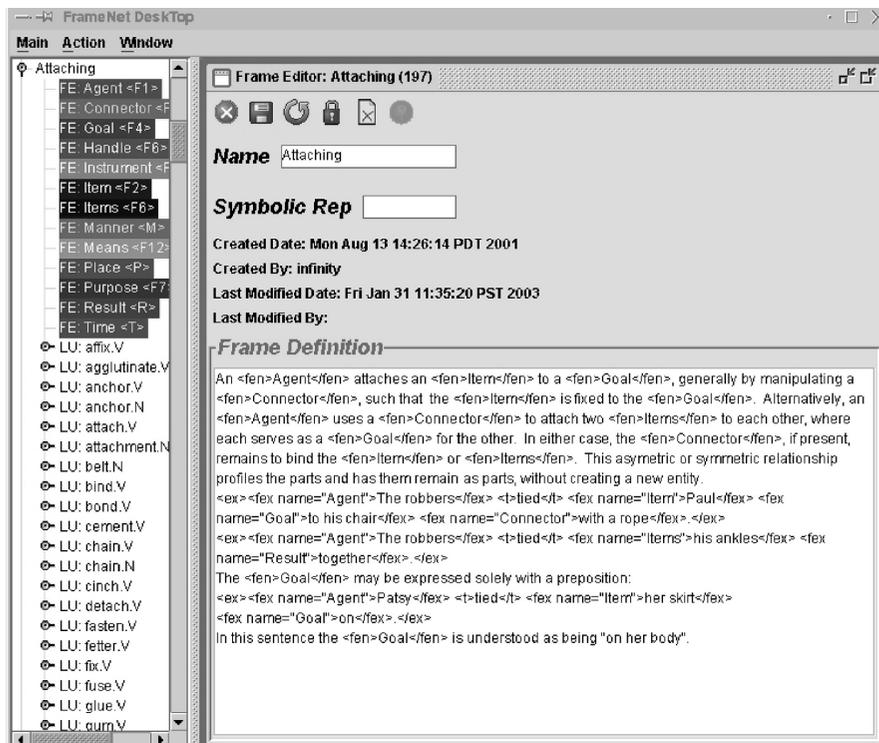


Figure 6: Frame Editor

Information is added to the database by selecting a frame name or frame element in the navigation window; clicking on the name will bring up the appropriate editor in the content space. In Figure 6, the content space is occupied by the frame editor. The frame editor is opened to the `Attaching` frame and we see the definition that we crafted. The definition is less readable here due to the addition of XML-markup that allows the names of frame elements to be displayed with a colored background in web browsers and in our reports.

Next, we use the Frame Element Editor, shown in Figure 7, to enter FE definitions. The most important parts of the FE Editor are (1) the definition box; (2) the radio button for choosing between core and non-core status; (3) the semantic role rank drop-down menu; and (4) the drop-down menu for choosing screen and printer colors as well as a field for specifying a shortcut key.

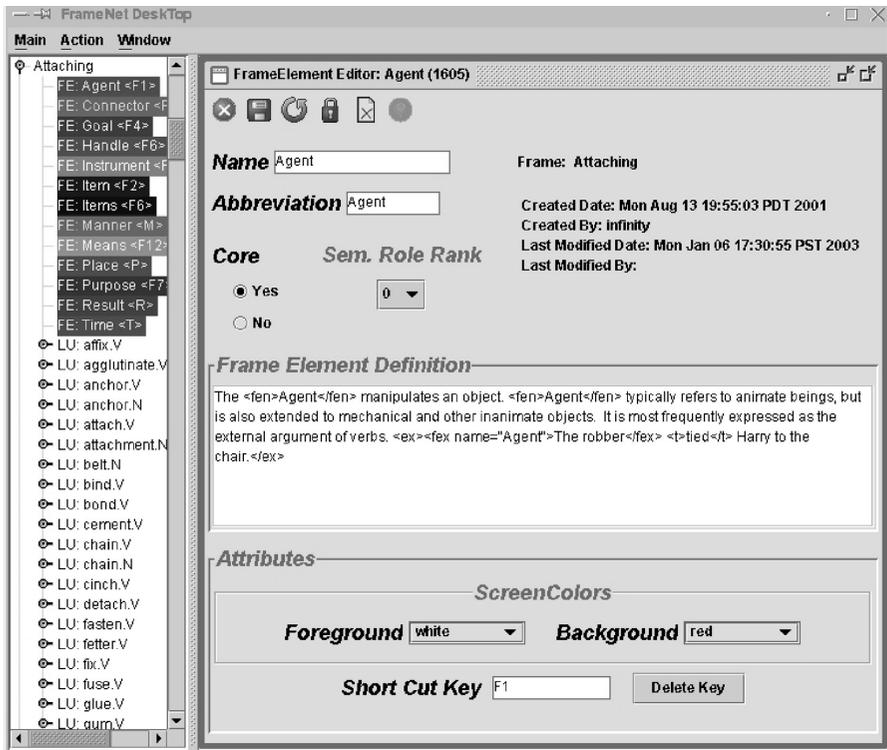


Figure 7: Frame Element Editor

The definition for each frame element contains XML-markup and example sentences just like the principal frame definition. The core/non-core value is assigned via a radio button. The core/non-core distinction corresponds roughly to the one between arguments and adjuncts in traditional grammatical analysis. When a core frame element such as AGENT in the *Attaching* frame is omitted, we record that in our annotation; we do not make a special point of indicating missing peripheral elements (see section 7.1).

The semantic role rank list allows us to indicate which grammatical relation we expect the frame element to have with respect to the target verb in a simple *active* clause. The AGENT in the case of the *Attaching* frame is assigned ‘0’ since it is the frame element that is expected to show up as subject in an active-voice sentence; ‘1’ is assigned to ITEM because it is the expected direct object, etc.⁹

After defining the frame and its frame elements, we can indicate relationships between the current frame and other frames by means of the Relation Editor. Recording such relationships allows us to economically capture semantic generalizations across frames regarding the type of participants involved, what happens, and what states result.

The Relation Editor connects the current frame to several other types of objects: frames that the current frame elaborates or inherits; frames that are

subframes of the current frame; and frames that the current frame *Evokes*. In addition, we can indicate that there are fine distinctions among frames within a group of closely related frames with the relation name *See-Also*. We will not go through all of these types of relationships here but only consider two of them, inheritance and subframe relationships.

First, we will discuss inheritance relationships between the current frame and one or more parent or child frames. Frame inheritance is a relationship in which a *child* frame is a more specific elaboration of the *parent* frame. In such cases, all of the frame elements, subframes, and semantic types of the parent have correspondents in the child frame. The child may, however, have additional subframes, frame elements, and semantic type constraints that are not found with the parent frame. Note also that FrameNet allows for multiple inheritance, that is, a frame can have more than one parent. The Frame Relation Editor for inheritance is pictured in Figure 8.

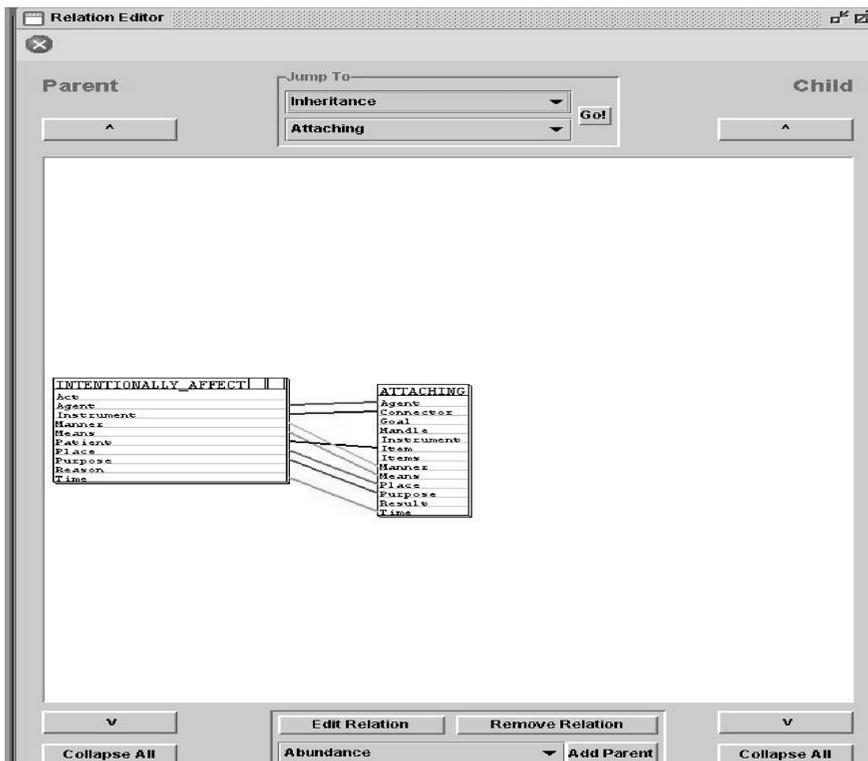


Figure 8: Frame Relation Editor – Inheritance

The inheritance relationships involving our current frame are displayed in the central viewing area of Figure 8. As can be seen, the *Attaching* frame has the *Intentionally_affect* frame as a parent. By clicking on that frame, we could select it as the current frame and then see all the frames to which it is related via inheritance.

Inheritance relationships between pairs of frames have to be specified one at a time. Parent or child frames of the current frame can be selected from the drop down box (in the center of the border area) at the bottom of the window. Clicking on the Add Parent/Child button will create the formal link between the two frames under consideration. Currently, we record correspondence relationships between frames only for their frame elements; as will be shown below, we also have the technical capability to record subframe relations.

Frame element mappings are established by selecting the frame in the display area and clicking on Edit Relation. Another window comes up (not pictured) and shows a set of default mappings between the FEs in the two frames; the default mappings may be accepted, modified, or edited, as appropriate. The existing mappings are displayed in the middle window. The AGENT of the Attaching frame is equated with the AGENT of the Intentionally_affect frame, the CONNECTOR frame element is equated with the INSTRUMENT frame element, and so forth.

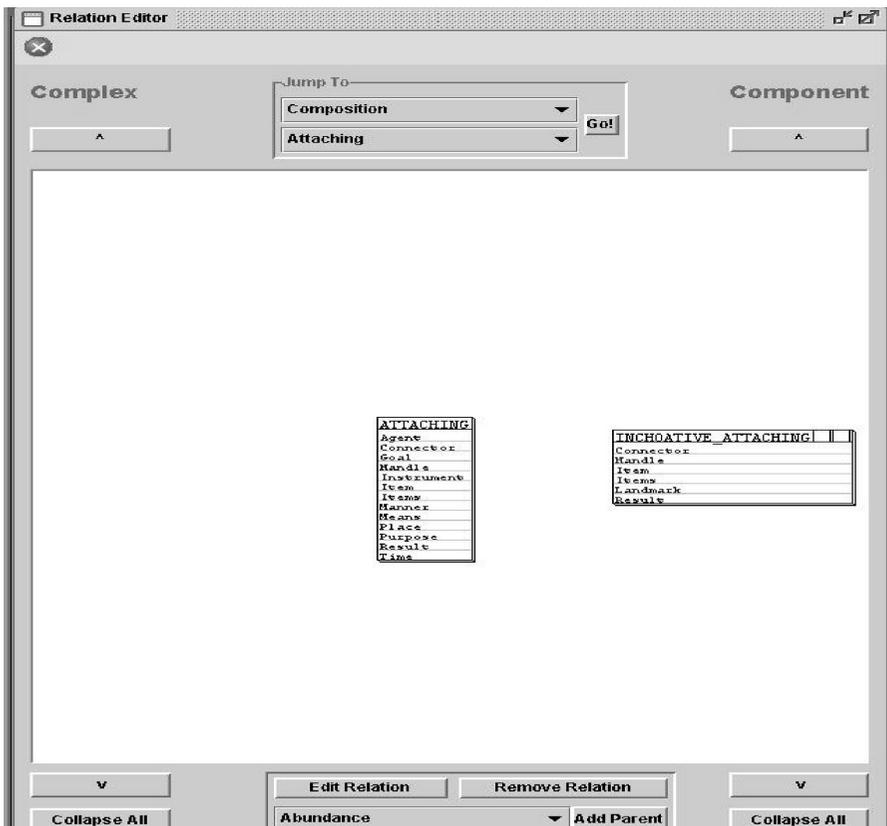


Figure 9: Frame Relation Editor – Composition

The subframe Relation Editor, shown in Figure 9, is virtually identical to the inheritance Relation Editor; however, notice Complex and Component at the top

of the frame (as opposed to Parent and Child for inheritance). In this mode, we indicate the frames that are related to each other via Composition. The concept of subframes allows us to deal with frames that are complex in that they designate sequences of states of affairs and transitions between them, each of which can itself be separately described as a frame. Each of the separate frames is related to the complex frames via the composition relation. For example, consider the verb *enter* in the *Arriving* frame, as in *Jack entered the room*. The entire event of entering is complex, characterized by two states of affairs (one that obtains before Jack enters the room and one that obtains afterwards) and a transition between the two states encoded by the verb itself. In cases of composition, frame elements of the complex frame may be identified with the frame elements of the subparts, although not all frame elements of one need to have any relation to the other. In this respect, composition contrasts with inheritance.

Our sample frame, *Attaching*, has the *Inchoative_attaching* frame as a subframe (Figure 9), which by default is taken to cover only cases where an *ITEM* comes to be attached to a *GOAL*, or is so construed, without the influence of an external cause. For a verb like *connect*, but not *tie*, we have instances of stative (*This wire connects my house to yours*), inchoative (*Through the new highway construction program Highway 3 will connect with Highway 10*) and causative (*We connected the hose to the spigot*) frames.

6. Adding the Lexical Unit *tie* to the *Attaching* Frame and Generating Subcorpora

Having entered the definitions for the *Attaching* frame and frame elements using the Frame Editor, the next step is to add the lexical unit *tie* to the frame. We begin by right-clicking on the *Attaching* frame in the navigation bar and selecting ‘Create LU’ from the menu that appears. This brings up the Lexical Unit Editor seen in Figure 10. Through this editor, we specify the following kinds of information about each lexical unit: (1) its name, (2) its part of speech, (3) its meaning, and (4) information about its formal composition. Like most of our lexical units, *tie* consists of only one lexeme, but we also recognize multi-lexeme lexical units. For these we identify (a) the head lexeme of the lemma, (b) places where the component lexemes can be separated by intervening words, and (c) the word forms of the head lexemes. For example, the head word of the multi-lexeme LU *tied up* (in the sense of being busy) is *tie*; the lexemes cannot be separated by intervening words; and its word forms are *tie*, *ties*, *tying*, and *tied*.

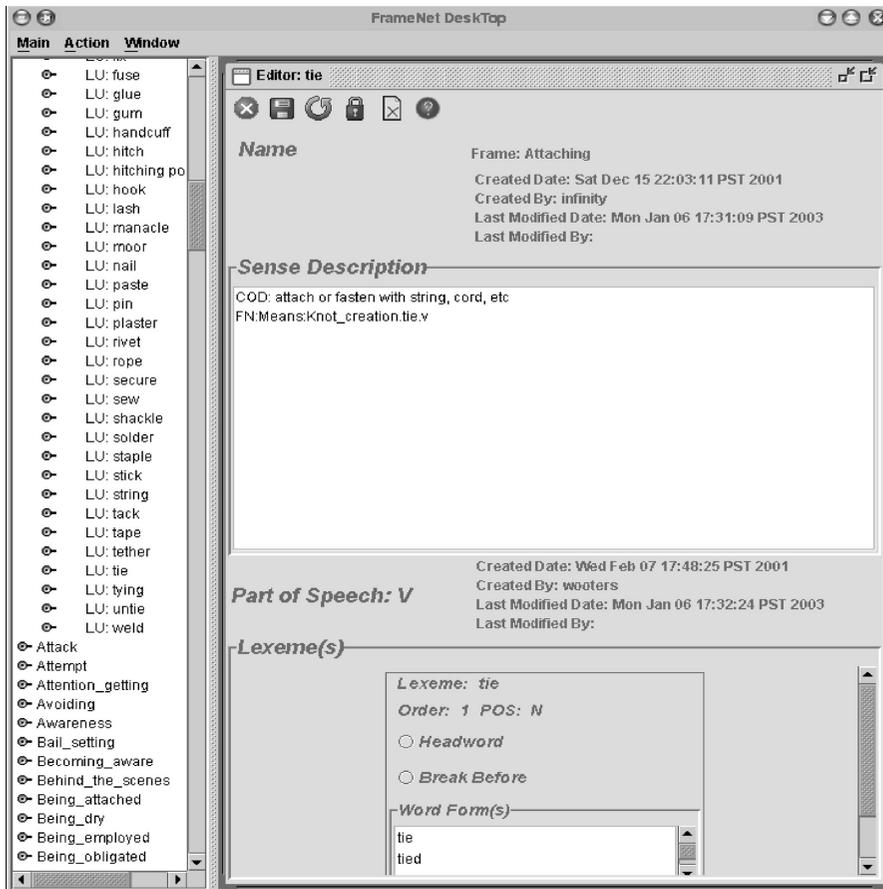


Figure 10: Lexical Unit Editor

Now we are ready to specify the searches that will enable us to extract from the BNC sentences containing *tie* whose grammatical form makes them likely to exhibit the sense we are trying to study. The Subcorporation¹⁰ Query Definition page, as shown in Figure 11, has three major sub-parts: (1) a collocate block at the top, (2) a constituent block in the middle, and (3) a prepositional phrase block at the bottom. As can be seen from the picture, we have included a large number of choices and specifications in the subcorporation process, many of which will be irrelevant for any given lexical unit. The full set of choices underlying any query definition allows us to (1) compensate for various errors in the tagging and lemmatizing in the copy of the BNC we are working from; (2) create subcorpora for collocates of the target and for the syntactic environments in which the target occurs; and (3) eliminate from the pool, through the ‘throwaway’ choices, those contexts which we are sure will yield sentences where the word is not in the intended frame.

lexunit_name

search_by lemma word

ignore_BNC_part_of_speech_tags

take_all

left_throwaway

left_collocate

right_throwaway

right_collocate

np kiss me

np_np brought her friend flowers

np_ap consider it foolish, wiggled his shoulder free

ap wriggle free

np_vp advised her to go

np_vp kept them laughing

np_sfin told them (that) he was back

np_wh told him what to do, told her where the money was

vping needs fixing

sw asked where she was, knew what he wanted

sfin claimed (that) it was over, agreed it was over

directquote no, said he / "no", he said / he said: "no"

sbrst demanded (that) he leave

vp agreed to do it over

vpbse helped make desert

np_pp make ice cream for her friends

np_ping prevent him from leaving, answer the qu

np_pwh asked her about what it all meant

np_particle put the dishes away/ put away the dish

particle touch down, turn over

particle_pp put up with it

particle_wh figured out how to get there

pp look at the picture

Figure 11: Subcorpus Query Definition Page

The buttons below the LU (lexunit) name allow us to carry out the search either by word form (e.g., by *tie*, *ties*, *tied*) or by lemma. We use the former when it is necessary to compensate for lemmatization errors in the corpus or when singular and plural forms of a noun have different meanings, as with *hostility* and *hostilities*. Furthermore, if we know there is a part of speech error in the corpus, it is necessary to check the ‘ignore_BNC_part_of_speech_tags’ box. Checking the ‘take_all’ box generates a subcorpus that includes all examples of a word form or lemma. We use this option if there are one hundred or fewer instances of a word in the BNC. If chosen, no other boxes or fields need to be checked or filled in. In subcorporating for *tie*, we do not use any of these options.

The fields in the collocate block, near the top of the figure from ‘left_throwaway’ to ‘right_collocate’, allow specification of throwaways and collocates to the left and to the right of the target. For our intended sense of *tie*, we want to eliminate clearly out-of-frame sentences from the pool from which

subcorpora are built; we choose *game* and *score* as throwaways because they are likely to signal a sense of *tie* – “have the same score” – which is not the one we’re interested in. But we choose *rope* and *string* as positive right collocates since each of these is likely to occur as the CONNECTOR frame element with our sense of *tie*.

The second block, from ‘np’ down to ‘vpbse’, consists of a series of checkboxes for various post-verbal complement patterns that do not involve any prepositional phrases.¹¹ The only pattern that we need to search for in the case of *tie* is the ‘np’ pattern at the top of the constituent block. The ‘np’ indicates that the verb can take a single noun phrase as its only post-verbal complement. Thus, this pattern should match simple transitive uses with a direct object, such as *They were tying her hands*. Checking this box will also result in the automatic generation of two sub-corpora of passives, one with an overt *by*-phrase and the other without.

The final block, from ‘np_pp’ to ‘pp’, is the prepositional phrase block consisting of another series of text fields. All except the last one are used to specify prepositions or particles that will occur with the target in a particular complementation pattern. For *tie* we include the prepositions *to*, *about*, *around*, *behind* and *with*. We know from earlier study of corpus attestations that the first four are likely to head phrases realizing the GOAL frame element; the last one is likely to head phrases realizing the CONNECTOR frame element. We use these prepositions in two patterns. In the first, ‘np_pp’, a prepositional phrase headed by one of these prepositions follows a post-verbal noun phrase (as in *tied the dog to a tree*). In the second, ‘pp’, the relevant prepositional phrase occurs immediately after the target word (as in *it was tied to a tree*). We also search for occurrences of *up* and *down* after *tie* since these combinations have literal uses that belong to our Attaching frame (as in *tied her hair up*).

Note that the subcorporation page for verbs provides all the possible options, or search parameters, for any kind of verb, i.e. no distinction is made between intransitives, transitives, and ditransitives. There are no default searches. The lexicographer has to decide, based on the evidence from corpus exploration, good judgment, and the examples provided to the right of the check boxes, which searches are linguistically meaningful for a given LU. The subcorporation forms for noun and adjective targets (not displayed here) are necessarily different from that for verbs since members of these word classes enter into different syntactic relationships with their frame elements than verbs do.

Once the lexicographer’s specifications are saved, a series of automatic processes takes place to generate annotation-ready subcorpora. The search specifications are used to fill in a template file containing regular expression searches. The fully specified searches are then run successively by a *corpus query* tool. In the process, the sentences that are matched against a particular search pattern are removed from the remainder corpus that is going to be used for further queries. This is done to prevent multiple occurrences of the same sentence and to prevent throwaways and collocates from obscuring the full

variety of noun phrase types that we would like to find in the pattern-based searches in the constituent and prepositional phrase blocks. The results for each search are saved to separate subcorpora, with the exception of throwaway searches, which are not saved at all. Sentences that are too long or contain hesitation markers or other indications of non-fluent speech are discarded. If there are still more than 20 sentences in a subcorpus afterwards, 20 are randomly chosen and the remainder discarded. Our reason for limiting the size of the subcorpora we examine is that our goal is to choose three to five examples of each of the patterns that we look for: we are interested in exemplifying the variety of patterns that exist, not in being statistically representative.

The subcorpora resulting from the preceding operations are then imported into our database where they are ready for annotation.

7. Annotation Software

Once the subcorporation software is run, the randomly chosen sentences for each subcorpus are available for annotation with the FrameNet Annotator, shown in Figure 12. In this section, we show the two kinds of annotation done in FrameNet. We will show annotation relative to our target LU *tie*, and annotation relative to a second LU, the slot filler *ribbon* from the Connectors frame. We will explain our method, and demonstrate the editing facilities in the course of annotation.

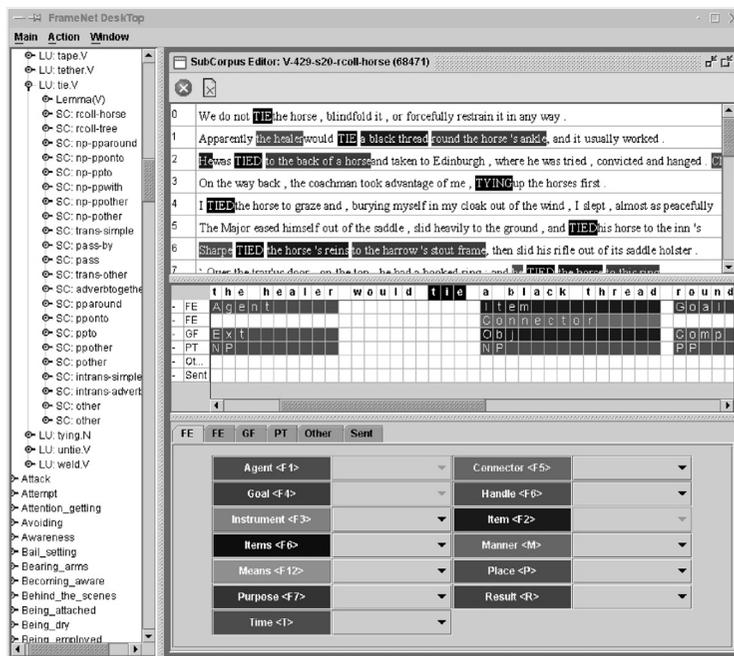


Figure 12: FrameNet Annotator

The annotation window is divided into three main sections: the top one is for viewing a subcorpus and selecting a sentence, the middle one is for annotating a selected sentence, and the bottom one provides lists of labels available for each of the different annotation layers. In Figure 12 the frame element (FE) layer is displayed, so the labels are those that have been specifically defined for annotating words from the `Attaching` frame.

Before annotation begins, a FrameNet lexicographer becomes familiar with the frame and frame element definitions, along with canonical examples, given in the Frame Information Report, as seen in Figure 4. Once an LU is chosen for annotation, a list of its subcorpora appears in the left frame of the FrameNet DeskTop, as seen in Figure 12; selecting a subcorpus makes it appear in the viewing space. The FrameNet lexicographer reads through the list of sentences in the viewing space and selects an appropriate sentence for annotation from the left navigation frame. The selected sentence appears in the annotation workspace and the FEs for the frame appear in the bottom section of the Annotation Window. For this figure the layers being used are Frame Element (FE), Grammatical Function (GF) and Phrase Type (PT).

FrameNet annotation is done relative to one lexical unit, the target, which is most often a single word, but may also be a multi-word expression (e.g. *give up*, *hold court*). We annotate whole constituents, not just the head word of the phrase that instantiates the frame element. These dependents are annotated, on their separate layers, for Frame Element identity, Grammatical Function (GF), and Phrase Type (PT).

The annotation space in Figure 12 shows annotation of a sentence selected for our lexical unit, *tie.v*. In the highlighted sentence, the FE `AGENT` has been assigned to *the healer*; `ITEM` has been assigned to *a black thread*, and `GOAL` has been assigned to *round the horse's ankle*, each of these being a core FE in the `Attaching` frame.

Sometimes a single constituent instantiates two frame elements. We call such situations *frame element conflation* and tag each frame element on a different FE layer. Notice that the FE `CONNECTOR`, also a core element, is tagged as a secondary FE on the noun phrase *a black thread*. We consider it a secondary FE because `ITEM` is the primary semantic role of the constituent, and its phrase type is typical for the primary FE, `ITEM`. No phrase type or grammatical function information for frame elements tagged on the secondary FE layer is provided. Note that a frame element tagged on the second annotation layer may also be instantiated as a constituent on its own. Compare our example sentence in Figure 12 with the one in (9), where `ITEM` and `CONNECTOR` are realized separately.

- (9) He fastened the panel from an old radio to the headboard with sticky tape and **ti**ed [the driving wheel `ITEM`] to Pete's cardboard box [with string `CONNECTOR`].

7.1. Annotation of Non-Core FEs

We also provide annotation for non-core frame elements, such as *TIME*, *PLACE*, *PURPOSE*, *RESULT*, and so on if they are present in the sentence.¹² Non-core FEs are not necessary conceptually; they are not part of what makes a frame unique. In other words, expressions of *TIME* and *PLACE* may occur with most events, and event frames are not differentiated by the fact that they occur in time and space. Similarly, while the *PURPOSE* or *RESULT* of an action can be expressed, doing so does not distinguish one action and its frame from another. Grammatically, non-core FEs cannot be subject or object of a target verb, and they frequently are prepositional phrases or adverbs.

Our initial exploration of the use of *tie* brought to our attention examples with the particles *up* and *down* and with the adverb *together*. We will end up using many of these in the *Attaching* frame, but we have not yet decided which will be used as separate phrasal LUs, e.g., the phrasal verb *tie+up*, and which will be our LU seen as occurring with an independent particle, e.g. *up*, yielding a compositional, rather than an idiomatic, interpretation of the whole. The following sentences from our corpus illustrate the problem, using both *tie* and another verb from the frame, *tack*.

- (10) **Tack** them [*up* GOAL] on a wall, and note the differences.
- (11) **Tack** a rug [*down* GOAL] lightly.
- (12) I **tyed** them [*together* RESULT] with rope.

Grammatical function and phrase type information are derived algorithmically by a chunk parser and instantaneously displayed on the screen. FrameNet lexicographers review the results of the automatic GF and PT tagging to make corrections, if necessary. Notice that the labels for each of the three layers (FE, GF, and PT) appear in the middle section of the annotation frame.

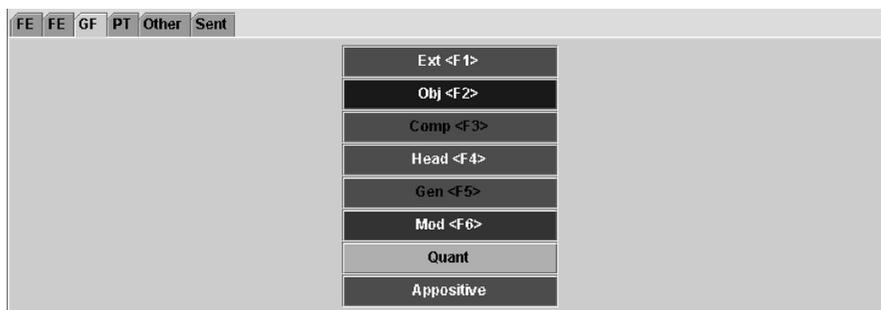


Figure 13: Grammatical Function Labels

The grammatical functions that can be assigned, displayed in Figure 13, describe the ways in which the constituents satisfy abstract grammatical requirements of the target word. In the pictured sentence in Figure 12, the GF

Ext(ernal)¹³ has been assigned to *the healer*; Obj(ect) has been assigned to *a black thread*, and Comp(lement) has been assigned to *round the horse's ankle*.

FE	FE	GF	PT	Other	Sent	
				N <F1>	VPhrst <F9>	Ssub
				NP <F2>	Sfin <F10>	AVP
				Poss <F3>	Swh <F11>	QUO
				PP <F4>	Swhether <F12>	AJP
				PPing <F5>	Sing	A
				VPfin <F6>	Sto	Num
				VPto <F7>	Sforto	Sun
				VPing <F8>	Sbrst	

Figure 14: Phrase Type Labels

The phrase types that can be assigned to the constituents we tag are displayed in Figure 14. In our annotations, AGENT and ITEM frequently appear as NP (when appearing as subject or object), GOAL, CONNECTOR and HANDLE frequently as PP (with such prepositions as *to*, *with* and *by*), and RESULT can be realized as AVP (adverb phrase).

It is also possible to tag frame elements that are conceptually necessary but do not occur as lexical or phrasal material; these we call *null instantiations*. We indicate in this way the absence of the core frame elements because doing so provides lexicographically relevant information about omissibility conditions for target words. We recognize three types of null instantiations – *constructional*, *definite*, and *indefinite* – each of which is discussed briefly here (see also section 2.4.2. of Atkins et al.'s article on lexicographic relevance in this volume).

Constructionally Null Instantiation (CNI) is licensed by a grammatical construction, as with the missing subject of an imperative sentence, which does not depend on the identity of any particular lexical item. Cases of CNI include: the omitted agent of passive sentences; the omitted subject of imperative sentences; the omitted subjects of independent gerunds and infinitives; and so on. For example, in *He was tied to the back of a horse* (sentence 2 in the subcorpus viewing section of Figure 12), we mark the absence of the core FE AGENT using the pull-down tab to the right of the frame element. We choose CNI since the missing FE is an omitted AGENT licensed by the passive construction. Notice that the symbol CNI appears (tagged in the color assigned to the FE) at the end of the sentence.

In addition, we use CNI for missing objects in instructional imperatives, as exemplified in (13), even though in this case the omission is not uniquely dependent on a particular construction but rather on a genre. The sentence below might be found in a cookbook.

- (13) **Tie** tightly with kitchen twine. (CNI ITEMS)

There are two types of lexically licensed null instantiation: Definite (or anaphoric) and Indefinite (or existential). With Definite Null Instantiation (DNI), the missing element must be understood in the linguistic or discourse context. In sentence (14) below, the *ITEM* is not expressed overtly in the sentence, but it is understood from the context. To be more specific, we understand that a boat-like entity was moored to the location.

- (14) Passing an oil rig and numerous naval vessels, we **moored** alongside a jetty on the estuary leading up to the town. (DNI *ITEM*)

Indefinite Null Instantiation (INI) covers missing objects of certain transitive verbs that are characterized as used intransitively, for example *eat*, *bake*, *drink*, and so on. The missing object of such a verb often has a special interpretation. For instance, the missing object of *eat* (*She hasn't eaten for three days*) is likely to be a meal, that of *bake* (*I spent the afternoon baking*) to be a flour-based product, and of *drink* (*I've stopped drinking*) to be an alcoholic beverage. In these cases, the semantic type of the missing object is understood through the conventional interpretation of this intransitive use; it is not necessary to retrieve a specific linguistic or discourse referent. In the *Attaching* frame, several LUs license the indefinite omission of the FE *CONNECTOR*, as illustrated in the following sentences from the FrameNet database. In each of sentences (15)–(18) below, we understand that a rope-like artifact was used for connecting the two entities. (15) Once he did, the man **bound** Corbett's hand tightly to the saddlebow. . . .

- (16) She rode up and **chained** her bicycle to tall railings.
 (17) She **sewed** buttons on to a white shirt she'd made.
 (18) It all began when a woman threatened to **tie** herself to a willow tree.

7.2. 'Gov-X' Annotation

While event nouns (e.g. *tying*, *connection*, *attachment*) evoke frames with the same kinds of event structures evoked by the verbs which semantically and morphologically underlie them, nouns that name artifacts and natural kinds (e.g. *skirt*, *chain*, *apple*, *giraffe*) generally do not. Rather, such nouns typically serve as slot fillers for frames evoked by verbs, adjectives or event nouns. Because we want to record information about (1) which frames certain entities typically appear in and (2) which entities fill the slots of particular frame elements, we provide annotations that identify typical governors of slot-filler nouns. We have noted the special affinity that certain nouns have with our *Attaching* frame, and we have compiled a separate list of them for a different sort of annotation in a *Connectors* frame, defined as in Figure 15. The annotation we refer to as 'Gov-X' selects the governing predicate (usually a verb) and marks it as governor, and labels the constituent containing the target word with 'X', with the intention that at some later stage when the frame for the governing word is itself annotated, the 'X' will be replaced with the name of the FE appropriate for

the governor's frame. In the meantime we are at least accumulating lexical information about typical governors: *tie*, *wrap*, etc., with *ribbon*, *cut*, *stab*, etc., with *knife*, and so on.

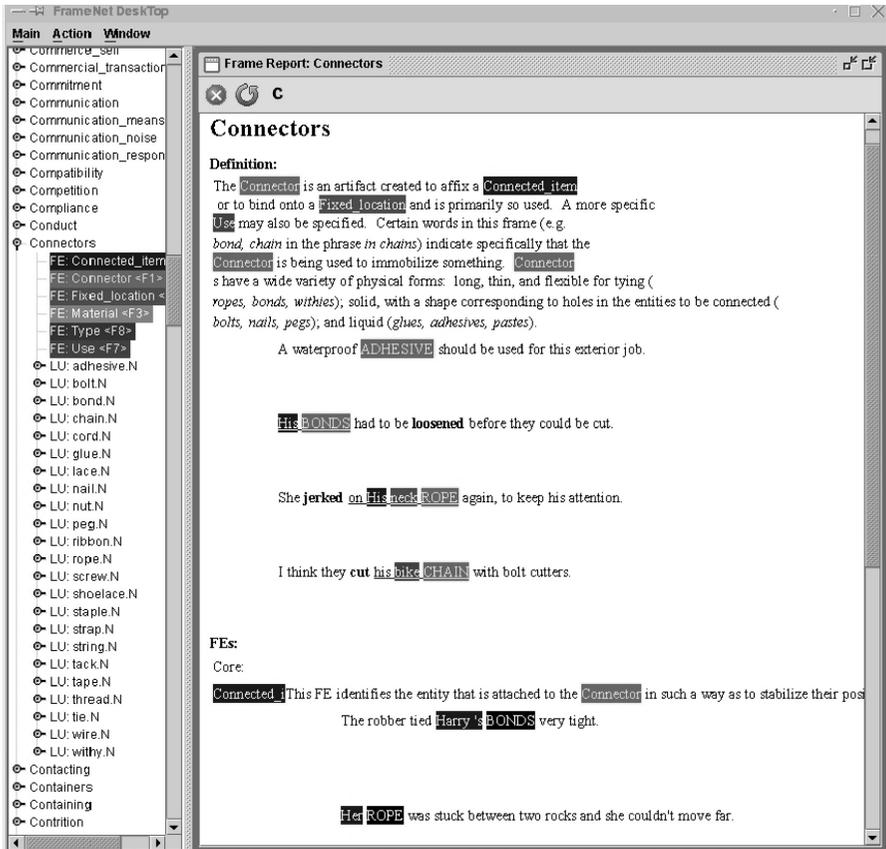


Figure 15: Connectors Frame Information Report

Subcorporation for the target nouns in the *Connectors* frame includes specific requests for sentences with several typical governing verbs, including *cut* and *tie*. Figure 16 shows the special annotation that we use for artifact nouns.¹⁴

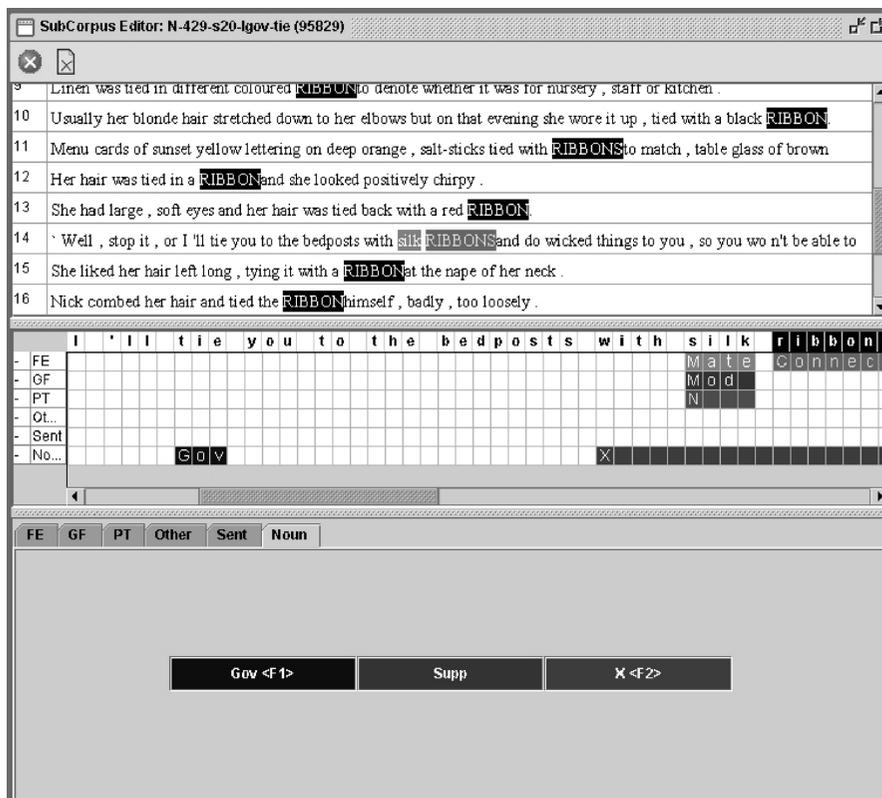


Figure 16: Gov-X Annotation

In contrast to annotation in event frames, here we annotate the target word without phrase type and grammatical function tags. We tag the governing verb (and particle for phrasal verbs, e.g. *cut up*), as well as the whole constituent containing the slot filler noun which we call 'X'. In addition we have also defined several frame elements that reflect a sort of qualia structure of the artifacts (see Pustejovsky 1995 for more information about qualia). For instance, we record information about the modifiers that denote the MATERIAL from which a connector is made (*constitutive quale*, as in *silk ribbon*) and the USE (*telic quale*, as in *hair ribbon*, *gift-wrap ribbon*) to which it is put.

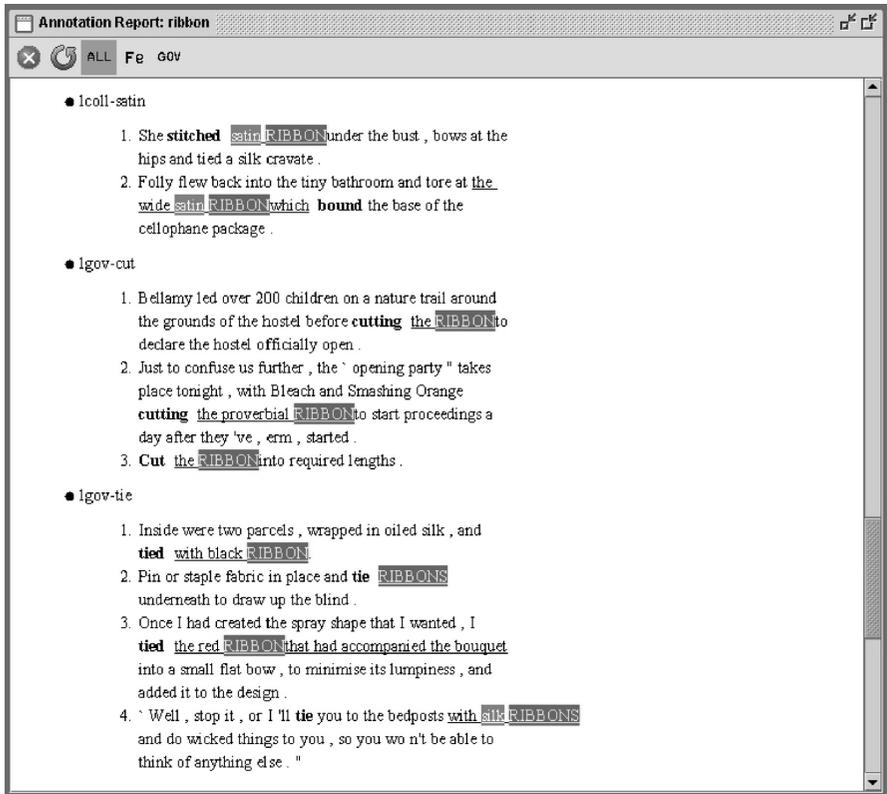


Figure 17: Annotation Report for *ribbon*

7.3. Support Expressions

Although in principle members of all major lexical categories can evoke a semantic frame, the dominant semantic frame of a sentence is usually evoked by the sentence's main verb. In some situations, however, it is a noun that provides the dominant frame; in fact, in certain styles of academic or political writing the dominant frame informing the meaning of the sentence is a noun. While the words in our *Attaching* frame do not welcome support expressions, we can nonetheless illustrate the phenomenon with a different sense of the noun *attachment*, specifically one that would be characterized in a frame concerned with emotional connections, and exemplified in the following sentences from the BNC.

- (19) ... and as a result we all became good friends without *forming* any sentimental **attachments**.
- (20) *Having* a sentimental **attachment** to them, I cannot resist mentioning...

- (21) A minority are unable to *make* any **attachments** at all and may need time in a therapeutic community ...

In each of these sentences, we understand that we are learning something about situations of emotional attachment, rather than acts of forming, having, or making. As a result, we want our annotation to be centered in the noun *attachment*, since it is the major frame-bearing word in each clause. We need to provide special treatment for the support verbs *form*, *have*, and *make* because the semantic information these words introduce is subordinate to the frame evoked by their syntactic object. Thus, we define *support verbs* as those verbs which, occurring in construction with a noun designating an event or state, result in a verb-like predicate with the frame that the noun belongs to; associated frame elements can either be in syntactic construction with the support verb (e.g., as its subject) or with the frame-bearing noun (e.g., as a prepositional or clausal complement). The support verbs are not meaningless, but they tend to serve registrational, aspectual or perspectivizing functions. In (20), above, the relationship is static; it is presented as merely existing. In (19) and (21) the relationship is coming into being (though in both cases negated). Compare these with (22), below, where we would not consider *promote* a support verb, since its subject is not a participant in the attachment relation.

- (22) ... the child's key worker draws up a renurturing program, to promote **attachment** or work on grief or identity.

Associating support verbs with event nouns is lexicographically relevant, and one of FrameNet's responsibilities is to record such associations. While there are some impressive regularities between individual support verbs and the semantic types of the nouns they accompany, it is generally the case that the noun selects its support verb, not the other way around. Moreover, different senses of a noun may be distinguished by which, if any, support verb it selects. While emotional *attachment* selects three different support verbs, recall that physical *attachment* (in our *Attaching* frame) tolerates none.¹⁵

Aside from support verbs, we also recognize support prepositions, which occur with certain nouns to produce an expression that is roughly equivalent to a predicative adjective, as in *at risk*, *in trouble*. As with support verbs, here too, the frame of the noun is dominant.

7.4. Criteria for Selecting Appropriate Sentences

The selection of sentences for annotation is guided by the following principles: frame-relevance; (relative) simplicity of structures; typicality of collocations; repetition avoidance; and preference for "world English" rather than expressions unique to UK English. Each principle is explained below.

7.4.1. *Frame-relevance*. The sentence chosen to exemplify the intended sense of the target lexical unit should show the word as clearly interpreted in its frame. In selecting sentences chosen to illustrate *rope* as the name of a connector, for

example, we would prefer a sentence about tying things up with rope over a sentence about, say, being interested in rope.

7.4.2. *Simplicity of structures.* One aspect of sentence complexity is length. Although matches of excessive length get eliminated by the corpus querying software automatically, there are still some sentences left that are quite long. FrameNet lexicographers are encouraged to avoid such sentences, on the grounds that they should not be needed to illustrate the basic grammatical behavior of the target LU. Another class of sentences that lexicographers avoid are those which dislocate the complements of the target word in ways that have nothing to do with the unique syntactic or semantic properties of the target. For example, in the invented sentence (23), of the kind generative linguists delight in composing, we find AGENT, ITEM and GOAL scattered throughout the sentence because of interrogative fronting, *find*-complementation, and *tough*-movement, in ways that tell us nothing interesting about the basic valence pattern of the target word *tie*.

(23) Which tree did you find the rope impossible to **tie** around?

7.4.3. *Typical collocations.* Although we make our data available with the disclaimer that the accompanying annotations do not constitute a statistically representative sample of the corpus, we nevertheless do seek to include sentences that we know show canonical uses of each target word. In choosing sentences with the *tie* of *Attaching*, we would surely want to include numerous examples with *string*, *cord*, *wire*, etc., used as connectors before including sentences in which something was tied in place with bacon strips.

7.4.4. *Repetition avoidance.* We avoid choosing multiple sentences of similar syntactic structure. Our main effort is to document the variety of syntactic constellations that a target word can enter into when used in a given frame, and, except for including a variety of frame-associated collocates, nothing would be gained by having dozens of examples of essentially the same syntactic pattern.

7.4.5. *World English.* We avoid choosing example sentences that are specifically British. We want our data to be equally useful to users in all parts of the English-speaking world.

8. Lexical Unit Reports

Section 7 explained our annotation procedure and software. Once annotation is complete, however, we need a tool for examining annotated sentences and the valence patterns they manifest. We have two such web-based reports, which are available internally as well as via the public web page; these are the Annotation by LexUnit and Lexical Entry Reports. These reports are automatically generated based on annotations completed using the annotation software.

The Annotation by LexUnit Report shown in Figure 18 displays all the annotated sentences for our lexical unit, *tie*.

Annotation Report: tie

ALL Fe Gov

tie.v

Agent
Connector
Goal
Handle
Instrument
Item
Items
Manner
Means
Place
Purpose
Result
Time

- roll-horse
 1. Apparently the healer would TIE a black thread round the horse's ankle, and it usually worked.
 2. He was TIED to the back of a horse and taken to Edinburgh, where he was tried, convicted and hanged. CNI
 3. Sharpe TIED the horse's reins to the harrow's stout frame, then slid his rifle out of its saddle holster.
 4. We TIED our horses to a small rail and knocked on the priest's house door.
- roll-tree
 1. Caspar was TIED to a tree with wire fixed to Philip's belt. CNI
 2. It all began when a woman threatened to TIE herself to a willow tree.
 3. Then Sapt TIED the rope round one of the trees near the water.
 4. He TIED his horse to a tree outside this bushy patch and left him grazing.

Figure 18: Annotation Report for *tie*

At the top of the report is a list of all the frame elements found within the frame; this is called the Frame Element Table. As displayed in Figure 18, the Attaching Frame has frame elements such as HANDLE, AGENT, ITEM, RESULT and so on. In a color display, these are seen highlighted in different colors. At the top of the report are three options: to display both frame element layers and Gov-X annotation (the latter was discussed in section 7.2); to show the frame element layers only; and to show Gov-X annotation only. The annotated sentences are displayed with the words instantiating frame elements highlighted in the same color in which the frame elements appear in the Frame Element Table. Thus, by matching the color highlighting a word or phrase with a color in the Frame Element Table, the user can tell which frame element the word or phrase instantiates. The sentences are grouped according to the subcorpus in which they appeared in the Annotator. Recall that subcorpora were defined during the subcorporation stage and pick out specific syntactic patterns and collocates which maximize the number of in-frame corpus examples selected for annotation. For example, in the first sentence in the subcorpus 'roll-

horse' (which means *tie* has a right collocate which is *horse*) *the healer* is the AGENT, *tie*, shown in capital letters, is the target for annotation, *a black thread* is the ITEM and *around the horse's ankle* is the GOAL.

The Lexical Entry Report summarizes the syntactic realization of the frame elements and the valence patterns of the lexical unit in two tables. The two parts of this report are illustrated in Figure 19 and Figure 20, respectively, for our lexical unit, *tie*.

The screenshot shows a window titled "Lexical Entry Report: tie" with a toolbar containing a close button, a refresh button, and a button labeled "C Fr". The main content area displays the following information:

tie.v

Frame: Attaching

Definition
COD: attach or fasten with string, cord, etc FN:Means:Knot_creation.tie.v

Frame Elements and Their Syntactic Realizations
The Frame elements for this word sense are (with realizations):

Frame Element	Number Annotated	Realizations(s)
Agent	27 exx	NP.Ext 20 exx --- 7 exx
Connector	5 exx	PP.Comp 5 exx
Goal	24 exx	PP.Comp 24 exx
Item	24 exx	NP.Ext 5 exx NP.Obj 19 exx
Items	3 exx	NP.Ext 1 exx NP.Obj 2 exx

Figure 19: Lexical Entry Report for *tie*, Realization Table

At the very top of the report two options appear. The first button, labeled 'C', allows one to control whether all frame elements appear or only core frame elements appear; the default for this report is to only display the core frame

elements. The second button, labeled 'Fr', allows one to choose the option to view all the annotated sentences, even those marked as belonging in a different frame.¹⁷

The table in Figure 19 shows all of the core frame elements, the number of annotated examples and their syntactic realization. So, for example, AGENT appeared a total of 27 times, 20 of those times as a noun phrase external argument (subject) and seven times as null-instantiated. CONNECTOR only occurred five times and always as a prepositional phrase complement of the verb. Clicking on the link indicating the number of examples displays the sentences which illustrate the syntactic pattern for that frame element.

Lexical Entry Report: *tie*

C Fr

Valence Patterns:

These frame elements occur in the following syntactic patterns:

Number Annotated	Patterns			
2 exx TOTAL	Agent	Connector	Goal	Item
1 exx	--	PP	PP	NP
	--	Comp	Comp	Ext
1 exx	NP	PP	PP	NP
	Ext	Comp	Comp	Obj
3 exx TOTAL	Agent	Connector	Items	
1 exx	--	PP	NP	
	--	Comp	Ext	
2 exx	NP	PP	NP	
	Ext	Comp	Obj	
22 exx TOTAL	Agent	Goal	Item	
4 exx	--	PP	NP	
	--	Comp	Ext	
1 exx	--	PP	NP	
	--	Comp	Obj	
17 exx	NP	PP	NP	
	Ext	Comp	Obj	

Figure 20: Lexical Entry Report for *tie*, Valence Table

The second table, shown in Figure 20, illustrates the valence possibilities for *tie*. By valence possibilities we mean the various combinations of frame elements and their syntactic realizations which might be present in a given sentence. For instance, there are two examples where the AGENT is realized as a noun phrase external argument, the ITEM is realized as a noun phrase object, the GOAL is realized as a prepositional phrase complement, as is the CONNECTOR. Clicking on the link indicating the number of examples will display the sentences which manifest that particular valence possibility.

Our web-based Lexical Unit reports allow a visually appealing and useful way to access Frame Net annotations and view a summary of the valence possibilities extracted from these annotations.

9. Conclusion

The preceding pages have given a sketch of the FrameNet process. We begin with a schematic description of a scene which seems to relate a group of words and end up with finished reports detailing the valence and frame element realization possibilities for each lexical unit within our now well-defined frame. For the purposes of this paper, we have followed the verb *tie* in the Attaching frame through the various stages of the FrameNet process as illustrated in the workflow diagram in Figure 1. As noted in the introduction, the progression through the stages is not strictly linear. That is, while we are annotating a lexical unit we may discover important facts about its frame or one of its frame elements. For example, as is frequently the case, during the annotation process we may discover the need for a new frame element that was not anticipated or the need to expand our definition of a frame element. This is, of course, the benefit of a corpus-driven approach to lexicography and one we fully exploit.

Our description of the FrameNet process was aimed to reveal what is so rewarding and exciting about FrameNet: FrameNet's lexical descriptions and their accompanying annotated sentences rely fundamentally on Frame Semantics; they are the product of much careful thought and consultation with other lexicographic resources; each sense of a lexical unit is connected to a well-defined Frame; and the lexicographic descriptions are based on a vast amount of corpus data. Thus, by virtue of its theory-driven but data-rooted nature, FrameNet offers a depth of semantic information and a breadth of examples not found in other lexicographic projects.

For more information about the FrameNet database and Frame Semantics' theoretical background please consult the other articles within this volume. To view more of our data, visit our website where the Annotation and Lexical Entry Reports for a large number of lexical units are available for public view.

Notes

¹ FrameNet-local terminology distinguishes *word forms*, *lexemes*, and *lemmas*. A *word form* is an inflectional variant. A *lexeme* is a word in a given part of speech instantiated by one or more word forms: thus the lexeme *bring* has forms *bring*, *brings*, *bringing*, *brought*, whereas the lexeme *up* has only the single word form *up*. A *lemma* is a unit made up of one or more lexemes seen as bearing one or more senses: thus the lemma *bring up* consists of the lexemes *bring* and *up* and supports several *lexical units*, among which one has to do with caring for a child to maturity, another with introducing a topic into a conversation. (We use *word* promiscuously for any of these.) It is possible, thus, for a lexeme to have a single word form, for a lemma to be made up of a single lexeme, and for a lemma which has only one sense to have one unique lexical unit.

² The Corpus Workbench was made available to us from the Institut für Maschinelle Sprachverarbeitung of the University of Stuttgart through Ulrich Heid and Oliver Christ.

³ The British National Corpus, information is available at <http://www.hcu.ox.ac.uk/BNC/>.

⁴ XML ('Extensible Markup Language') is widely used as a format for the exchange of data between different computer systems, programs, etc. See <http://www.xml.com> for more information.

⁵ Numbered examples that appear in square brackets refer to sentences from KWIC results in Figure 2. All other examples in this paper are numbered separately; these numbers appear in parentheses.

⁶ The phrase that we label as instantiating a given frame element includes whatever grammatical markers serve to set it off as a syntactic complement of the head word: as seen here, that includes the prepositions *to* and *around*.

⁷ The shading in the figure conceals the different colors that represent, in the original, the distinct frame element types and the phrases that exemplify them.

⁸ Confusingly, here the word *frame* refers to a section of a FrameNet Desktop.

⁹ Semantic Type features associated with FEs will indicate the kinds of semantic entities that are their typical realizations (e.g. animate, abstract, concrete, etc.), and when associated with lexical units can indicate their semantic types; this information will be useful in NLP applications.

¹⁰ This awkward term, which over the years has come to seem perfectly normal to us, refers to the process of creating syntactically or collocationally delimited subsets of the concordances for a given target word.

¹¹ The phrasal patterns are suggested by the categorial information and the example 'kiss me' reminds the lexicographer to decide whether to include sentences in which the verb is immediately followed by a noun phrase.

¹² Core vs. non-core status is not indicated in the annotation tool.

¹³ The GF label Ext roughly covers, for verbs, the notion of surface subject, in finite clauses; but we also use it for labeling the controlling NP in various predictive complements. *John* is the Ext of *cut* in *John decided to cut the rope* but also in *We made John cut the rope*.

¹⁴ Gov-X annotation is also done for transparent nouns, including those that name aggregates, parts, types, etc. Often such nouns occur as the first noun of a N-of-N construction in contexts where the governing verb selects the second noun semantically, rather than the syntactic head. For example, while a sentence containing the phrase *piece of ribbon* would be annotated with respect to the target word *piece* in the Part_piece frame, FrameNet lexicographers would also tag a typical governing verb such as *cut* or *tie*.

¹⁵ On support constructions see Mel'čuk 1996, Fillmore & Sato 2002, Ruppenhofer et al. 2002.

¹⁶ This is a purely technical issue. Due to certain limitations of the subcorporation and annotation software, sentences may have been annotated in the wrong frame and left to be moved into the right frame in the future.