INTERNATIONAL COMPUTER SCIENCE INSTITUTE

1947 Center St. • Suite 600 • Berkeley, California 94704-1198 • (510) 666-2900 • FAX (510) 666-2956

FrameNet: Theory and Practice

Christopher R. Johnson
Charles J. Fillmore
Miriam R. L. Petruck
Collin F. Baker
Michael Ellsworth
Josef Ruppenhofer
Esther J. Wood

TR-02-009

Version 1.0
September 13, 2002

Abstract

Describes Frame Semantics as applied in the FrameNet project, what is annotated and why, how annotators deal with missing or conflated frame elements, and the differences in annotating sentences with verb, noun or adjective target words. Explains the phrase types and grammatical functions used in FrameNet annotation, and briefly describes lexical entries and frame-to-frame relations.

Preface

FrameNet is a lexicographic project housed and administered at the International Computer Science Institute (http://www.icsi.berkeley.edu), in Berkeley, California. We are grateful to the National Science Foundation for funding the work of the project through two grants, IRI #9618838 "Tools for Lexicon Building" March 1997–February 2000, and ITR/HCI #0086132 "FrameNet++: An On-Line Lexical Semantic Resource and its Application to Speech and Language Technology" September 2000–August 2003. (We refer to these two three-year stages in the life of the project as FrameNet I and FrameNet II.)

We also wish to acknowledge the support of Oxford University Press, which, through Timothy Benbow, made it possible for us to use the British National Corpus as the evidential basis for our inquiry into the behavior of English words, and through Robert Scriven, gave us permission to select definitions from the Concise English Dictionary to serve as as parts of the FrameNet I lexical entries.

We are also indebted to the Institut für Maschinelle Sprachverarbeitung of the University of Stuttgart, through the kindness of Ulrich Heid and Oliver Christ, for the IMS Corpus Workbench software, which made it easy for us to extract, explore, and sort example lines and sentences from the BNC.

The Project is administered at the International Computer Science Institute: we are indebted to Jerome Feldman, Director of the Institute during most of the FrameNet I period - and a year and a half before our initial funding - and to Nelson Morgan, current ICSI Director, for their efforts in making available to us the services and facilities of this remarkably welcoming and well-run research community.

Charles J. Fillmore

September 10, 2002

PI:	Charles J. Fillmore
Head Programmer:	Beau Cronin
Project Manager:	Collin Baker
Software Designer:	Charles Wooters
GUI Developer:	Hiroaki Sato
Consultants:	B. T. Atkins
Techies:	Collin Baker, Beau Cronin, Olya Gurevich, Carol Hays, Hiroaki Sato, Wendy Wooters
Lexicographers:	Caitlin Cota, Ellen Dodge, Michael Ellsworth, Olya Gurevich, George Kawamoto, Russell Lee-Goldman Miriam Petruck, Josef Ruppenhofer, Susanne Stadtlebauer Eugenia Teytelman, Gloria FanBei Yang, Esther Wood, Abby Wright.

Contents

1 Introduction to the Project **7**

2 FrameNet Annotation **9**
- 2.1 Introduction . . . 9
- 2.2 Annotation with verbs as targets . . . 10
- 2.3 Annotation with nouns as targets . . . 16
- 2.4 Annotation with adjectives as targets . . . 22
- 2.5 Annotation relative to slot fillers . . . 22
- 2.6 Text annotation . . . 24

3 Identifying Phrase Types **25**
- 3.1 List of phrase types . . . 25
- 3.2 Phrase Type Labels for Noun Phrases . . . 27
- 3.3 Phrase Type Labels for Prepositional Phrases . . . 29
- 3.4 Phrase Type Labels for Verb Phrases . . . 32
- 3.5 Phrase Type Labels for Clauses . . . 33
- 3.6 Phrase Type Labels for Adjective Phrases . . . 35
- 3.7 Phrase Type Labels for Adverb Phrases (AVP) . . . 36
- 3.8 Phrase Type Labels for Numerals (NUM) . . . 37
- 3.9 Phrase Type Labels for Quotes (QUO) . . . 37

4 Assigning Grammatical Functions **39**
- 4.1 List of Grammatical Functions . . . 40
- 4.2 Assigning GFs for Verbs . . . 41
- 4.3 Assigning GFs for Adjectives . . . 44
- 4.4 Assigning GFs for Prepositions . . . 46
- 4.5 Assigning GFs for Nouns . . . 47

5 Lexical Entry Structure **49**
- 5.1 Lexical Entries . . . 49
- 5.2 Incorporated Frame Elements . . . 50

6 Semantic Relations **51**
- 6.1 Frame-to-frame Relations . . . 51
- 6.2 Semantic Type . . . 53

A Publications related to FrameNet **55**

Index **59**

Typographic Conventions

We use the following typographic conventions in this text.

- The first mention of a technical term appears in **bold face**:

  A **semantic frame** is a script-like structure.

- Data not set off from the text appear in *italics*:

  Note that *the children take naps* is not treated as a clause.

- In the text, names of Frame Elements are capitalized:

  A Speaker communicates a Message to an Addressee in some Medium.

- In example sentences set off from the text, target words are in **bold face**:

  Bob **told** a story.

- Constituents which represent frame elements are in square brackets:

  [Bob] **told** [a story].

  This shows that *Bob* and *a story* are elements in the frame evoked by the target word *told*.

Chapter 1

Introduction to the Project

The Berkeley FrameNet project is creating an online lexical resource for English, based on **frame semantics** and supported by corpus evidence. The aim is to document the range of semantic and syntactic combinatory possibilities (valences) of each word in each of its senses, through manual annotation of example sentences and automatic capture and organization of the annotation results. The FrameNet database is in a platform-independent format, and can be displayed and queried via the web and other interfaces.

A 'starter lexicon' became available to the public in May, 2001, and contained approximately 2000 items – verbs, nouns, and adjectives – representative of diverse areas of the vocabulary. The analysis was supported by about 40,000 annotated sentences. At the time of the Fall, 2002 release, the database includes approximately 6,000 lexical items, for which there are about 130,000 annotated sentences.

A **lexical unit** is a pairing of a word with a meaning. Typically, each sense of a polysemous word belongs to a different **semantic frame**, a script-like structure of inferences that characterize a type of situation, object, or event. In the case of **predicates** or **governors**, each annotation accepts one word in the sentence as its **target** and provides labels for those words or phrases in the sentence which fill in information about a given instance of the frame. These phrases are identified with what we call **frame elements (FEs)** – participants and props in the frame whose linguistic expressions are syntactically connected to the target word. A frame semantic description of a predicating word, derived from such annotations, identifies the frames which underlie a given meaning and specifies the ways in which FEs, and constellations of FEs, are realized in structures headed by the word. In the case of **dependents**, most typically nouns, the annotations serve to identify the most common predicates that govern phrases headed by them, and thus to illustrate the ways in which these targets function as elements within frames evoked by those predicates. Formally, these are provided as constellations of triples that make up the frame element realization for each annotated sentence, each triple consisting of a frame element (say, Patient), a grammatical function (say, Object) and a phrase type (say, NP). **Valence descriptions** of predicating words are generalizations over such structures.

The annotated sentences are the building blocks of the database: marked up in XML, they form the data from which the lexical entry descriptions are derived. This format supports searching by lexical unit, frame, frame element, and combinations of these.

The FrameNet database serves both as a dictionary and a thesaurus. The dictionary

features include definitions (from the Concise Oxford Dictionary, 10th Edition, courtesy of Oxford University Press, or a definition written by a FrameNet staff member), tables showing how frame elements are syntactically expressed in sentences containing each word, annotated examples from the corpus, and an alphabetical index. Like a thesaurus, words are linked to the semantic frames in which they participate, and hence to the other words which evoke those frames; and frames in turn are related to other frames.

The main FrameNet corpus is the 100-million-word British National Corpus (BNC), though several current subprojects are using U.S. newswire texts made available through the Linguistic Data Consortium. The semantic annotation is carried out using software (called "Annotate") developed in-house. The syntactic annotation, which adds grammatical function and phrase type to each annotated phrase, is handled by an in-house tagging program.

The project's deliverables will consist of the FrameNet database itself:

- lexical entries for individual word senses
- descriptions of frames and frame elements, and
- annotated subcorpora

The documentation for the project is distributed in two parts. One is the document you're reading now. It contains discussion and description of frame semantic theory and practice. The second document contains the frame and frame element descriptions contained in our database. Both documents can be downloaded from our web-site: `http://framenet.icsi.berkeley.edu/~framenet/`

Editing, annotating, and viewing software are currently being combined into a single software suite which will be released to researchers fairly soon. The adaptability of the FrameNet software to a different language has been demonstrated in the Spanish FrameNet project headed by Carlos Subirats. Researchers interested in obtaining tools for doing similar annotation work should contact the FrameNet Project directly.

Chapter 2

FrameNet Annotation

2.1 Introduction

This chapter describes what we annotate and how we carry out our annotation task.

Before we start let us briefly consider the Revenge frame, which will serve as our example frame throughout this chapter. The definition of this frame is as follows:

> An AGENT performs a RESPONSE_ACTION on an VICTIM as a punishment for an earlier action, the INJURY, that was inflicted on an INJURED_PARTY. The AGENT need not be identical to the INJURED_PARTY but needs to consider the prior action of the VICTIM a wrong. Importantly, the punishment of the VICTIM by the AGENT is seen as justified by individual or group opinion rather than by law.

Lexical units in this frame include *avenge.v, avenger.n, get even.v, retaliate.v, retaliation.n, retribution.n, retributive.a, retributory.a, revenge.v, revenge.n, revengeful.a, revenger.n, vengeance.n, vengeful.a*, and *vindictive.a*. Some example sentences with the lexical unit *avenge* are given here.

> [His brothers AGENT] **avenged** [him INJURED_PARTY].
>
> With this, [El Cid AGENT] at once **avenged** [the death of his son INJURY].
>
> [Hook AGENT] tries to **avenge** [himself INJURED_PARTY] [on Peter Pan VICTIM] [by becoming a second and better father RESPONSE_ACTION].

Returning now to the topic of annotation, there are several characteristics to note.

- FrameNet annotation is always done relative to one particular linguistic unit, the target, which is most often a single word but can also be a multi-word expression such as a phrasal verb (e.g., *give in*) or an idiom (e.g., *take into account*).
- We annotate whole constituents rather than just the head words of the target's syntactic dependents. For instance, in the third of our introductory examples above, the preposition *on* is included in the constituent expressing the VICTIM frame element.

- The dependents are annotated for Frame Element identity, Grammatical Function, and Phrase Type relative to the target word.

- We annotate sentences rather than running text.

- Sentences are preselected because they contain a predetermined target word.

- We seek to exemplify the whole range of combinatorial possibilities that the target lexical unit exhibits.

- The data is meant to be lexicographically relevant, not statistically representative.

Against the backdrop of these general principles, we produce annotation of two different types, reflecting the two different kinds of target words we have. There is:

- annotation relative to a **frame bearing** syntactic governor, either a predicate, modifier, or a referring expression and

- annotation relative to a **slot filler**, that is, relative to a referring expression that is a frame element of a frame determined not by itself but by a governor

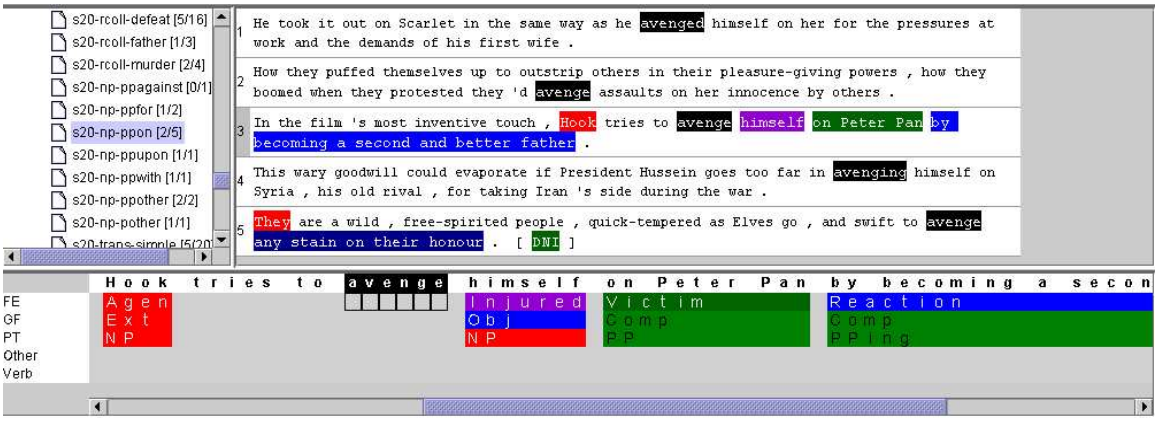
Both kinds of annotation will be discussed. Since the annotation relative to syntactic governors is the main kind we do, we will discuss it first.

2.2 Annotation with verbs as targets

In principle, members of all three major lexical categories (verb, noun, adjective) can be frame bearing, that is, to evoke a semantic frame. However, the most prominent semantic frame evoked in a particular sentence is usually one evoked by a verb. For instance, the example sentences given in the introduction are centrally concerned with acts of Revenge and in each case the Revenge frame is brought into play by the meaning of the verb *avenge* rather than by the denotations of the noun phrases. Therefore, we will start by considering annotation relative to a verbal syntactic governor.

2.2.1 Easy cases

Annotation is easiest when all and only the **core** frame elements, that is the conceptually necessary participants of the frame that a syntactic governor evokes, find syntactic expression as separate immediate syntactic dependents of the governor. Under such circumstances, we simply annotate the syntactic dependents for the three kinds of information mentioned earlier: Frame Element role, Grammatical Function, and Phrase Type. The various kinds of information are recorded on separate layers, as shown below, where three of the layers are labeled FE, GF, and PT.



The annotators actually only need to apply a Frame Element label; a Grammatical Function and Phrase Type are derived algorithmically by a chunk parser but may require manual correction.

In addition to core frame elements, we also provide annotation for **non-core** frame elements expressed in the sentences we select for annotation. Non-core frame elements are conceptually not necessary in the sense that they don't make the frame uniquely what it is. For instance, TIME and PLACE expressions can occur with most states and events but (almost) no frame is differentiated from other frames by the fact that the event or state it talks about is located in space or time. Similarly, actions often have a PURPOSE the AGENT intends to accomplish by performing the action indicated by the target word. Again, having a purpose doesn't usually distinguish actions amongst themselves and so PURPOSE is often a non-core frame element, as in the case of *give* in the following example:

> When I accepted he suggested I might like to **give** him a certain sum [to help the wine flow PURPOSE].

Grammatically, we can notice that non-core elements cannot be nuclear arguments, that is subject and object, of target verbs and that they frequently are adverbs or prepositional phrases.

Unfortunately, not all sentences that we encounter are as straightforward to annotate as the ones we have seen so far. We will now discuss how we deal with various challenging cases.

2.2.2 Missing frame elements

Sometimes FEs that are conceptually necessary do not show up as lexical or phrasal material in the sentence chosen for annotation. Nevertheless, we indicate their absence as it provides lexicographically relevant information regarding **omissibility conditions**. The FE tagged indicates which semantic role the missing element would fill, if it were present.

Not all cases of frame element omission are alike. We recognize three different cases, which are recorded with the annotation tool. Notice the tab to the right of the frame element VICTIM in the above picture: it shows that the FE was omitted under definite null instantiation. We will now discuss the three types of omission, starting with the type that is not lexically specific and continuing with the two types that are.

Constructional Null Instantiation Constructionally omitted constituents (also called *structurally* omitted) have their omission licensed by a grammatical construction in which the target word appears and are therefore more or less independent of lexically specific information. Cases of CNI include: the omitted subject of imperative sentences; the omitted agent of passive sentences; the omitted subjects of independent gerunds and infinitives (PRO); and so on. The symbol that we use for constructionally null instantiated constituents is **CNI**. In each of the following two examples, the FE Agent is tagged with the symbol CNI.

> Family feuds last for generations, and [slurs on honor INJURY] are **avenged** [by murder RESPONSE_ACTION] . [CNI AGENT]
>
> **Get even** [with her VICTIM] [for this INJURY] [CNI AGENT]

In addition, we use CNI for missing objects in instructional imperatives such as exemplified below, even though in this case the omission is not uniquely dependent on a particular construction but rather on a genre.

> **Cook** on low heat until done. [CNI FOOD]
>
> **Tie** together loosely. [CNI ITEMS]

Definite Null Instantiation We now come to the first lexically specific type of null instantiation. Cases of definite (also called *anaphoric*) instantiation are those in which the missing element must be something that is already understood in the linguistic or discourse

context. In the following example, the VICTIM is not expressed overtly in the syntax, but its referent has to be known in the context. The symbol that we use for definite null instantiation is **DNI**.

> [The monkey AGENT] **avenged** [himself INJURED_PARTY] [by growing to the size of a giant and setting fire to the city RESPONSE_ACTION]. [VICTIM DNI]

Indefinite Null Instantiation The indefinite cases (sometimes also referred to as *existential*) are illustrated by the missing objects of verbs like *eat*, *sew*, *bake*, *drink*, etc., that is, cases in which these ordinarily transitive verbs can be spoken of as **used intransitively**. As is well known, there are often special interpretations of the existentially understood missing objects: for example, with **eat** the missing entity is likely to be a meal, with **bake** it is likely to be flour-based foods, with **drink** it is likely to be alcoholic beverages, etc. However, the essential difference between indefinite/existential and definite/anaphoric omissions is that with existential cases the nature or semantic type of the missing element can be understood given interpretational conventions, but there is no need to retrieve or construct a specific discourse referent.

In the Revenge frame, all lexical units happen to allow the frame element RESPONSE_ACTION to be omitted under indefinite null instantiation.

> He took it out on Scarlet in the same way as [he AGENT] avenged [himself INJURED_PARTY] [on her VICTIM] [for the pressures at work and the demands of his first wife INJURY]] . [INI RESPONSE_ACTION]

Note that usually verbs in a frame differ in this respect. For instance, while *eat* allows its object to be omitted, *devour* does not, even though they are both in the Ingestion frame.

Although we do have the symbol **INI** available for tagging frame elements that are omitted under indefinite null instantiation, we do not normally use it. Instead we ask users of our database to consider core frame elements as omitted under INI if they are neither instantiated nor accounted for by one of the other two kinds of null instantiation.

2.2.3 Undefined frame elements

Undefined frame elements represent the opposite problem from omitted arguments: there are elements present in the sentences selected for annotation that are not currently defined as frame elements but which annotators regard as likely to be elements in a revised and improved version of the frame. Eventually, we will replace all such labels with normal FE labels or to remove them.

2.2.4 Expletives

Some syntactic constructions require the presence of non-referential material in an argument position even though the non-referential material has no semantic relationship to the predicate. In a subset of cases this happens while a semantic argument is dislocated to a non-canonical position. The non-referential items appearing in such constructions are called expletives. *It* and *there* are the two kinds of non-referential NPs in English. They are never

frame elements and so are not not given GF and PT tags. We do, however, record their presence with the **Null** tag on the Other Layer. Some typical examples of expletives are given in the following examples.

Subject extraposition

[It NULL] is **clear** that we won't finish on time.

Object extraposition

I **hate** [it NULL] when you do that.

Existential construction

[There NULL] are more cookies in the jar.

Subject requirement of zero-place predicates

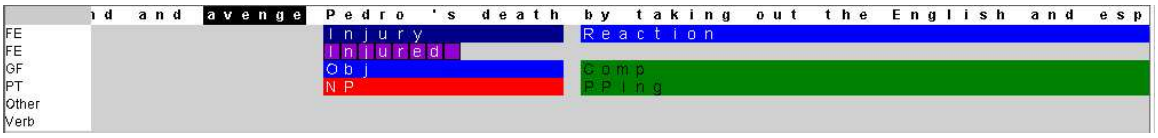
[It NULL]'s **raining**.

2.2.5 Frame element conflation

In some cases, information about two frame elements is expressed in a single constituent, a sitation we call **conflation**. For instance, the concept of ousting somebody from office requires an understanding of the (former) incumbent of the office and the identity of the office, and these both can be represented separately in a sentence like *We ousted Jones as mayor*. But in a sentence like *We ousted the mayor*, the direct object stands for both the office and the incumbent.
We also find examples of frame element conflation in the Revenge frame. In particular, the INJURED_PARTY may be contained as a possessive in the phrase that realizes the INJURY frame element as seen in the following example:

> [He AGENT] **avenged** [Pedro 's death INJURY] [by taking out the poker-faced Guards Officer RESPONSE_ACTION].

In this example, the possessive *Pedro's* realizes the frame element INJURED_PARTY, the person who suffered the INJURY. In such cases, the annotation tool allows for the creation of an additional FE layer, enabling the secondary annotation of (parts of) constituents in the same frame.



For camparison, consider our earlier example.

Note that conflation does not mean that the element on the secondary annotation layer cannot be instantiated as a constituent of its own in the syntax. In the following example, information about the person hit is encoded in the possessive of the NP complement of the preposition *on*, as well as in the direct object *me*.

He **hit** me on my hand.

Also note that there is never a phrase type or grammatical function indicated for the frame elements on the secondary FE layer.

2.2.6 Syntactic locality

As mentioned in the introduction to this chapter, our annotation is directed towards the dependents of a target word. We do not annotate constituents that we understand only through context as referring to the filler of a particular frame element role. An example of this is the following:

Smith was surprised when Lowry **retaliated** immediately for the attack.

Here, the target *retaliate* evokes the Revenge frame but the surface syntax does not contain constituents representing the VICTIM and the INJURED_PARTY. One plausible understanding of the sentence as a whole is that Smith was the one that Lowry retaliated against. One might thus want to tag *Smith* as the VICTIM. However, there is an alternative interpretation: Smith could just be an observer of a conflict involving Lowry and a third party. In cases like these, where there is no syntactic construction present that guarantees the interpretation we have in mind, we do not allow ourselves to annotate the non-local phrases in question. In our particular example, we would register the VICTIM as omitted under definite null instantiation (DNI).

In general, we try to select sentences for annotation where, with the exception of subjects, we find all frame elements realized by constituents that are part of the maximal phrase headed by the target word. In addition, there are several clear cases where, unlike in the above example, grammatical structure guarantees a particular interpretation of phrases that are not dominated by the maximal phrase headed by the target. In such cases, we choose to annotate the non-local constituents realizing frame elements. Some of the most common of the relevant constructions are illustrated below. For a more comprehensive list of contexts, see the appendix.

Raising

Subject to object
We expect [John AGENT] to **retaliate** [against us VICTIM] [INI RESPONSE_ACTION] [DNI INJURY].

Subject to subject
[John AGENT] seems to have **avenged** [the death of his brother INJURY] [by luring Smithers into a trap RESPONSE_ACTION].

Control

Subject control
[They AGENT] are hoping to **get even** [with Smithers VICTIM] [for the insult INJURY].

Object control
The commander ordered [the troops AGENT] not to **retaliate** [against the rebels VICTIM]

Tough-movement

[The defeat INJURY] was difficult to **avenge**. [CNI AGENT]

2.2.7 Aspect

When verb targets co-occur with particles that are used productively to indicate aspectual information, we tag the particle with the label **aspect** on the POS layer. Some examples follow.

They were **chattering** [away ASPECT] in the kitchen, when the door bell rang.

Mo **talked** [on and on ASPECT].

Particles that seem aspectual but form entrenched lexical units with the verb are not treated in this way. For instance, in examples such as the following *pull through* would be treated as one lexical unit, rather than as an instance of the verb *pull* accompanied by aspectual *through*.

It's amazing what they can do these days and luckily the sick dog **pulled through**.

2.3 Annotation with nouns as targets

FrameNet also annotates relative to noun targets. In the course of our work, we have been paying attention to several types of noun: nouns that denote events such as *withdrawal* and *replacement*; relational nouns such as *brother* or *girlfriend*; artefact nouns such as *house* and *vest*; and some others. Of these different types, it is the event nouns and the relational nouns that most clearly are frame-bearing. Consider some examples from the Revenge frame.

The Americans must have felt as if he was taking **revenge** on them for what had happened.

King Menephta took awful **revenge** on a Libyan army he defeated around 1300 BC

In both of these examples, we have the impression that what is reported is an act of revenge rather than an act of taking, the frame evoked by the noun *revenge* clearly being more dominant. Thus, nouns like *revenge* are legitimate objects for annotation as frame-bearing targets. Many of the issues we discussed earlier for verbs, e.g. undefined or conflated frame elements, apply to nouns in the same way as to verbs. However, the annotation of nouns also brings some new challenges to the annotation enterprise, which we now discuss in turn.

2.3.1 Support expressions

When we discussed the examples at the end of the previous section, we noted that the noun *revenge* in each case provides the dominant frame as opposed to the verb *take*. However, we would also like to give verbs like *take* a special treatment as **support verbs**. The intuition behind this move is basically that support verbs do not introduce any semantics of their own. Sentences with support verbs are about the same event or state denoted by the noun occurring by itself. We thus define support verbs as those verbs that turn a target event or state noun into a verb phrase-like predicate, allow for the expression of a Frame Element as their subject, and are semantically neutral (to the degree that that is possible). In the examples above, the subject of *take* realizes the AGENT frame element, and *take* contributes no semantics: all that is talked about is an act of revenge. In contrast, *encouraged* significantly changes what the sentence is about in the following example.

The Prime Minister encouraged **revenge** against Absurdistan.

Encouraging somebody to do something is different from doing it and the encourager is possibly no participant at all in the encouraged act. We thus would not consider *encourage* a support verb. Similarly, if the verb in the preceding example had been *called for*, we would not consider it a support verb.

Recognizing support verbs not only allows us to annotate their subjects as frame elements. It is actually also lexicographically necessary to record them given that it is the support verbs that are selected by the noun rather than the other way around. Support verbs also have other interesting properties that we want to be able to study. Some vary with the sense of the noun, that is, a noun may take different support verbs depending on the frame it belongs to. For instance, in the first of the following examples the noun *argument* takes the support verb *have*, and has a meaning related to conversational exchange; in the second example, the noun takes the support verb *make*, and has a meaning related to reasoning.

John and I [had SUPP] a terrible **argument** last night.

John [made SUPP] a convincing **argument** that the project should be funded.

Some support verbs have registral associations. In the Revenge frame, for instance, *wreak* clearly belongs to a higher, more formal register than *take* and *have*.

A short while later Joseph [had SUPP] his **revenge** on Watney 's.

The Australians magnificently [took SUPP] **revenge** for their defeat.

He [wreaked SUPP] a terrible **revenge** on them for their betrayal.

When de Raimes had [exacted SUPP] his **revenge** he would hand her over to his knight, who would complete her destruction.

In addition to support verbs we recognize a second type of support expression, namely **support prepositions**. Support prepositions combine with certain target nouns to yield a phrase that is more or less equivalent to a predicative adjective.[1] As in the case of support verbs, the frame of the noun is dominant and it is the noun that selects the particular support preposition rather than the other way around. Notice that in the case of target noun with support preposition, we also tag as support expressions the copular verbs that combine with them to form a verb phrase.

The painting [is SUPP] [on SUPP] **loan** from Mr Smithers

About 650 oil wells [are SUPP] [on SUPP] **fire**.

This old book [is SUPP] now [in SUPP] my **possession**.

The documents [came SUPP] [into SUPP] my **possession** by inheritance.

Notice also that there need not be a copular verb for a target noun occurring with a support preposition. When the combination of support preposition and noun is used as a depictive or resultative secondary predicate, there will be no copular verb. Here are some illustrative examples from the Facial_expression frame.

His face was twisted [in SUPP] a **grimace**.

He wrinkled his brow [into SUPP] a **frown**.

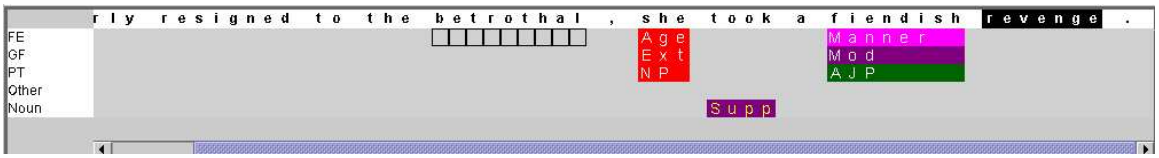
Such sentences are usually roughly parallel to sentences with a support verb. For the two preceding examples, *make* would be a good candidate.

His face was twisted [making SUPP] a **grimace**.

He wrinkled his brow [making SUPP] a **frown**.

Finally, notice that in accordance with their status as syntactic augments for target nouns, support expressions are not frame elements. They are tagged on their POS-Layer (part-of-speech) and have no grammatical function or phrase type assigned to them.

[1] We might want to recognize another possibility: a preposition may turn a noun into an adverb, e.g. *at X's risk/peril*



2.3.2 Copula and Copular sentences

Our use of the term **copula** is highly idiosyncratic and runs counter to grammatical tradition so careful explanation is in order. Usually, the term copula or linking verb is used for the verb *be* (and a few others) when they occur in constructions of the form: NP_1 *V* NP_2. [2]

Examples include:

John **is** a sailor.

Smithers **is** the vice-president of the armchair division.

However, the two uses above are actually of two distinct types. The first use is predicative: the property of being a sailor is predicated of John. The second use is specifying or equational: what is asserted is the identity of the referents of *Smithers* and *the vice-president of the armchair division*. In specifying sentences, NP_1 and NP_2 can be switched and the resulting sentence is still meaningful and even has the same truth conditions. The same is not true of predicational sentences as is shown by the following pair of examples:

*A sailor is John.

The vice-president of the armchair division **is** Smithers.

The importance of the above facts to our annotation is twofold. Certain event nouns occur in equational sentences where what's on the other side of *be* expresses a frame element of the frame evoked by the target noun.

His **decision** is to leave the company.

To leave the company is his **decision**.

In these equational sentences, we want to be able to tag the non-target NP as a frame element but we also want to be able to distinguish the two possible orderings of target NP and frame element. To accomplish this we made the decision to call *be* a copula only in cases where it follows the target. When *be* precedes the target we call it a support verb in analogy to regular support verbs like *take*, which in active sentences normally precede their target. This is illustrated by the following examples:

His **decision** [is COPULA] to leave the company.

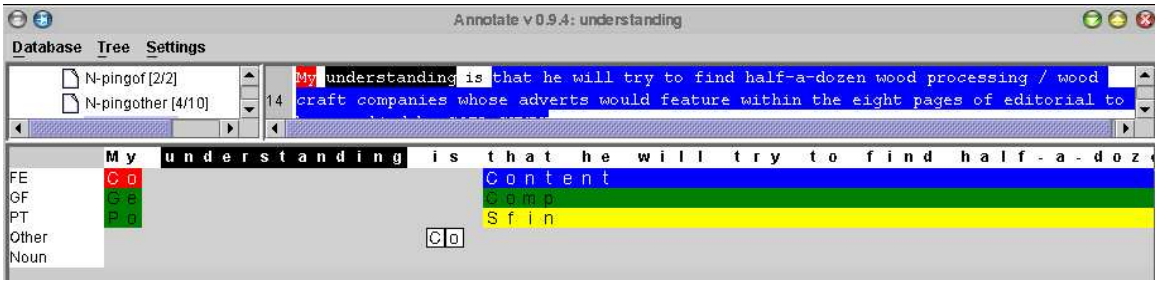
To leave the company [is SUPP] his **decision**.

[2] There are some cases of structural ambiguity where *be* is a copular verb under one reading, and an auxiliary under the other, as in: *His pastime is annoying the girls.*

We also use copula in some cases where the target names an artefact, rather than an event, and where a descriptive predicative adjective fills a frame element role.

Sue's **feet** [are COPULA] narrow.

Finally, note that the COPULA tag, unlike the SUPP tag, is applied to the Other-Layer.



2.3.3 Existential sentences

A subset of nouns can be used to report the occurrence of an event in a *there*-construction. We record these cases by applying the label **Existential** to the copular verb on the Other-layer. Usually the verb is *be*, but others, such as *come*, *occur*, and *ensue* are found as well.

There [was EXISTENTIAL] an **argument** between the players.

There [ensued EXISTENTIAL] a **discussion** about the current program.

In existential sentences like these, the word *there* is marked with the Null tag on the Other-Layer (see 2.2.4 for more information on the Null tag).

2.3.4 Null instantiation

In the case of noun targets, null instantiation is very common and it is much more difficult to decide what licenses the absence of the noun's conceptually necessary arguments. Quantification and generic use, as in the following pair of examples, make the notion of DNI inapplicable since by definition they do not allow for the individuation of events or states and their participants.

Every time Max did something like that to me, I knew how to **get even**.

Revenge is sweet.

However, even when there is no quantification, there are no reliable clues from definiteness marking whether frame elements are contextually known or not. Consider the following sentence uttered in a courtroom context.

His role in that **conviction** is now under investigation.

It is clear that in the Verdict frame to which the noun *conviction* belongs, the CHARGES are a core frame element. This can be illustrated with an example containing the verb *convict*.

Smithers was **convicted** after two hours of deliberation.

The preceding example is felicitous only if the CHARGES is contextually recoverable.

Returning now to our example about the noun *conviction*, we note that that sentence could be preceded by either one of the following discourses:

We have reason to believe that this agent has tampered with evidence before to get a conviction. Just two months ago the murder **conviction** of Howie Cheatham was overturned. Agent Smith also was a witness there.

We know that this agent has tampered with evidence before to get a **conviction**. There is for instance the case of a man in Missouri that had to be released after it was found that agent Smith had manipulated fingerprints.

While in the first example, the CHARGES (*murder*) are explicitly mentioned, in the second they are not identified at all, just existentially bound to the event of the earlier trial and conviction. Thus, definite reference to an event does not require that all aspects of an event be known to speaker and hearer. We are now left to decide whether the absent frame elements should be assigned to the INI or the CNI class. Since the omission does not seem lexically specific we should probably call it a case of CNI. However, we chose not to record such cases at all because, as in the case of INI with verb targets, they would be very frequent and laborious to annotate, and we are not sure how to characterize the licensing construction.

There are a few cases, such as the following example, where we record DNI for frame elements of nouns. In this example, a SOURCE location seems to have to be understood in the context.

[His THEME] **departure** had been delayed by two hours [DNI SOURCE].

But even in this case, one could construct an admittedly less plausible context that allows for mere existential binding of the SOURCE frame element:

Sam lost his wallet in an airport cafeteria last year. It took him an hour to retrieve it and he thought he had missed his flight. But luckily, ...

In sum, we currently do not have a good account of omitted arguments with target nouns. Use our annotation of null instantiated arguments of nouns with care.

2.4 Annotation with adjectives as targets

Adjectives can also evoke frames, that is, be frame-bearing. Certain semantic areas such as emotion or evaluation-related frames actually have a considerable share of adjectival lexical units. FrameNet annotates adjectives both when they are used attributively and when they are used predicatively. For the most part, adjectives take similar kinds of phrases as complements as verbs or nouns: finite clauses (*happy that*); infinitival clauses (*eager to*); gerundive forms (*busy writing*); prepositional phrases (*curious about*); adverbial phrases (*very/not in the least disappointed*); etc. However, predicative adjectives are different from verbs in that they do not take direct objects, with the possible exception of *worth* (as in *It's not worth the trouble*). And attributive adjectives, of course, modify nominals (N or N') rather than complete noun phrases.

2.4.1 Support & Copula

When adjectives are used predicatively, the annotation of verbs like *be* is the same as with predicate nominals: they get tagged as SUPP. Notice that in the case of adjectives, unlike with nouns, there is no possibility of a specifying use since adjectives are not referential. Thus, the copular verbs occurring with adjectives are tagged as SUPP in all cases.

> Smithers [is SUPP] very **clever** but he's no Einstein.
>
> **Clever** though he may [be SUPP], Smithers is no Einstein.

2.4.2 Relational modification

There is a distinct subclass of adjectives like *medical, military, judicial* that can never be used predicatively as is shown in the following examples.

> The White House announced a new **economic** policy.
>
> *The policy that the White House announced is **economic** .

We call these adjectives **relational** modifiers. Other names used for them include **pertainyms** or **domain** adjectives. Relational adjectives are comparable to the non-head noun in noun-noun compounds. They do not modify the referent or the extension of a head noun but rather its sense or intension, hence the name. Many of them have a meaning roughly like 'having to do, relating to, characteristic of' [some abstract or concrete entity]. These adjectives are not frame-bearing, at least not in a way that is concrete enough to allow for a clear definition of a scenario and of frame elements. Although they are placed in the frames that they are broadly associated with—the adjective *retributory*, for instance, lives in the Revenge frame—we do not provide full annotation for them.

2.5 Annotation relative to slot fillers

Some nouns—for example, kind and artefact nouns like *tomato, hammer*, or *pants*—do not evoke frames by themselves, or do so only marginally. These nouns mostly just occur as

slot fillers in frames evoked by verbs, adjectives, or other nouns. Nevertheless, we tag a governing verb or preposition on their POS-Layer as a governor for some of these slot filler nouns.

The motivation for doing this is as follows. We would like to know about certain entities in which frames they appear as slot fillers, and conversely what a typical slot filler for a particular frame element is. Imagine that we had a Cutting frame. We might wonder what kinds of things are usually sliced as opposed to simply cut. Conversely, thinking about tomatoes, we might be wondering what kinds of events they are participants in. We may intuit that they are often mentioned as objects of *slice* but are there other kinds of cutting, for instance, that apply to tomatoes?

We could attempt to automatically derive this kind of information from our annotation relative to governors like *slice, cut, chop*, etc. However, for that to be useful we would have to annotate many more sentences for each of the governors than we currently do.[3] For lexicographic purposes, it is not necessary to document that in addition to onions, carrots, and many other things, tomatoes can be sliced as well and so we have not done it. Thus, currently the easiest way to address questions like the ones above is to look for sentences with nouns such as *tomato*, *onion* etc. and treat them as targets. We can then record what syntactic governors (Gov) take phrases containing the artefact noun targets as arguments and what kind of syntactic constituents contain the target. (For lack of a better term, we call the constituent containing the target slot filler noun 'X'.)

He [grabbed GOV] [two **tomatoes** and the biggest apples X].

Skin and [quarter GOV] [the **tomatoes** X].

Notice that we tag both the verb and the particle as GOV when a phrasal verb is the Governor of a target noun, whether verb and particle are contiguous or not.

Leslie [chopped GOV] [up GOV] [the **tomatoes** X].

Leslie [chopped GOV [the targettomatoes X] [up GOV].

In several of our frames of artefact nouns, we also have defined some frame elements reflecting a kind of qualia structure of the artefacts. For instance, we may record modifiers that denote the material from which an artefact is made (constitutive quale) or the purpose for which it is used (telic quale). Here are some examples from the Clothing frame, which contains nouns denoting items of clothing.

Dot always [slept GOV] [in [her WEARER] [**vest** GARMENT] and knickers X]

I [put on GOV] [[my WEARER] [smart DESCRIPTOR] **suit** X] for the last of the interviews

In addition, to Gov-X annotation and qualia-type annotation, we can also attach a semantic type to some or all the lexical units in a frame. For instance, we could assign the semantic type 'Dress' to all items of clothing. Similarly we could give nouns like *wood,*

[3] We aim for around 3 -5 examples per subcorpus, resulting in an average of around 20-25 sentences per lexical unit.

concrete, ivory in a Materials-frame a semantic type 'Material'. This kind of information would then give users of our database information about classes of words that meet certain selectional restrictions.

One special kind of noun that received Gov-X annotation are what we call **transparent** nouns. Nouns like *top, pound, bunch* can appear as the first noun in N_1-of-N_2 constructions in contexts where the governing verb actually selects N_2 semantically rather than N_1, the syntactic head. In these contexts, we call N_1 transparent.

Sue drank a **cup** of hot coffee.

He pinned a **square** of fabric on the back.

Semantically, the nouns that can be transparent fall into the following classes:

- Aggregates (*bunch, group, collection*)
- Quantities (*flood, number, scores*)
- Types (*breed, class, ilk, kind*)
- Portions and Parts (*half, segment, top, bottom*)
- Unitizers (*glass, bottle, box*)
- Evaluations (*gem, idiot*)

While these words were annotated in appropriate frames such as Aggregate in the case of *group, bunch*, annotators added Gov-X annotation for those sentences where the target noun appears as a transparent N_1:

Leslie [ate GOV] [a [**sliver** PIECE] [of lemon SUBSTANCE]X].

Robin [drank GOV] [[a COUNT] [**pint** UNIT] [of beer STUFF] X].

2.6 Text annotation

As we said in the introduction to this chapter, we annotate individual sentences that we selected for containing a target word of interest to us. Annotation of running text is not a part of the main project but it is technically possible thanks to the annotation layering technique: one could one by one declare each word in a sentence a target, select a frame relative to which the new target ought to be annotated, get a new set of annotation layers (frame element, grammatical function, phrase type) and appropriate frame elements tags, and begin to annotate.

Chapter 3

Identifying Phrase Types

The syntactic metalanguage used in the FrameNet project is intended as a framework for **lexical description**–i.e. to describe the syntactic valence properties of individual lexical items. It is not intended as a framework for the complete syntactic description of sentences. In choosing the phrase types and grammatical functions to use, the major criterion was whether or not a particular label might figure into a description of the grammatical requirements of one of the **target words** of the project. The emphasis on what is relevant to lexical descriptions means that we limit ourselves, for the most part, to those phrase type labels which might appear in subcategorization frames. We do not include a complete list of all phrase types as would appear in more theoretically oriented syntactic descriptions.

FrameNet syntax also differs from traditional treatments of subcategorization, in its inclusion of certain modifiers. For example, the FrameNet description of a noun like *clinic* makes reference to the types of nouns which can modify this noun in compounds like *allergy clinic*. In theoretical treatments, modifiers of this sort are typically assumed to be outside the realm of subcategorization. We include them because they frequently express the same semantic roles (or frame elements) with respect to the modified heads as complements express with respect to their governors. For example, the frame associated with the verb *treat* includes a role for the Disease which is treated, and this role is typically expressed as the object of *treat*. Thus, you can *treat an allergy*, *treat the flu*, and so forth. Modifiers occurring with the noun *treatment* often express the same role; thus, there are *allergy treatments*, *flu treatments*, etc. We consider such observations to be of lexicographic relevance, and our syntactic descriptions reflect this perspective.

Target words are not assigned a PT tag.

3.1 List of phrase types

What follows is a list of phrase types used in FrameNet. Phrase types are assigned automatically during the classifying process, but may require manual correction.

3.1.1 Noun Phrase Types

Non-referential Noun Phrase

There (Expletive *there*)
It (Expletive *it*)

Possessive Noun Phrase (Poss)

Non-maximal Nominal (N)

Standard Noun Phrase (NP)

3.1.2 Prepositional Phrase Types

(**Prepositional Phrases (PP)**
Two types of Prepositional Phrases are assigned the phrase type **PP**.

Standard Prepositional Phrase (with NP object)

Particle (with no object)

PPing (Preposition with gerund object)

3.1.3 Verb Phrase types

Finite Verb Phrase (VPfin)

Nonfinite Verb Phrase

VPbrst (Bare Stem Verb Phrase)
VPto (*To*-Marked Infinitive Verb Phrase)
VPing (Gerundive Verb Phrase)

3.1.4 Complement Clause Types

Finite Clause

Sfin (Finite Clause (with or without *that*))
Swh- (*Wh*-Clause)
Swhether (*Whether/if*-Clause)

Nonfinite Clause

Sing (Gerundive Clause)
Sto (*To*-marked Clause)
Sforto (*For-to*-marked Clause)
Sbrst (Bare Stem Clause)

3.1.5 Subordinate Clause (Ssub)

3.1.6 Adjective Phrase Types

Non-maximal Adjective (A)

Standard Adjective Phrase (AJP)

3.1.7 Adverb Phrase (AVP)

3.1.8 Number Phrases (NUM)

3.1.9 Quote (QUO)

3.2 Phrase Type Labels for Noun Phrases

3.2.1 Non-referential NPs

The first distinction to worry about with NP types is that between **referential** and **non-referential** NPs. Expletive *it* and *there* are the two kinds of non-referential NPs. These are **not** given PT tags, although we record their presence on the Other Layer.

Some examples are given below.

[It] is clear that we won't finish on time.

[It] is odd that George is winning.

[There] are more cookies in the jar.

[There] is a fly in my soup.

3.2.2 Possessive Noun Phrase (Poss)

Referential NPs are either possessive NPs (marked **Poss**) or standard (non-possessive) NPs (marked **NP**). Possessive NPs, which may either be possessive pronouns or noun phrases marked with *'s*, often express frame elements of predicating nouns. For example, in the Communication frame, possessive nouns express The Speaker role when they are the determiners of target nouns such as *claim*, *remark*, *reply*, etc.:

I question [your] **claim** that the car was already damaged.

[The President's] **remarks** surprised the reporters.

[Leslie's] **reply** was well-timed.

Note: The label 'possessive' is not restricted to NPs denoting actual possessors. It is a morphosyntactic type rather than a semantic type.

3.2.3 Non-maximal Nominal (N)

In some situations it is necessary to tag nominal expressions which are not complete (i.e. maximal) noun phrases. For example, consider nominal modifiers of target nouns, or the modified nouns in sentences showing target adjectives used attributively in what follows here.

The judge dismissed the [forgery] **allegations**.

[Cancer] **treatments** are advancing rapidly.

Allergic [patients] benefit from this medicine.

The senator gave a **polemical** [speech].

These non-maximal nominal expressions are given the phrase type **N** (for 'nominal').

In contrast, head nouns that are frame elements of post-nominal modifiers are treated as if the post-nominal modifier was used with a copula, i.e. they are treated as full NPs with respect to Phrase Type, and as External arguments with respect to Grammatical function. They are not treated as non-maximal nominals.

The problem seems to affect [people NP/External] **sensitive** to primulas.

3.2.4 Standard Noun Phrase (NP)

With the exception of possessive[3.2.1]poss) and of referential noun phrases[3.2]nonref), all noun phrases are marked with the phrase type **NP**. The sections below discuss special circumstances which arise with the tagging of NPs.

Nouns with complements

Some nouns take prepositional or clausal complements. These are included in the PT tag for the relevant NP. In the examples below, noun complements appear in italics.

I **heard** [a story *about a man named Jed*].

I **dropped** [the lid *of my vitamin jar*]

[The fact *that moles are blind*] is totally **irrelevant**.

Nominals with relative clauses

Relative clauses containing the target word If the target word is inside the relative clause, we include the relative pronoun or relativizer with the head nominal, as in the following examples:

[the doctor who] **cured** my insomnia

[the joke that] got **repeated** over and over

Relative clauses with the target word outside If the target word is not inside the relative clause, we tag the whole relative clause modifier along with the nominal head, as in the following examples. (Relative clauses are in italics.)

[The acupuncturist *(whom) I saw last month*] **cured** my insomnia.

Other post-nominal modifiers

We also tag other post-nominal modifiers inside NPs. These include reduced relative clauses headed by prepositions and participial forms of verbs. (Note, however, that, in general, if possible we choose sentences for annotation without reduced relative clause.)

[The cat *in the corner*] likes celery.

I have [a cat *with orange stripes*].

[The cat *running down the hall*] is my favorite.

I'm talking about [the cat *bitten by a mouse*].

If there is more than one post-nominal modifier, they are all tagged with the phrase type NP, as illustrated below.

Stop [that cat *with orange stripes running down the hall*].

3.2.5 Coordinate Noun Phrase (Coord)

While a Noun Phrase can consist of coordinated NPs (e.g. *John and Sue*, it cannot consist of an NP and a following PP headed by *with along with*, or *together with.*

Compare the following two sentences, where only the first one is a coordinate noun phrase and tagged **Coord**.

The police **arrested** [Jim and me COORD].

The police **arrested** [me SUSPECT] along with Jim.

3.3 Phrase Type Labels for Prepositional Phrases

PP is assigned to ordinary prepositional phrases with nominal objects and to particles, the latter under the assumption that particles can be regarded as prepositional phrases which lack objects. **PPing** is assigned to prepositional phrases with gerundial objects rather than nominal ones. Here are some examples:

The passengers **looked** [at the monitors]. PP

The players began to **spread** [out]. PP

The fog **prevented** us [from seeing anything]. PPing

In addition we assign the phrase type PP to the second piece of certain discontinuous degree phrases, as exemplified below.

Jo is so **wicked** [as to be beyond redemption PP].

Stealing paper clips from work is less **immoral** [than to take them from a store PP].

These are closely related to the *than-* and *as*-phrases in the following sentences, which we tag also tag as PP.

Jo is as **wicked** [as John PP].

Jo is less **moral** [than Dave PP].

3.3.1 Particles

Particles like those in the following examples are treated as prepositions without objects and are assigned the Phrase Type **PP**.

Did you **figure** the problem [out]?

Look the number [up] in the phone book.

He **took** his hat [off] and put it on the table.

Note that particles of this kind may occur before NPs and therefore give the appearance of being the heads of regular PPs with NP objects.

Did you **figure** [out] the problem?

Look [up] the number in the phone book.

He **took** [off] his hat and put it on the table.

However, given the fact they are separable, as shown in the earlier examples, they cannot plausibly be treated as the heads of PPs in these contexts . Therefore, they are assigned a separate label **PP**. Generally, lexicographers are informed in advance of the particles that can appear with particular target words.

Test: If you are uncertain about whether or not to treat a word W as a particle, perform this test:

1. Think of a simple VP of the form V W NP.
2. Transpose the W and the NP: V NP W.
3. If the transposed version is an acceptable paraphrase of the original VP, then the word W is a particle.

While some particles, like the ones above, are equivalent in form to prepositions and may therefore misleadingly appear to head PPs in certain contexts, other particles do not resemble prepositions and are therefore less likely to be mis-analyzed that way:

Throw [away] those old things!

The librarian told me to **put** [back] the books.

There is no syntactic reason to distinguish these particles from the ones which resemble prepositions, and they are therefore given the same label (PP).

3.3.2 Prepositional Verbs

Some verb-preposition combinations are clearly conventional, as shown here.

The passengers *looked at* the information monitors.

Let me know if you *come across* that reference I asked you about.

We analyze the prepositions in such expressions as heading PPs.

The passengers **looked** [at the information monitors]

Let me know if you **come** [across that reference].

Though these verb-preposition combinations are units in the lexicon, we do not capture their unitary status in terms of *constituent structure*. That is, we do not analyze *look at* and *come across* as syntactic constituents.

In accordance with the Construction Grammar analysis of these expressions, their unitary status is captured in the valence representations of lexical entries. For example, there will be a lexical entry for *look at* which states that the verbal head *look* requires a PP headed by the preposition *at*.

3.3.3 Complex Prepositions

Some prepositions function as individual lexical units but consist, orthographically, of more than one word (complex prepositions shown in italics):

Put the birthday cake *next to* the other desserts.

We had tofu *instead of* veal.

Expressions of this kind are treated as single complex prepositions which head normal PPs. The PPs in the above sentences should be tagged in the following way:

Your birthday cake is [next to the other desserts].

We had tofu [instead of veal].

3.3.4 Preposition Stranding

A preposition and its complement may be separated from each other, with the preposition appearing in a canonical post-verbal position and its complement noun phrase appearing a in a pre-verbal position higher in the clause.

[John] we **laughed** [at].

[The man] you **screamed** [at] is my father.

Since allowing for preposition stranding is not lexically relevant information, annotators are discouraged from marking such sentences. If, however, sentences with preposition stranding have been annotated, then the two parts are assigned their normal phrase type values, NP and PP.

3.3.5 Preposition Phrases with Relative Clauses

If the target word is inside the relative clause and one of its frame elements is a prepositional phrase containing the relative pronoun, then we include the phrase containing the relative pronoun or relativizer inside the tagged constituents with the head nominal, as in the following examples:

[the house out of which] I was **evicted**

[the operator to whom] he had **spoken**

Notice that the bracketed constituent is treated as a PP.

If preposition stranding occurs within the relative clause, we mark the antecedent and relativizer as well as the stranded preposition.

[the house that] I was **evicted** [out of]

[the operator that] he had **spoken** [to]

3.4 Phrase Type Labels for Verb Phrases

Every verb phrase has at least a head verb, which may be a main verb or an auxiliary. VPs headed by main verbs may also contain one or more auxiliaries. A verb phrase may also have a negative marker, an infinitive marker, a pre-verbal adverb phrase, one or more complements of the verb, and one or more post-verbal adjuncts. A VP may be headed by the main verb in a sentence or it may be embedded as a complement under another verb. The following examples show a variety of VPs.

I *have*. (In response to "Have you taken out the trash?")

This book *really stinks*.

I didn't expect you *to eat your sandwich so quickly*.

3.4.1 Finite Verb Phrases (VPfin)

Any VP containing a verb (including auxiliaries) which (1) expresses information about tense and (2) is not in a separate embedded clause is tagged as a finite VP. Finite VPs are not generally subcategorized for, but it is nonetheless necessary to tag them in certain contexts, as illustrated here.

Who do you **think** [ate the sandwich]?

What did you **say** [fell on your hat]?

This pattern seems to be limited to a fairly small number of verbs of belief and assertion which subcategorize for clausal complements: *think*, *believe*, *say*, *claim*, *assert*, etc.

3.4.2 Non-finite Verb Phrases

Among non-finite VPs it is necessary to recognize bare stem infinitives (**VPbrst**), *to*-marked infinitives (**VPto**), and gerunds (**VPing**).

Bare stem infinitives (VPbrst)

Bare stem infinitives are non-tensed verb phrases headed by verbs in the bare stem form without the infinitive marker *to*. Examples of bare stem infinitives (VPbrst) are given below.

We **made** the children [take naps].

Management **let** the employees [set their own hours].

Note that *the children take naps* and *the employees set their own hours* are not treated as clauses in the FrameNet project, though that is how they are sometimes analyzed.

***To*-marked Infinitives (VPto)**

To-marked infinitives are VPs that begin with the infinitive marker *to*. Otherwise they are identical to bare-stem infinitives. Examples of *to*-marked infinitives appear below.

The cat **wants** [to go outside].

The mayors **persuaded** the President [to support the cities].

It is **hard** for infants [to tie their own shoes].

Gerundive Verb Phrases (VPing)

Gerundive VPs are VPs headed by verbs in the *-ing* form. They often occur in syntactic contexts in which nominal expressions also occur. Examples of Gerundive VPs are provided here.

My friend **likes** [running barefoot].

[Inhaling pepper] **makes** most people sneeze.

We **watched** the dogs [playing].

Gerunds present a challenge because they are sometimes verb-like and sometimes noun-like. Annotators are advised to consider both syntactic and semantic criteria to determine if the automatic classification of gerundive verb phrases is correct.

3.5 Phrase Type Labels for Clauses

Expression types that are treated as clauses in some syntactic theories are treated in the FrameNet syntax as combinations of smaller constituents. For example, the sequence *Pat leave* in a sentence like *They made Pat leave* is sometimes analyzed as a 'small clause,' but in the FrameNet metalanguage it is treated simply as an NP followed by a bare stem infinitive VP. This strategy has been adopted for two reasons. First, it simplifies the lexicographers' task of annotation, making it unnecessary to decide in certain cases which combinations of constituents should be treated as clausal and which should not. Second, it makes the lexical descriptions produced by the FrameNet project relatively theory-neutral. While the

question of which verbal complements are clausal and which are not is answered differently in different syntactic theories, the analysis of clauses into their major constituents is in most cases uncontroversial.

3.5.1 Finite complement clauses

Declarative Finite Complement clauses (Sfin)

Declarative finite complement clauses are full sentences that may begin with the complement marker *that*. In this PT, the entire clause, including the complement marker, is tagged.

Pat **knew** [Kim would never agree]

Pat **knew** [that Kim would never agree]

Wh-interrogative Clauses (Swh)

Structurally, a *wh*-interrogative clause may be a sentence or a verb phrase. Although not full clauses, these phrases only occur in constructions which allow a full Swh and therefore a single PT is used for both. Note that we treat *how* as a *wh*-expression. *Wh*-expressions are included in the tag for the clause.

I **heard** [what you said].

I **forgot** [what to say].

I **know** [how you feel]. I don't **know** [how to react].

I **asked** [who came]. She **told** me [who to invite].

Whether-if Interrogative Clauses (Swhether)

Structurally, a *Whether-if* interrogative clause may be a sentence or, in the case of *whether*, a verb phrase. Although not full clauses, these phrases only occur in constructions which allow a full *Whether-if* clause and therefore a single PT is used for both.

I **wonder** [whether the Indian restaurant delivers]

He **wondered** [whether to turn back]

Kim didn't **know** [if Pat liked the show]

3.5.2 Non-finite Clauses

Gerundive Clauses (Sing)

Sequences of object-form noun phrase and gerundive verb phrase are treated as single clauses by FrameNet. The reason for the analysis as a clause is that the noun phrase cannot be separated from the gerundive verb phrase, for instance, in passivization.

I don't like [him being here all the time]

[*He] wasn't liked [being there all the time]

Notice that similar-looking gerundive forms with a possessive subject are treated as noun phrases:

I don't **like** his being here all the time.

To-marked clauses (Sto)

I'd **like** [you to meet my mother] Certainly , but I should **hate** [you to forget that he has scored more runs in Test cricket than any other Englishman].

In sentences like the above example, *you* cannot be the subject of a passive and therefore is treated as part of the non-finite clause.

*[You] would be liked [to meet my mother]

For-to-marked clauses (Sforto)

I'd **like** [for you to meet my mother] I would **prefer** [for John to stay in the 250 class].

Bare stem clauses (Sbrst)

The manager **demanded** [that employees be on time].

3.5.3 Subordinate Clauses

Certain clauses introduced by subordinators can be frame elements and consequently need to be tagged. Such clauses receive the PT value Ssub (Subordinate Clause) rather than Sfin (finite complement clause). In the following sentence, the *because*-clause expresses the Reason frame element of the target word *admire*, which belongs to the Cognition_Judgment frame.

I **admire** her [because she is an actress who can also sing]

Certain adjective targets welcome discontinuous Degree phrases, as in the following example, where we assign the phrase type **Ssub** to the *than*-phrase.

That wine is more **delectable** [than I could imagine SSUB].

3.6 Phrase Type Labels for Adjective Phrases

Adjective Phrases typically occur as pre-nominal modifiers, as non-Subject complements of copular *be* and a small number of related verbs (*seem*, *become*, etc.), and as predicate complements of verbs like *find*, *consider*, etc.:

They were eating [very large] sandwiches.

The house is [empty].

You seem [sad] today.

The company considers these documents [extremely valuable].

3.6.1 Standard Adjective Phrase (AJP)

An Adjective Phrase may consist of just a single adjective, an adjective with some modifying expression (such as an adverb or an intensifier), or a conjunction of adjective phrases:

We **found** the play [dull].

We **found** the play [extremely dull].

We **found** the play [extremely dull and too long].

3.6.2 Non-maximal Adjectival (A)

Some adjectival expressions to be tagged are not treated as complete (i.e. maximal) adjective phrases. This is typically the case with relational modification, as shown below.

[marital] **bliss**

*very marital bliss

These expressions are given the phrase type A.

3.6.3 Adjectives with Complements

Some adjectives take complements other than the nouns they modify and these should be included as part of the Adjective Phrase. For example, consider the comparative adjective illustrated below:

Leslie is [taller than Kim].

An adjective and its complement may form a discontinuous constituent:

We need to find a [taller] player [than Kim].

In such cases, both the adjective and its complement are enclosed in brackets and assigned the label AJP.

3.7 Phrase Type Labels for Adverb Phrases (AVP)

Sometimes an adverb expresses a frame element of a target verb, as illustrated here.

The President **answered** the question [affirmatively].

In this sentence, the adverb *affirmatively* expresses the frame element Message, because it tells us that the President said *Yes*, or something equivalent in meaning, to the reporter's question. It is therefore tagged and assigned the phrase type AVP, the grammatical function Mod(ifier) and the frame element Message.

3.8 Phrase Type Labels for Numerals (NUM)

When annotating target nouns, the preceding number or quantifier is given the phrase type NUM. For example,

Bob poured [two NUM] **cups** of coffee.

Bob poured [thirty seven NUM] **cups** of coffee at the brunch.

Bob drank [a NUM] **glass** of wine.

Note that we treat cardinal numbers and *a* (= 1) in the same way.

3.9 Phrase Type Labels for Quotes (QUO)

Some verbs of communication take quoted material as a complement and are assigned the PT QUO. For example:

["Get out of here!"] she **cried**.

["But, I, er, uh..."] he **stammered**.

Quoted material can be of any syntactic form, or syntactically ill-formed, for that matter. Because the distribution or 'external syntax' of quoted material does not depend on its internal syntactic structure, we use a separate phrase type to tag it. Only direct quotes are given the phrase type QUO. Indirect quotes always take the form of some other kind of specific phrase type, e.g.

They **asked** us [what we were doing there].(*Wh*-clause)

The President **said** [that he would support the inner city].(*That*-clause)

Quoted material is easy to identify because it almost always appears in quotation marks, which should be included inside the brackets marking the Quote constituent.

Sometimes quoted material forms a discontinuous constituent:

["Cities,"] he **said**, ["are a very high priority."]

In such cases, both portions of the quote should be enclosed in square brackets and assigned the PT QUO. The tagging and annotation software will automatically coindex the parts and treat them together as a single unit.

Chapter 4

Assigning Grammatical Functions

In FrameNet, we annotate example sentences from the point of view of one particular **target word** in a given sentence. Each constituent tagged with a frame element in respect to a **target word** is also assigned a grammatical function tag in respect to that target. Only constituents tagged with frame elements are assigned grammatical functions. While target words are occasionally tagged with frame elements, they are never assigned a grammatical function.

The grammatical functions (**GFs**) that are assigned describe the ways in which the constituents satisfy abstract grammatical requirements of the target word. FrameNet grammatical function labels do not describe surface-syntactic positions of the constituents to which we assign them.

For example, suppose the following sentence is selected to exemplify grammatical properties of the target word *treat*:

Circumstances forced the doctor to **treat** her enemies.

Although the word *circumstances* is the subject of the sentence as a whole, this fact is not of interest to us and is not marked in any way in the example sentence. Instead, the NP *the doctor* is tagged as the external argument (Ext) of *treat*, even though it is not the surface subject of the sentence, because it satisfies a valence requirement of the verb *treat* outside the phrase headed by *treat* (thus 'external'). That is, it satisfies a semantic role, associated with *treat*, which would be realized in a simple declarative main clause by the subject of the clause.

The doctor also **treated** her enemies.

The combinations of grammatical function labels that occur with particular lexical items provide both a way of encoding the syntactic constructions a lexical item occurs in, and also a way of retrieving certain narrower distinctions between grammatical functions than those covered by the set of GF labels used.

For example, the verb *like* can occur in sentences with 'extraposed objects':

I **like** it [that you speak French].

FrameNet annotation of example sentences demonstrate that as one of its valence options, the verb *like* may take a null object *it* followed by a complement clause.

Eventually, it will be possible to retrieve examples of extraposed objects from the database by searching for combinations of null object and complement clause.

4.1 List of Grammatical Functions

What follows is a list of all of the grammatical functions used in FrameNet. It is followed by sections providing detailed criteria for the assignment of each GF. As with phrase types, GFs are assigned automatically during the classifying process, but may require manual correction.

4.1.1 Grammatical Functions assigned by Target Verbs

External Argument (Ext)

Object (Obj)

Complement (Comp)

Modifier (Mod)

4.1.2 Grammatical Functions assigned by Target Adjectives

External Argument (Ext)

Head noun modified by attributive adjective (Head)

Complement (Comp)

Modifier (Mod)

4.1.3 Grammatical Functions assigned by Target Prepositions

External Argument (Ext)

Object (Obj)

4.1.4 Grammatical Functions assigned by Target Nouns

External Argument (Ext)

Complement (Comp)

Genitive determiner (Gen)

Modifier (Mod)

Appositive (Appos)

4.2 Assigning GFs for Verbs

4.2.1 External Argument (Ext)

External covers situations in which phrases outside of the maximal phrase headed by the target word are functionally linked to the target word. This includes anything that satisfies an FE requirement of a target word in any of the following syntactic contexts:

- as a subject of a finite target verb, as in

  [The physician] **performed** the surgery;

- as a subject or object of a controlling governor of the target, as in

  [The physician] **decided** to **perform** the surgery;

- as the possessive modifier of a governing noun, as in

  [The physician's] **decision** to **perform** the surgery....

Notice that the phrase *the physician's* is in the Genitive GF relation to the noun *decision*, but bears the Ext relation to the verb *perform*.

In addition, any constituent which controls the subject of a target verb is assigned the GF External. This constituent might be a subject, an object, or a prepositional object in its local syntactic context, as shown below.

[The doctor] tried to **cure** me.

They persuaded [the doctor] to **treat** me.

They gestured to [us] to **leave**.

In the last example, note that it is just the NP *us*, and not the PP *to us*, which is tagged as the External Argument. This is because here we are interested in the valence properties of *leave* and not in the valence properties of *gesture*. Only the latter are responsible for the prepositional marking of the prepositional phrase *to us*.

A sub-case of a controlling verb is that of a verb that can be seen as specifically dedicated to providing external representation for an element of the conceptual structure associated with the meaning of a nominal or adjectival target, which we refer to as **support verbs**. Subjects and objects of support verbs can be taken as Ext for the dependent word; other complements of the support verb, where relevant, can be treated as complements of the dependent noun or adjective.

This last provision allows us to finesse the problem of disputed constituency decisions in the case of support-verb constructions. Thus, in the following sentence it is contestable whether the phrase *to the press* is a complement of the noun *statement* or the verb *make*.

He wrote a **statement** to the press about the bribery case.

Regardless of the analysis, FrameNet tags the phrase with the FE Addressee as a complement of the target word *statement*. By allowing this phrase under either of the disputed analyses, we avoid the constituency decision completely.

It would have been theoretically justifiable to omit selecting phrases outside of the standard **subcategorization frame** of a target word, or we could have limited such excursions to the subjects of finite verbs. However, since one of our goals is to provide a database that includes samples of phrases capable of satisfying particular FE requirements of the words we analyzed, our decision increased the size of such a sample.

In the process of noting the function of such words, we have also taken on the obligation of recording the actual support verbs we encountered. Thus, the database is also a resource for identifying the support verbs that most often accompany particular nouns and adjectives. In doing this, we have taken a relaxed rather than a strict view of identifying support verbs, including alongside of *make* (as in *make an attack*), also such semantically richer verbs as *launch* (as in *launch an attack*), etc. An extension of FrameNet could be defined which sought to include the full range of **Lexical Functions** in the sense of I. Mel'cuk. [1]

4.2.2 Object (Obj)

Any normal object, any *wh*-extracted object, or any post-target-verb NP which controls the subject of a complement of the target verb is assigned the grammatical function Obj, as shown here.

Voters **approved** [the stadium measure].

[What] did you **cook** for dinner?

They **expect** [us] to finish soon.

They **made** [us] eat our vegetables.

In some syntactic theories, the NP *us* in the last two examples would be treated as the subjects of small clause complements, and not as Objects of the target words. We have decided to treat all post-verbal NPs of this type as Objects in order to simplify the task of tagging.

The GF Object is also assigned to any subject of a *tough*-predicate which satisfies the Object role of a verb or preposition in the complement of the *tough*-predicate.

[Artichokes] are hard to **eat**.

4.2.3 Complement (Comp)

Complement is the general grammatical function assigned to PPs, VPs, Clauses (and a small number of NPs) which occur after their governing verbs, adjectives or nouns in normal declarative sentences. This grammatical function includes both what are usually referred to as *complements* and what are usually referred to as *obliques*. One reason we do not

[1] I. Mel'cuk (1996) Lexical Functions: A Tool for the Description of Lexical Relations in the Lexicon. In L. Wanner (ed.): Lexical Functions in Lexicography and Natural Language Processing, Amsterdam/Philadelphia: Benjamins, 37-102.

recognize Oblique as a distinct grammatical function is that it seems merely to be reserved for PPs that are complements. In other words, it redundantly encodes phrase type. For those who wish to maintain the distinction between complement and oblique, keep in mind that any PP which is assigned the GF Complement can be considered an oblique.

PP Complements

The GF Complement is assigned to any particle or any PP, optional or obligatory, which expresses a semantic role belonging to the frame associated with the target word. This does not include setting adjuncts of PLACE or TIME, purpose clauses, or other such expressions which can occur with very large classes of predicators. Here are some examples of of the Complement GF.

Give the gun [to the officer].

Pat **spoke** [to me].

Pat **lives** [in Cleveland].

Note: A Locative expression may be a Complement if it expresses a role belonging to the frame of the target predicator. In the third example above, the PP *in Cleveland* is a Complement because the frame of the verb *live* (='reside') includes a role for the place in which a person lives.

Some NPs are marked as Complements rather than as Objects. These NPs are not passivizable, and they often express Place, Time, and other meanings normally associated with adjuncts and PP complements (e.g. Measurement). Such NPs are often Complements in the same contexts in which comparable PPs might be used instead.

I **run** [ten miles] every day.

Come [this way]!

I **expect** your papers [the moment you walk into class].

In keeping with conventions of Construction Grammar, the second object of ditransitives is treated as a Comp. FrameNet does not have a category for second object and does not assign the grammatical function Object a second time in ditransitive sentences.

They **gave** the children [candy].

The children were **given** [candy].

A note on particles: Even if a preposition seems to 'go with' a verb to form a phrasal verb, if it can plausibly be considered the head of a post-verbal PP it is analyzed that way, and the PP is assigned the GF Comp. Separable verb particles, like *up* in *pick up the package* (and *pick the package up*), cannot plausibly be treated as the heads of PPs. They are therefore marked with the GF Comp by themselves, and the NPs with which they occur are treated as GF Obj. In passive sentences, these NPs are treated as GF Subj.

Pat **picked** [up Comp] [the package Obj].

Pat **picked** [the package Obj] [up Comp].

[The package Subj] was **picked** [up Comp] by Pat.

For more information on particles, see the appropriate section in the chapter on Phrase Types.

Locative adverbs such as *here* and *there*, when used as true complements, are also assigned the GF Comp.

Chuck **went** [there Comp].

Verbal and Clausal Complements

Any verbal or sentential complement, regardless of whether or not it is passivizable is assigned the GF Complement, as shown here.

They **want** [to stay home].

They **expect** us [to stay home].

I **believe** [that you are the winner].

They **think** [you are the president].

You **persuaded** me [to finish early].

I **wonder** [who will finish first].

4.2.4 Modifier (Mod)

Modificational adverbs such as those indicating manner or ending in *-ly*, as well as related modificational adverbs denoting speed, are assigned the GF Mod.

The board [quickly] **resolved** the problem.

4.3 Assigning GFs for Adjectives

4.3.1 External Arguments of adjectives

When an adjective appears in a clausal predication, one of its arguments is expressed as the subject of a support verb (indicated with underlining) and of the clause as a whole, as in the following examples.

[The chair] is **red**.

[My sister] seems more **interesting** than yours.

This constituent is assigned the GF External Argument (Ext). Also, the noun in object-control constructions with adjectives is assigned the GF External Argument, illustrated below.

We consider [Pat] very **intelligent**.

Note that copulas and support verbs are not assigned GFs.

4.3.2 Modified head nouns with pre-nominal adjectives

In pre-nominal uses of adjectives, the modified head nouns are assigned the GF Head (Head).

the **small** [children Head]

Only some types of sentences in which a target adjective appears in pre-nominal position will actually be annotated and included in the database of corpus examples. These are the sentences in which the adjective has a *qualitative* as opposed to a *relational* use.

In a qualitative use, the modified noun expresses an element of the frame associated with the adjective, and this is the same frame element which is typically expressed by the subject of the verb BE or other clausal predication (in which the adjective occurs in predicate position):

The children are **small**.

Note that in these cases, the annotators tag the verb with the label Supp on the POS-Layer.

In a relational use of an adjective, it is much more difficult to identify a specific frame element which is expressed by the modified noun. In fact, the semantic relation between the adjective and noun may be more strongly determined by the modified noun than by the adjective. In any event, the relation between modifier and head is much less tightly constrained than in cases of qualitative modification, resembling the relation between nouns in a noun compound. Very often relational uses of adjectives do not have corresponding predicative uses:

Pat had an immune response to the virus.

*Pat's response to the virus was immune.

We do not annotate relational uses of target adjectives because doing so is unlikely to reveal much that is interesting about the frames associated with the adjectives.

Many cases of relational modification (like many noun compounds) are highly conventionalized. Such cases will be treated as idioms when they are relevant to a particular domain, and will be identified with respect to their head nouns rather than their adjectives. For example, *immune response* will be treated in the health care domain as a lexical unit headed by *response*, rather than as a particular use of *immune*.

4.3.3 Post-nominal adjectives

In post-nominal uses of adjectives, their arguments are treated as if the adjective were used with a copula. In particular, modified head nouns are assigned the GF External (Ext) and the PT Noun Phrase (NP) rather than the GF Modifier (Mod) and the PT Noun (N). Thus, in the following example, *people* has the grammatical function Ext and the phrase type NP.

The problem seems to affect [people] **sensitive** to primulas.

4.3.4 Adverbial Modifiers of Adjectives

Adverbial modifiers of adjectives are assigned the GF Mod.

The [very MOD] **beautiful** woman left the office.

Certain degree adverbs may also modify adjectives, and are therefore assigned the GF MOD with adjective targets, as shown below.

That was [so Mod] **silly** of you.

[How Mod] **offensive** that movie was!

It was a [quite Mod] **remarkable** feat.

4.3.5 PP Complements of Adjective

Some adjectives welcome PP Complements, to which we assign the GF Comp.

Jo is not **aware** [of his bad behavior COMP].

Lee is **certain** [of his innocence COMP].

4.4 Assigning GFs for Prepositions

For most frames prepositions are not among the target words. However, they are targets in frames that cover the vocabulary of space, time, and motion. In these frames, we need to be concerned with assigning GFs for the relevant constituents.

4.4.1 Objects of prepositions

Any constituent which occurs immediately after a preposition and which expresses an element of the frame associated with the preposition is tagged as an Object (**Obj**). Typically this constituent is an NP, but it can also be a gerund or a clause:

We had a glass of wine **before** [the meal].

Wash your hands **before** [returning to work].

He left **before** [I had a chance to say anything].

4.4.2 External Arguments of prepositions

A constituent which expresses an element of the frame associated with a preposition, but which is outside the PP, is tagged as an External Argument (**Ext**). The most easily identified prepositional External Arguments are those which occur with so-called reduced relative clauses (i.e. post-nominal modifiers) and with copular predications:

[the day] **before** yesterday

[the trail] **to** our campsite

[the ball] **under** the table

[The ball] is **under** the table.

4.5 Assigning GFs for Nouns

4.5.1 Assigning GFs for Nouns (Comp)

The GF Comp is assigned to any post-nominal complement of a target noun, example sof which are shown here.

the **fact** [that cats have fur]

a **letter** [to the President]

a **story** [about a young entrepreneur]

your **attempt** [to scare me]

our **arrival** [at the station]

Sometimes complements of nouns are realized as the predicates of copular sentences, e.g.

The **fact** is [that cats have fur].

The **letter** was [to the President].

The **attempt** was [to scare you].

In order to assist the computer in this process it is necessary to tag the copula which follows the target and introduces the 'predicate complement' in these expressions. For this purpose annotators use the simple tag Copula (Cop) on the Other-Layer.

4.5.2 Genitive determiner of noun (Gen)

The GF Gen is assigned to any possessive NP functioning as determiner of a target noun, as indicated here.

[your GEN] **book**

[your work's GEN] **influence** on the field

Note that we use the term Possessive (Poss) to refer to the phrase type of Genitive NPs.

4.5.3 Modifier of noun (Mod)

The GF MOD is assigned to any pre-nominal modifier (whether a noun, adjective, gerund, or participle), as shown below.

[allergy] **treatment**

[monthly] **stipend**

[sleeping] **cat**

[broken] **lamp**

4.5.4 Quantification Modifiers (Quant)

The GF Quant is assigned to a pre-nominal modifier of a target noun, including the indefinite determiner when it functions as a number, illustrated below.

[two] **cups** of coffee

[a] **glass** of wine

4.5.5 Appositives

(Appos)

The GF Appositive is assigned to post-target appositional Ns or NPs.

Libel **lawyer** [Jonathan Crystal APPOSITIVE] represented the plaintiff.

Chapter 5

Lexical Entry Structure

5.1 Lexical Entries

FrameNet lexical entries are sets of lexical sub-entries, each of these being a record of what FrameNet has recorded for the lemma in one of its senses. (Since our work has proceeded one frame at a time, rather than one lemma at a time, there are not many instances of multiple sub-entries in the database.

An individual lexical entry, then, covers a lemma in a particular part of speech, e.g., as verb or as noun. A lexical sub-entry is intended to represent a single lexical unit, i.e., a lemma in a given part of speech in a single sense.

A lexical sub-entry comprises the following components:

1. Headword: the lexeme to be defined
2. Frame: a path to the individual background frame, e.g. Attaching, Judgment_communication, or Weather
3. A definition, taken from the Concise Oxford Dictionary, 10th Edition, or written by a FrameNet lexicographer
4. Table of Frame Element Realizations: a full list of the syntactic ways, in terms of grammatical function and phrase type, in which Frame Elements have been expressed in the annotated sentences
5. Table of Valence Patterns: a list of the groupings of Frame Elements and their syntactic realizations as found in the annotated sentences
6. Annotated sentences (where each sentence is annotated in respect to a single target word and the semantic roles which neighboring phrases bear to that word)

The Frame Element Realization table and the Valence Pattern table are derived automatically from the sentence annotations. Each item in each of these tables is linked to the annotations that exemplify it.

In short, a FrameNet entry provides information, for each sense, about frame membership and the syntactic means by which each Frame Element is realized in the word's

surroundings, and summarizes, as Valence Patterns, the full range of combinatorial possibilities as attested in the Corpus.

5.2 Incorporated Frame Elements

There are many verb frames which involve a particular kind of entity, in general, and some of these words incorporate information about a particular entity type in their meaning. Consider verbs of body movement, where a typical case is a verb which is expected to co-occur with the name of a body part, even when the identity of the body part is clear from the meaning of the verb. A dog *wags its tail*, people *arch their brows*, *bat their eyes*, *purse their lips*, etc. However, in the case of *smile*, *grimace*, *frown*, *pout*, and *scowl*, the affected body-part is not separately presented. We say that it is **incorporated**. Notice that with some verbs in this frame the affected body part can either be present or not: we can say either (*She blinked* or (*She blinked (her eyes)*. In the near future, for such cases, the information about incorporated FEs will be given in the lexical entry of the word.

We will provide complete lexical entries for each lexical unit in the FrameNet database in the final release of the data.

Chapter 6

Semantic Relations

There are several different kinds of semantic relations in the FrameNet database. These consist primarily of frame-to-frame relations which indicate semantic relationships between collections of concepts. In addition, **semantic types** are employed for a variety of purposes, marking features on frames, frame elements, and lexical units. The exact number and use of these relations is an area of active research.

6.1 Frame-to-frame Relations

The FrameNet database records information about **Subframe**, **Inheritance**, **SeeAlso**, and **Use** relationships between frames. Marking of these relationships is a work in progress, especially in the case of the SeeAlso and Use relations which were only recently added to our repertoire.

6.1.1 SubFrames

Some frames 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. The separate frames are related to the complex frames via the SubFrame relation. In such cases, frame elements of the complex frame may be identified (mapped) to the frame elements of the subparts, although not all frame elements of one need have any relation to the other. (In this respect, it contrasts with inheritance; see below.) Also, the ordering and other temporal relationships of the subframes can be specified using binary precedence relations. To illustrate, consider the complex Criminal_process frame, defined as follows: A Suspect is arrested by an AUTHORITY on certain CHARGES, then is arraigned as a DEFENDANT. If at any time the DEFENDANT pleads guilty, then the DEFENDANT is sentenced, otherwise the DEFENDANT first goes to trial. If the VERDICT after the trial is guilty, then the DEFENDANT is sentenced. In the end, the DEFENDANT is either released or is given a SENTENCE by a JUDGE at the sentencing. For each step in the process, there is a separate frame in the database, including Arrest, Arraignment, Trial, Sentencing, and so on. Each of these frames is related to the Criminal_process frame via the SubFrame relation in the frame editor. Moreover, subframes (of the same complex frame) are related to each other through their ordering. Notice that a given subframe may itself be a complex frame. For example,

the Trial frame is a subframe of the Criminal_process frame, and has its own rich structure, some of which can be decomposed into simpler frames that are related to each other. A Trial is made up of court appearances, and involves opening arguments, presentation of evidence and testimony, and closing arguments. The system of subframe links is also quite complex. At present, the SubFrame relation is not indicated across the entire database.

6.1.2 Frame Inheritance

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 equally or more specific correspondents in the child frame. For example, the Perception_active frame, evoked by concepts such as *listen* and *watch*, inherits from the more general Perception frame, and thus all the frame elements of Perception have correspondents in the Perception_active frame as follows:

Perception_active	Perception
Ground	Ground
Perceiver_agentive	Perceiver
Phenomenon	Phenomenon

This means that Perception_active is a more specific kind of Perception, and that the PERCEIVER_AGENT role is a more specific instance of the PERCEIVER role, in which additional restrictions involving the PERCEIVER's active participation are imposed. As an example of how subframes correspond in an inheritance relationship, the Cause-to-move frame has Motion as a subframe, and thus its descendant, Carrying, must also have Motion (or a more specific kind of Motion) as a subframe.

In some cases a single word may evoke multiple frames simultaneously. In such cases, we say that the lexical unit belongs to a frame which inherits from both. This is called **multiple inheritance** (formerly referred to as frame blending).

An example of such a case is Judgment_communication, which inherits from Communication and Judgment. The two judgment frames share the following canonical structure

```
A __ed B for C-ing
```

representing the JUDGE (A), the EVALUEE (B) and the REASON for the judgment (for C-ing). As part of the inheritance relationship between these two frames, the Judgment_communication frame elements JUDGE, EVALUEE, and REASON are identified or **mapped** to the corresponding frame elements in the Judgment frame. In contrast, Judgment_communication and Communication, share the roles of SPEAKER, ADDRESSEE, and TOPIC; in the Judgment_communication frame the SPEAKER role is identified with the JUDGE and for some lexical units, such as *scold*, *berate*, *flatter*, and *compliment*, the ADDRESSEE role is, identified with the EVALUEE.

Ultimately, it will be possible to view the inheritance relations in a web report, but this is not yet implemented.

6.1.3 SeeAlso

In cases where there are groups of frames which are similar and should be carefully differentiated, compared, and contrasted, each of the frames in question will have a SeeAlso relation with a representative member of the group. In the frame definition of the representative member, there will be a comparison which will contrast the frames to make clear the intended boundaries between them. For example, since the Scrutiny and Seeking frames are similar, there is a SeeAlso relation from Scrutiny to Seeking, and text in the Seeking frame that explains the difference.

This relationship does not denote any particular relationship between the Frame Elements or Subframes of the frames involved.

6.1.4 Use

Often a particular frame makes reference to the structure of a more abstract, schematic frame. Typically, this occurs in cases where the Specific frame provides a specific perspective of the Schematic frame. For these, we say the that the Specific frame has a Use relationship with the Schematic frame, and bindings between the Frame Elements and Subframes may be specified. For example, the Commercial_transaction frame specifies a complex schema involving an exchange of multiple Themes (the MONEY and GOODS) between the BUYER and SELLER. The BUYING frame has a Use relationship with the Commercial_transaction frame in which the MONEY, GOODS, BUYER, and SELLER are identified.

This type of relation can capture the FrameNet I concept of domains. In FrameNet I, frames were organized by domains - very general categories of human experience and knowledge - to cover various areas of the English vocabulary and to provide useful groupings of semantic frames. Domains also had a degree of theoretical significance: they were broad-level generalizations over the frame network that we are constructing.

Many frames will have Use relationships simultaneously with inheritance relationships. In such a case, the Use relation will specify something like the basic domain of relevance, such as Cognition, Communication, Society, Body, Crime and Justice, Life and Death, Education, etc. (The frames that serve as a basis for Use relationships have alternatively been called *background frames* or *big frames.*) The frame or frames from which a frame inherits would be more akin to image-schemas, including Giving and Receiving, Containing, Emitting, Motion, etc. This combination of relationships will facilitate incorporating the notion of perspective into the database.

6.2 Semantic Type

We employ the mechanism of Semantic Type in order to capture semantic facts about frames, FEs, or LUs that don't necessarily fit into our developing hierarchy of frames. One of the major uses of semantic types is recording important semantic differences between lexical units that cut across frames, such as "Positive evaluation" and "Negative evaluation".

For example, lexical units across a range of frames may incorporate positive versus negative evaluation. Consider (again) the Judgment verbs *praise* and *criticize*, the Experiencer_subject verbs *like* and *hate*, and the Frugality adjectives *generous* and *stingy*. We can indicate the

semantic type "Positive evaluation" on the first word of each pair, and "Negative evaluation" on the second, thus recording the information in the database.

Other uses of semantic types include:

- Basic typing of fillers of frame elements, e.g. "Sentient" for the Cognizer FE
- Useful, functional marking on frames, such as the type "Non-lexical" on frames which are present purely to participate in inheritance, subframe, or use relations with other frames
- Useful, functional marking on lexical units, especially the case where a lexical unit (e.g. *Wednesday* in the Calendric_unit frame) is not to be annotated because it is essentially identical to other lexical units which have been annotated.

Appendix A

Publications related to FrameNet

Frame Semantics Background

Fillmore, Charles J. (1968): The case for case. In Bach and Harms (Ed.): Universals in Linguistic Theory (pp. 1-88), Holt, Rinehart, and Winston, New York.

Fillmore, Charles J. (1976): Frame semantics and the nature of language. In Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech, Volume 280 (pp. 20-32).

Fillmore, Charles J. (1977): Scenes-and-frames semantics, Linguistic Structures Processing. In Zampolli, Antonio (Ed.): Fundamental Studies in Computer Science, No. 59, North Holland Publishing.

Fillmore, Charles J. (1977): The need for a frame semantics in linguistics. In Karlgren, Hans (Ed.): Statistical Methods in Linguistics

Fillmore, Charles J. (1982): Frame semantics. In Linguistics in the Morning Calm (pp. 111-137), Hanshin Publishing Co., Seoul, South Korea.

Fillmore, Charles J. (1985): Frames and the semantics of understanding. In Quaderni di Semantica, Vol 6, No. 2 (pp. 222-254).

Fillmore, Charles J. (1992): "Corpus linguistics" vs. "computer-aided armchair linguistics". In Directions in Corpus Linguistics: Proceedings from a 1991 Nobel Symposium on Corpus Linguistics (pp. 35-66), Mouton de Gruyter, Stockholm.

Fillmore, Charles J. and B. T. S. Atkins (1994): Starting where the dictionaries stop: The challenge for computational lexicography. In Atkins, B. T. S. and A. Zampolli, eds. Computational Approaches to the Lexicon. Clarendon Press.

Recent Papers by FrameNet Staff

Many of these papers are available for down-load as .pdf or word files at the FrameNet web-site: http://framenet/icsi.berkeley.edu/~framenet.

Baker, Collin F. and Josef Ruppenhofer (To appear): FrameNet s Frames vs. Levin s Verb Classes. In Proceedings of the 28th Annual Meeting of the Berkeley Linguistics Society held February 15-18, 2002.

Baker, Collin F., Charles J. Fillmore, and John B. Lowe (1998): The Berkeley FrameNet project. In Proceedings of the COLING-ACL, Montreal, Canada.

Boas, Hans C. (2002): Bilingual FrameNet Dictionaries for Machine Translation. In M. Gonzlez Rodrguez and C. Paz Surez Araujo (eds.), Proceedings of the Third International Conference on Language Resources and Evaluation. Las Palmas, Spain. Vol. IV: 1364-1371.

Boas, Hans C. (2001): "Frame Semantics as a framework for describing polysemy and syntactic structures of English and German motion verbs in contrastive computational lexicography," in: Rayson, Paul, Andrew Wilson, Tony McEnery, Andrew Hardie, and Shereen Khoja (eds.), Proceedings of the Corpus Linguistics 2001 conference. Technical Papers, Vol. 13. Lancaster, UK: University Centre for computer corpus research on language

Chang, Nancy, Srini Narayanan, and Miriam R.L. Petruck (To appear): From Frames to Inference. In Proceedings of the First International Workshop on Scalable Natural Language Understanding. Heidelberg, Germany.

Chang, Nancy, Srini Narayanan, and Miriam R.L. Petruck (2002): Putting Frames in Perspective. In Proceedings of 19th International Conference on Computational Linguistics. Taipei, Taiwan.

Dodge, Ellen, and Abby Wright (In press): Herds of Wildebeest, Flasks of Vodka, Heaps of Trouble: An Embodied Construction Grammar Approach to English Measure Phrases. In Proceedings of the 28th Annual Meeting of the Berkeley Linguistics Society.

Fillmore, Charles J.(In press): Double-Decker Definitions: The Role of Frames in Meaning Explanations, Sign Language Studies, Vol. 3. No. 3.

Fillmore, Charles J. and B.T.S. Atkins (1998): FrameNet and lexicographic relevance, Proceedings of the First International Conference on Language Resources and Evaluation, Granada, Spain.

Fillmore, Charles J., Collin F. Baker, and Hiroaki Sato (2002a): The FrameNet Database and Software Tools. In M. Gonzlez Rodrguez and C. Paz Surez Araujo (eds.), Proceedings of the Third International Conference on Language Resources and Evaluation. Las Palmas, Spain. Vol. IV: 1157-1160.

Fillmore, Charles J., Collin F. Baker, and Hiroaki Sato (2002b): Seeing Arguments through Transparent Structures, In M. Gonzlez Rodrguez and C. Paz Surez Araujo (eds.), Proceedings of the Third International Conference on Language Resources and Evaluation. Las Palmas, Spain. Vol. III: 787-791.

Fillmore, Charles J. and Collin F. Baker (2000): FrameNet: Frame semantics meets the corpus. Poster session at The Linguistic Society of America, January, 2000.

Fillmore, Charles J. and Collin F. Baker (2001): Frame Semantics for Text Understanding. In Proceedings of WordNet and Other Lexical Resources Workshop, NAACL, Pittsburgh, June, 2001.

Fillmore, Charles J.and Hiroaki Sato (To appear): Transparency and Building Lexical Dependency Graphs. In Proceedings of the 28th Annual Meeting of the Berkeley Linguistics Society held February 15-18, 2002.

Fillmore, Charles J., Charles Wooters and Collin Baker (2001): Building a Large Lexical Databank Which Provides Deep Semantics. In Proceedings of the Pacific Asian Conference on Language, Information and Computation. Hong Kong.

Gahl, Susanne (1998): Automatic extraction of Subcorpora based on Subcategorization Frames from a Part-of-Speech tagged Corpus. In COLING-ACL '98 Proceedings of the Conference held August 10-14, 1998, at the University of Montral, Canada, pp.428-32.

Gahl, Susanne (1998): Automatic extraction of Subcorpora for Corpus-based Dictionary-building. In Thierry Fontenelle et al. (eds.) EURALEX '98 Proceedings: Papers submitted to the Eighth EURALEX Conference, at the University of Lige, Belgium, pp.445-452.

Gildea, Daniel and Daniel Jurafsky (2001): Automatic Labeling of Semantic Roles, ICSI Technical Report 01-005.

Gildea, Daniel and Daniel Jurafsky (2000): Automatic Labeling of Semantic Roles. In Proceedings of ACL 2000, Hong Kong.

Johnson, Christopher R. and Charles J. Fillmore (2000): The FrameNet tagset for frame-semantic and syntactic coding of predicate-argument structure. In Proceedings of the 1st Meeting of the North American Chapter of the Association for Computational Linguistics (ANLP-NAACL 2000), April 29-May 4, 2000, Seattle WA, pp. 56-62.

Johnson, Christopher R. (1999): Multiple frame inheritance in lexical descriptions. Presentation at the 1999 Annual Meeting of the Linguistic Society of America. Los Angeles, January 9.

Lowe, J.B., Collin F. Baker, and Charles J. Fillmore (1997): A frame-semantic approach to semantic annotation. In Proceedings of the SIGLEX workshop "Tagging Text with Lexical Semantics: Why, What, and How?" held April4-5, in Washington, D.C., USA in conjunction with ANLP-97.

Narayanan, Srini, Charles J. Fillmore, Collin F. Baker, and Miriam R.L. Petruck (2002): FrameNet Meets the SemanticWeb: A DAML+OIL Frame Representation.

Petruck, Miriam R.L. (1996): Frame Semantics. In Jef Verschueren, Jan-Ola Östman, Jan Blommaert, and Chris Bulcaen (eds.). Handbook of Pragmatics 1996. Philadelphia: John Benjamins.

Ruppenhofer, Josef, Collin F. Baker, and Charles J. Fillmore (2002): Collocational Information in the FrameNet Database. In Proceedings of EURALEX 2002. Copenhagen, Denmark.

Index

Adjective Phrase, *see* AP
adjectives
    external argument of, 44
    non-maximal, 36
    post-nominal, 45
    with complements, 36
    with support verbs, 44
adjuncts, 43
Adverb Phrase, *see* AVP
adverbs, 36
AP, 35
    non-maximal, 27
    standard, 27, 36
    types, 27
artefact nouns, 16, 23
aspect, 16
AVP, 27

bare stem clauses, *see* clauses
Bare stem infinitives, *see* VP

clauses, 33
    bare stem, 35
    declarative finite complement, 34
    finite complement, 34
    for-to marked clauses, 35
    gerundive, 34
    how-clauses, 34
    non-finite, 34
    subordinate, 35
    that-clauses, 34
    to-marked clauses, 35
    wh-clauses, 34
    whether-if interrogatives, 34
CNI, 12
Complement clause
    types, 26
complement clauses, *see* clauses
complements, 42
    of nouns, 28
    versus obliques, 42
compounds, 25
conflation, 14
constituency
    prepositional verbs, 31
constructional null instantiation, 12
control, 16
copula, 19, 22, 47
copula tag, 20
core frame element, 10

definite null instantiation, 12
discontinuities
    adjectives and their complements, 36
    quotes, 37
ditransitives, 43
DNI, 12
domains, 53
double object sentences, 43

event nouns, 16
existential tag, 20
expletives, 13
external argument, 39, 41
    and support verbs, 41
    of adjective, 44
    of preposition, 46
    of verbs, 41

finite complement clauses, *see* clauses
Finite VP, *see* VP
for-to clauses, *see* clauses
frame bearing, 10
frame elements
    incorporated, 50
frame-to-frame relations, 51

frames
perception, 52
perception_active, 52

genitive determiner, 47
gerunds
as complements of prepositions, *see* PP
gerundive clauses, *see* clauses
gerundive verb phrases, *see* VP
grammatical functions, 39
with object control, 44
with tough predicates, 42

how-clauses, *see* clauses

incorporated frame elements, *see* frame elements
indefinite null instantiation, 13
infinitives, *see* VP
inheritance, 52
INI, 13
instructional imperatives, 12

lexical entry
structure of, 49
lexicographic relevance, 25

modifier
pre-nominal, 47, 48
modifiers, 25, 44
AP, 35
post-nominal, 28, 29
qualitative, 45
relational, 45

nominals
non-maximal, 28
with relative clause, 28
non-core frame element, 11
non-finite clauses, *see* clauses
Non-finite VP, *see* VP
non-maximal adjective, *see* adjective
nouns
and copular sentences, 47
complements of, 28
GFs for, 47

NP
coordinate, 29
non-referential, 27
possessive, 27
standard, 28
types, 26
null instantiation, 12, 20
definite, 12
Null tag, 20
Number, 27
numerals, 37

object, 42
object control
grammatical functions, 44
obliques, *see* complements
Other layer, 20

particles, *see* PP
POS layer, 18
PP
types, 26
complex prepositions, 31
particles, 29, 30, 43
preposition stranding, 31, 32
prepositional verbs, 31
constituency, 31
test for identifying particles, 30
with gerundive complement, 29
with relative clauses, 32
preposition stranding, *see* PP
prepositional verbs, *see* PP
prepositions
external argument of, 46
purpose clauses, 43

Quote, 27
quotes, 37
direct, 37
discontinuities, *see* discontinuities
indirect, 37

raising, 15
relational modification, *see* modifiers
relational modifiers, 22
relational nouns, 16

relative clauses
    nominals with, 28
    reduced, 29, 46

semantic roles, 25
semantic type, 53
semantic types, 23
setting expressions, 43
slot filler, 10
small clauses, 33, 42
subcategorization, 25
subframes, 51
subordinate clauses, *see* clauses
support prepositions, 18
support verb, 22
support verbs, 17, 41
    with adjectives, 44

tags
    general
      copula, 20
text annotation, 24
that, *see* clauses
that-clauses, *see* clauses
to-clauses, *see* clauses
tough-movement, 16
tough-predicates, 42
transparent nouns, 24

undefined frame elements, 13

Verb Phrase, *see* VP
VP, 32
    bare stem infinitives, 33
    finite, 32
    gerundive, 33
    non-finite, 32
    to-marked infinitives, 33
    types, 26

wh-clauses, *see* clauses
whether-clauses, *see* clauses