From Theory to Practice: Frame Semantics and the Design of FrameNet

Hans C. Boas
University of Texas at Austin

1. Introduction

Many of the major research efforts in Natural Language Processing (NLP) over the past two decades has focused on the design of systems for information extraction, question answering, and machine translation. Due to the combination of sophisticated statistical software with template-based algorithms tailored to particular domains, significant progress has been made in these areas of research. A large part of this progress has been made possible because of more refined models that rely on abstract grammatical rules responsible for parsing and generating sentences. As these systems became more sophisticated, it soon became obvious that the lexicon was a major bottleneck in NLP. For example, it was unclear how the vast amount of seemingly unstructured lexical information should be organized in such a way that it would be useful for a wide range of NLP applications. Fellbaum (1998c) summarizes the three main challenges underlying the design of lexical databases in terms of the following three questions:

> "First, how should the lexicon be constructed – by hand or automatically? Second, what kind of information should the lexicon contain? Third, what should the design of the lexicon be – that is, how should its contents be organized and made accessible?" (Fellbaum 1998c: 3)

In this paper I discuss a range of linguistic issues that arise in the design and implementation of large lexical resources for NLP applications. Focusing on two lexical resources, namely WordNet (Miller (1990), Fellbaum (1998b, 1998c)) and FrameNet (Lowe et al. (1997), Baker et al. (1998), Fillmore et al. (2003a)), I examine the different theoretical approaches underlying their design and implementation in order to compare the different types of solutions offered to the three questions above. This discussion not only reveals crucial differences in the structure of the theoretical concepts applied in creating the two lexical resources, it also highlights how these differences have influenced the design of the two databases.

The paper is structured as follows. Section 2 reviews the theoretical concepts underlying the design of WordNet. It also shows how the structure of the WordNet database reflects these theoretical concepts and then discusses the organization of lexical entries in WordNet. Section 3 introduces the key concepts of Frame Semantics and compares and contrasts them with those underlying WordNet. Section 4 discusses how frame-semantic concepts have guided the design of the FrameNet database, demonstrating how the structure of the database represents a straightforward practical application of frame-semantic principles. Section 5 presents a summary and provides an outlook for further research, in particular focusing on the question of how information contained in the two lexical resources can be combined in order to provide users with the greatest possible breadth and depth in semantic and syntactic information.

2. The Structure and Scope of WordNet

In contrast to lexical databases such as COMLEX (MacLeod (1998)) that exclusively focus on the syntactic properties of words, WordNet is a lexical database of English organized in terms of a large semantic network.[1] It combines features of dictionaries and thesauruses in a novel way that makes it possible to discover the subtleties of the internal structure of the lexicon. Describing the meanings of nouns, verbs, adjectives, and adverbs in terms of their links to other words within their syntactic categories via conceptual-semantic and lexical relations, WordNet includes single words, collocations, compounds, and idiomatic phrases. The data included in WordNet come from a variety of sources such as the Brown Corpus, different thesauruses, and the COMLEX lexicon (cf. Miller (1998a: xix)).

Central to the organization of WordNet is the concept of a *synset* (short for synonym set) that encompasses a group of words referring to the same concept. WordNet is thus organized like a thesaurus, but one with built-in hierarchies consisting of conceptual-semantic and lexical relations that link words and synsets to other words and synsets. As such, the meaning of a word is not only accounted for in terms of other words belonging to the same synset, but also in terms of its relation to other words located at different places within the same semantic net. Before reviewing other design features of WordNet, we first look at the workflow and organizational principles underlying the analyses of different parts of speech in WordNet.

The WordNet system consists of four parts: "the lexical source files; the software (the Grinder) used to convert these files into the lexical database; the WordNet lexical database; and the software tools used to access the database." (Tengi (1998: 105)) The core part of WordNet are the lexical source files which are compiled by lexicographers conducting detailed relational analyses of the lexical semantic relations holding between words and their senses. These source files contain synsets from a syntactic category including synonymous words, relational pointers, a gloss, and example sentences. (cf. Tengi (1998: 106)).[2]

The conceptual-semantic relations connecting words differ between the syntactic categories. For nouns, a central relation is that of hyponomy, the IS-A relation, which can have different levels of granularity (cf. Cruse (1986: 88-92)). To exemplify, consider one of the meanings of the noun *fork* denoting cutlery used for serving and eating food. In WordNet, a part of its meaning is given by all the terms referring to different kinds of *forks*, such as *carving fork*, *salad fork*, *table fork*, and *toasting fork*. Similarly, the meanings of *carving fork*, *salad fork*, *table fork*, and *toasting fork* are each defined in part by their superordinate concept *fork*. Distinguishing between three different types of hyponymy, Fellbaum (1998b: 211) points out that WordNet lists "separable parts such as *blade* and *knife* (…) members of groups (like *professor* and *faculty*), and substances such as *oxygen*, which is a component of *air* and *water*." Meronymy, or the part-whole relation

[1] This section gives a brief summary of some of the most important features of WordNet in order to set the stage for our discussion of the organization of FrameNet in subsequent sections. It does not address non-literal language or semantic disambiguation (for this and other details, see Fellbaum (1998a)). Many of the examples in this section are taken from Fellbaum (1998b, 1998c).

[2] For a more detailed description of the workflow and organization underlying WordNet, see Tengi (1998).

(cf. Saeed (cf. 1997: 70-71)), is another important relation holding between noun synsets. Meronymy reveals hierarchical classifications in the lexicon that are similar to taxonomies. For example, *a valve* is a part of *an engine*, which in turn is a part of *a car*.[3] One advantage of structuring a semantic net for nouns in this way is that it allows for the discovery of lexical gaps in cross-linguistic research. Fellbaum (1998b: 211) cites the example of French *rentrée* whose English counterpart consists of a multi-word expression, namely "the start of the school or university year." She points out that "these gaps are not always structural artifacts; they are often lexicalized in other languages and they reveal conceptual structures as distinct from lexical structures." (1998c: 6)[4]

In contrast to nouns, adjectives in WordNet are classified according to different conceptual-semantic relations for the following reasons: (1) their meanings crucially depend on the head noun they modify; (2) high frequency adjectives like *good*, *new*, and *big* are highly polysemous whereas lower frequency adjectives such as *steep* and *international* are less polysemous; (3) adjectives do not exhibit an organizational hierarchy similar to that of nouns and verbs (cf. Fellbaum (1998b: 212)). These differences have led WordNet to capture the conceptual-semantic relations among qualitative adjectives not in terms of synsets, but rather in terms of so-called "direct" and "indirect" antonymic relations. The idea is that semantically related adjectives constitute a cluster whereas their antonymic counterparts constitute an opposite cluster (Gross et al. (1989), Fellbaum (1998b), K.J. Miller (1998)). Direct antonyms such as *long* and *short* exhibit high frequency, are psychologically salient, and "constitute a conspicuous but small part of the adjective lexicon." (Fellbaum 1998b: 212) As Fellbaum (ibid.) points out, adjectives "like *abbreviated* and *eternal* are considered "indirect antonyms" of *long* and *short*, respectively. The organization of adjectives in WordNet can thus be visualized in terms of barbell-like structures, with a direct antonym in the center of each disk surrounded by its semantically similar adjectives (which constitute the indirect antonyms of the adjectives in the opposed disk)." In contrast to qualitative adjectives such as *long* and *short*, relational adjectives such as *atomic* and *industrial* are not represented in terms of antonymy, but rather in terms of the nouns from which they are derived (*atom*, *industry*) (cf. Fellbaum (1998b: 213). Adverbs in WordNet are treated in terms of *–ly* affixation to corresponding adjectives, thereby following their lexical organization. WordNet does not propose any specific implementation for the treatment of lexical adverbs such as *hard* and *even* (cf. Fellbaum (1998b: 213)).[5]

In contrast to other parts of speech, verbs are organized in terms of a different lexical relation for the construction of lexical hierarchies. Troponymy is a lexical relation that relates one verb in terms of a more specific instantiation of an associated, more abstract verb. For example, lower-level verbs such as *limp* or *amble* are troponyms of *walk* because they represent a more specific way of walking. Troponymy can be organized according to various types of manner or means distinctions. As Fellbaum (1998b: 213) points out, verbs of motion "are semantically elaborated along such dimensions as speed (*walk – run*), direction (*rise – fall*), and means of displacement (*walk – drive*)." Troponymy is a polysemous relation that also represents semantic opposition among verbs (*rise – fall*, *come – go*) (cf. Fellbaum (1998b: 213)). Another important

[3] On the transitivity relations of hyponymy and meronymy relations, see Fellbaum (1998b: 211).
[4] For more details on the treatment of nouns in WordNet, see Fellbaum (1998b) and Miller (1998b).
[5] For more details on the treatment of adjectives and adverbs in WordNet, see K.J. Miller (1998).

relation captured by WordNet is that of lexical entailment.[6] "For example, eating entails swallowing: When people eat, they necessarily swallow." (Fellbaum (1998b: 214), (1998d: 77-79))

This brief review of the organizational principles underlying WordNet's treatment of different parts of speech has shown two important points. First, each syntactic category is analyzed separately from other syntactic categories. Second, the organizational principles used for the construction of lexical hierarchies and relations such as meronymy, hyponymy, and troponymy are largely based on "traditional" concepts in semantics (e.g., Lyons (1977), Cruse (1975, 1977, 1986)). As such, WordNet presents a practical implementation of semantic concepts that have been applied by a variety of frameworks.

Another "traditional" organizational concept reflected by WordNet's structure is that of polysemy (for an overview, see Ravin & Leacock (2000)). WordNet makes fine-grained distinctions between different senses of words by including traditional lexicographic glosses and definitions. Whenever groups of words exhibit regular and predictable polysemy, they are grouped together in order to illustrate the tight semantic relation holding between the individual senses (Fellbaum (1998b: 213)). For example, the nouns *magazine* and *newspaper* are so-called cousins in WordNet because they exhibit a type of regular polysemy due to the types of superordinates they share (*publication*, *publishing house*, and *product*) (cf. Fellbaum 1998b: 214)). Similarly, polysemous verbs whose senses share certain semantic features are part of the same group.

Having outlined the theoretical concepts underlying the organization of WordNet, we now turn to their practical implementation by discussing different types of lexical information contained in the database.[7] We first review the treatment of verbs, and then we turn to nouns.

2.1 Verbs in WordNet

The organization of WordNet is reflected by the multitude of search options provided by the on-line interface, which allows for verb searches (including their synsets) in many ways. For example, it allows to search for "synonyms, ordered by estimated frequency," "synonyms, grouped by similarity," "coordinate terms," "hypernyms," "troponyms," "V entails doing", "derivationally related forms", "sentence frames," "domains," and "familiarity." Consider *to cure* whose partial entry has four senses including synonyms, with the most frequent one listed first.

[6] For more details on the treatment of verbs in WordNet, see Fellbaum (1998b, 1998c, 1998d).

[7] The WordNet entries are from WordNet 2.0 via its on-line interface at http://www.cogsci.princeton.edu/~wn/. For different methods for accessing WordNet, see Tengi (1998: 122-127).

(1) Partial information for *to cure* in WordNet 2.0, including synonyms ordered by estimated frequency and sentence frames.[8]

Sense 1: bring around, **cure**, heal -- (provide a cure for, make healthy again; "The treatment cured the boy's acne"; "The quack pretended to heal patients but never managed to") => help, aid – (improve the condition of; "These pills will help the patient")
*> Somebody ----s something
*> Somebody ----s somebody

Sense 2: **cure** -- (prepare by drying, salting, or chemical processing in order to preserve; "cure meats"; "cure pickles") => preserve, keep – (prevent (food) from rotting; "preserved meats"; "keep potatoes fresh")
*> Somebody ----s something

Sense 3: **cure** -- (make (substances) hard and improve their usability; "cure resin") => harden, indurate – (become hard or harder; "The wax hardened")
*> Somebody ----s something
*> Something ----s something

Sense 4: **cure** -- (be or become preserved; "the apricots cure in the sun") => change -- (undergo a change; become different in essence; losing one's or its original nature; "She changed completely as she grew older"; "The weather changed last night")
*> Something ----s

The structure of the entry for *to cure* illustrates that for each sense of a verb, WordNet provides a number of synonyms that are part of the same synset, followed by a gloss that includes a short definition as well as examples sentences illustrating the use of the word(s). "In most cases, each line of search results is preceded by a marker (usually =>), then a synset. If a search traverses more than one level of a tree, subsequent synsets are indented by spaces corresponding to their level in their hierarchy." (Tengi (1998: 122)) WordNet also provides information about hypernyms higher up in the hierarchy. For *to cure* this means that above the level of *to help* and *to aid* we find *to better*, *to improve*, *to amend*, among others. Above that level, WordNet provides *to change*, *to alter*, and *to modify* as hypernyms.[9]

One important point concerns the treatment of syntactic information in WordNet. That is, although the meaning of each word and its senses is defined in detail (in combination with a gloss and an illustrative example sentence), the syntactic information provided by WordNet is rather scarce. The reason is that WordNet "was conceived as a

[8] Note that it is difficult to present the multitude of information that WordNet has to offer in print. The information in (1) represents the combined output from searches for "synonyms, ordered by estimated level of frequency" and "sentence frames" for *to cure*. For formatting reasons, the original web-based layout has been slightly altered here. For more details on how to search WordNet, see Tengi (1998).

[9] As already pointed out, WordNet offers such a multitude of different search options providing information about how a word is related semantically to other words, that a full discussion would go well beyond the scope of this paper.

semantic database only." (Fellbaum (1998c: 11)) To illustrate, consider the verb *to cure*. Besides the example sentences included in the gloss, the sentence frames "Somebody ----s something" and "Something ----s" state that different senses of *to cure* may occur with a transitive as well as with an intransitive syntactic frame. It does not mention that sense 1 may also occur with other syntactic frames including additional constituents such as Prepositional Phrases (expressing the treatment as in *They cured his hangover with herbs*).

2.2 Nouns in WordNet

With this brief review of how WordNet describes semantic and syntactic properties of verbs, we now turn to the treatment of nouns in WordNet. To illustrate, consider the noun *cure*, which has one sense in WordNet.

(2) Partial information for *cure* in WordNet 2.0, including hypernyms (this is a kind of …)

Sense 1
remedy, curative, cure -- (a medicine or therapy that cures disease or relieves pain)
=> medicine, medication, medicament, medicinal drug --
((medicine) something that treats or prevents or alleviates the symptoms of disease)
=> drug -- (a substance that is used as a medicine or narcotic)
=> agent -- (a substance that exerts some force or effect)
=> causal agent, cause, causal agency --
(any entity that causes events to happen)
=> entity --
(that which is perceived or known or inferred to
have its own distinct existence (living or nonliving))

The information in (2) lists *remedy* and *curative* as members of the synset of *cure* and provides a gloss with a definition. WordNet also lists five levels of hypernyms of *cure* with their respective glosses representing hierarchical families of lexicalized concepts. Thus, we learn that *cure* may be a type of *medicine*, which in turn is a type of *drug*, which is a type of *agent*, and so on. Besides information on hypernyms, WordNet offers a wealth of other data on conceptual-semantic and lexical relations among nouns. They include "synonyms, ordered by estimated frequency," "coordinate terms," "hyponyms (… is a kind of …), brief," "hyponyms (… is a kind of …), full," "holonyms (… is a part of …), regular," "meronyms (parts of …), inherited," "derivationally related forms," and "familiarity."

Comparing the information provided by WordNet for nouns with that provided for verbs, we find two main differences. First, verbs and nouns differ in the number and types of conceptual-semantic and lexical relations. Second, WordNet provides some syntactic subcategorization information for verbs in the form of "sentence frames" and example sentences. In contrast, it contains little information about the types of syntactic

structures in which nouns occur by only including occasional example sentences in the definitional glosses.

2.3 Summary

Our discussion of the many different types of information provided by WordNet about verbs and nouns has shown that its organization relies heavily on "traditional" lexical semantic relations such as hyponymy, synonymy, meronymy, and polysemy, among others. As such, the unique architecture of WordNet reflects a straightforward implementation of many linguistic and psychological concepts that had been the focus of research in other areas of (psycho-)linguistic research before the inception of this large-scale relational lexicon.[10] Keeping in mind the fact that WordNet "was conceived as a semantic database only" (Fellbaum (1998c: 11), we now turn to a comparison of the design and implementation of WordNet with that of a different lexical resource organized by semantic principles, namely FrameNet. We begin with a discussion of Frame Semantics, the semantic theory underlying the design of FrameNet.

3. Frame Semantics

Frame Semantics is "a research program in empirical semantics and a descriptive framework for presenting the results of such research" (Fillmore (1982: 111)) developed by Fillmore and his associates over the past three decades (Fillmore (1970, 1975, 1976, 1977a, 1977b, 1982), Fillmore & Atkins (1992, 1994, 2000), among many others). This approach differs from other theories of lexical meaning in that it builds on common backgrounds of knowledge (semantic 'frames') against which the meanings of words are interpreted.[11] A "frame is a cognitive structuring device, parts of which are indexed by words associated with it and used in the service of understanding." (Petruck (1996: 2)) The central ideas underlying Frame Semantics are summarized by Fillmore & Atkins (1992) as follows.

> A word's meaning can be understood only with reference to a structured background of experience, beliefs, or practices, constituting a kind of conceptual prerequisite for understanding the meaning. Speakers can be said to know the meaning of the word only by first understanding the background frames that motivate the concept that the word encodes. Within such an approach, words or word senses are not related to each other directly, word to word, but only by way

[10] Besides its usefulness as a dictionary and thesaurus, it has inspired the creation of WordNets for a variety of other languages which, in turn, has led to the discovery of differences in cross-linguistic patterns of lexicalization (Vossen (1998)). In addition, WordNet has formed the basis for practical applications such as word sense disambiguation (Leacock & Chodorow (1998)), information retrieval (Al-Halami & Kazman (1998), Voorhees (1998)), and knowledge engineering (Burg & van de Riet (1998)).

[11] Fillmore's use of the concept of "frame" is somewhat related to work in artificial intelligence. For example, Minsky (1975: 212) describes a frame as a "data-structure representing a stereotypical situation." Work in psychology employs a similar concept that refers to knowledge structures for sequences of events, cf. Schank & Abelson's (1975) "restaurant script." (cf. Boas (2003: 164)) For differences between Frame Semantics and semantic field theories, see Fillmore & Atkins (1992: 76-79).

> of their links to common background frames and indications of the manner in which their meanings highlight particular elements of such frames.
>
> (Fillmore & Atkins (1992: 76-77))[12]

3.1 Frame Descriptions and Frame Elements: The Theft Frame

To illustrate, consider the Theft frame, which involves several semantically related verbs such as *steal*, *snatch*, *shoplift*, *snitch*, *pinch*, *filch*, *purloin*, and *thieve*, among others. The Theft frame represents a scenario with different frame elements that can be regarded as instances of broader semantic roles such as AGENT, UNDERGOER, INSTRUMENT, etc.[13] Giving precise definitions for frame elements is important because the entirety of frame elements comprises the frame description, which in turn represents a schematic arrangement of the situation type that underlies the meanings of semantically related words. Compare the following sentences.

(3) a. Nikki stole the watch from Carolyn.
b. Jana nicked the book from Vaughan.
c. Guido pinched the disk from the table.
d. Ingrid filched the snack from Karen.

In (3a) – (3d), the Theft frame is evoked by the verbs *steal*, *nick*, *pinch*, and *filch*. This frame represents a scenario with different core frame elements such as GOODS (anything that can be taken away), PERPETRATOR (the person or other agent that takes the goods away), SOURCE (the initial location of the goods before they change location), and VICTIM (the person (or other sentient being or group) that owns the goods before they are taken away by the perpetrator). The frame description defines the relationships between frame elements of the same frame. For the Theft frame, the frame description states that a PERPETRATOR takes GOODS that belong to a VICTIM.

To exemplify, *stole* in (3a) is the target word that evokes the Theft frame. *Nikki* is the PERPETRATOR frame element, *the watch* is the GOODS frame elements, and *from Carolyn* is the VICTIM frame element. In example (3c), *from the table* is the SOURCE frame element. The needed background to interpret the verbs in (3a) – (3d) as belonging to the THEFT frame requires an understanding of illegal activities, property ownership, taking things, and a great deal more. This example illustrates that with a frame-semantic description it is possible to describe the meanings of a variety of semantically related verbs in relation to the same frame.

The key role of the frame as a conceptual underpinning for a frame-based organization of the lexicon can also be seen in the description of other parts of speech. For example, the properties of nouns that are semantically related to the verbs in (3a) – (3d) can also be described by appealing to the Theft frame. That is, agentive nouns such as *shoplifter*, *snatcher*, *stealer*, *thief*, and *pickpocket* evoke the same Theft frame as the

[12] For a more detailed review of the main principles of Frame Semantics, see Petruck (1996) and Fillmore et al. (2003a).

[13] For an overview of different characterizations of semantic roles (also known as theta-roles), see, e.g., Fillmore (1968, 1975, 1985a), Jackendoff (1990), Langacker (1990), Ravin (1990), Dowty (1991), and Van Valin & Wilkins (1996).

verbs above. In a sentence such as *The cattle stealer rode home* the noun *stealer* evokes the Theft frame while simultaneously instantiating the PERPETRATOR frame element. In contrast, *cattle* represents the frame element GOODS. Appealing to the Theft frame also allows us to capture the semantic and syntactic properties of event nouns such as *stealing*, *thieving*, and *pilfering*. Consider the sentence *The stealing of food is a problem*. Here, *stealing* is the target word evoking the Theft frame, while *of food* represents the frame element GOODS. Besides describing the properties of verbs and nouns, adjectives such as *light-fingered*, *thieving*, and *stolen* can also be analyzed using the Theft frame. In a sentence like *A stolen car was left behind*, the adjective *stolen* is the target word evoking the Theft frame, whereas *car* represents the frame element GOODS.

Besides so-called core frame elements there are other frame elements that are peripheral from the perspective of the Theft frame such as MEANS (e.g. *by trickery*), TIME (e.g. *two days ago*), MANNER (e.g. *quietly*), or PLACE (e.g. *in the city*). These frame elements do not belong to the set of core elements of the Theft frame because they are also found among other frames of agentive action. With this short overview of the basic concepts of Frame Semantics let us now look at how this semantic theory was developed and elaborated to eventually become a blueprint for the design and implementation of FrameNet.

3.2 Frame-semantic Influences on Lexicology and Lexicography

Among the key concepts further developed in Frame Semantics is the idea of creating a description of the relations between lexical meanings and the syntactic structures of sentences containing words with those meanings. Capturing detailed syntactic information and linking them to semantic descriptions has its roots in Fillmore's seminal (1968) paper *The Case for Case*, and was refined in his subsequent work (1970, 1975, 1977a), eventually leading to his (1977b) paper on *Topics in Lexical Semantics.* In this paper, Fillmore points out a number of important issues that semantic theory needs to address with respect to syntax: "The difference between obligatory and optional elements; the difference between nuclear and nonnuclear elements; the grammatical markings of the elements in the periphery." (Fillmore (1977b: 102)) The "labeled box notation initially suggested as an informal representation system for the lexicon (Fillmore (1977b)) was refined and used for the representation of grammatical constructions in the grammatical framework developed by Fillmore and his colleagues, construction grammar (Fillmore (1985b, 1988))." (Petruck (1996: 7)) The crucial link between frame-semantic and syntactic information was not only central to subsequent work in construction grammar (Fillmore & Kay (1993), Goldberg (1995)), but also became the focus of application of frame semantic principles to lexicology and lexicography. The value of incorporating syntactic information in a word's frame-based lexical entry is made explicit in a number of papers (Fillmore & Atkins (1992, 1994), Fillmore (1994), and Atkins (1994, 1995)) dealing with lexical description and dictionary design.

3.2.1 The Importance of Lexico-syntactic Patterns

In understanding the development of Frame Semantics and its direct application to the design of FrameNet, Fillmore & Atkins (1992) is the most significant work because it is

here that it is first explicitly proposed that the types of information contained in a frame-based lexicon should include "relationships between particular lexical meanings and specific lexico-syntactic patterns." With respect to the overall structure of such a frame-based lexicon, Fillmore & Atkins point out that "each of these lexico-syntactic patterns will have its components indexed with specific parts or aspects of the associated frame." (1992: 75)

The authors exemplify their approach by utilizing a large data corpus in addition to traditional print dictionaries to provide a detailed analysis of the lexeme *risk* (both noun and verb). The description of the verb *risk* consists of a list of more than twenty entries that explicitly combine frame-semantic with syntactic information. The box in (4) represents only a partial lexical description of the verb *risk* and illustrates that in one of its usages, it may be followed by a noun phrase representing the frame element VALUED OBJECT (VO) and a prepositional phrase representing the frame element SITUATION (Sit).[14]

(4)

RISK $^{\text{VO}}${**NP**} $^{\text{Sit}}${**Prep NP**}

He was being asked to risk
$^{\text{VO}}${his good name}
$^{\text{Sit}}${*on* the battlefield of politics}.

Others had risked
$^{\text{VO}}${all}
$^{\text{Sit}}${*in* the war}. (Fillmore & Atkins (1992: 87))

The lexical description of *risk* advocated by Fillmore & Atkins is not only exemplary in that it combines syntactic and frame-semantic information with example sentences taken from an electronic corpus. It is also different from previous approaches in that it provides a detailed account of all the different ways in which frame elements can be realized syntactically (the authors list a total of 25 syntactic frames for *risk*). Providing an exhaustive account of a word's semantic and syntactic combinatorial possibilities is one of the major advantages of FrameNet, as we will see in section 4.

Fillmore & Atkins' (1992) paper represents a crucial milestone on the way to the creation of FrameNet because it is in this paper that we first find a vision of how a frame-based dictionary will look like one day. The authors suggest that a frame-based dictionary should be "housed on a workstation with multiple windowing capabilities" (1992: 75) that would function as follows:

> "A user's keying in of a word to be looked up will cause a window to appear that will display relationships between particular lexical meanings and specific lexico-syntactic patterns. Each of these lexico-syntactic patterns will have its

[14] For a detailed description of the frame elements CHANCE, HARM, VALUED OBJECT, SITUATION, DEED, ACTOR, GAIN, PURPOSE, BENIFICIARY, and MOTIVATION in the RISK frame, see Fillmore & Atkins (1992: 79-84).

components indexed with specific parts or aspects of the associated frame." (Fillmore & Atkins (1992: 75))

The advantage of such a system is that it would go well beyond the limitations of traditional commercial dictionaries in terms of types and amounts of data provided: "Another reason is that research into the lexicon inevitably uncovers much more information about words than standard dictionaries have room for." (1992: 76) This idea is expanded upon in a subsequent paper (Fillmore & Atkins (1994)), where the authors point out that "a description of word meaning in terms of frame semantics requires a new kind of dictionary, one which is not limited to the printed page, in which access to the information is not constrained by alphabetical order, and approach to the contents is not linear." (1994: 375) As we will see in section 4, these programmatic proposals formulated by Fillmore & Atkins (1994) form the basis for the architecture of FrameNet.

3.2.2 Polysemy

Another key point raised by Fillmore and Atkins concerns the treatment of polysemy in an on-line frame-based lexicon. They point out that "the concept of 'frame' makes it possible to reconsider the notion of polysemy, and to develop the consequences of this reconsideration for lexicography." (1992: 101) On this view, a frame-semantics approach makes it possible to move beyond traditional concepts such as dictionary senses that have to discover, record, and define so-called central and extended senses.[15] To illustrate, Fillmore & Atkins (1994: 357) discuss the following three senses of *risk*.

(5) a. risk doing something that could result in losing his life
b. risk losing his life
c. risk his life (1994: 57)

They point out that whereas traditional dictionaries identify sense (c) as the primary sense, it is really sense (a) that is the basic sense, whereas (b) is metonymically derived from (a), and (c) is metonymically derived from (b). In a frame-semantic approach, it is possible to recognize the interphrasability of these related senses and to describe them with respect to the same frame. In other words, by specifying the interrelations between semantic frame and syntax, differences in syntactic realization can be neatly captured by appealing to the common conceptual underpinning of the same frame.

An added benefit of this approach is that it is possible "to link obviously related noun and verb senses (…) and to show clear relationships among the senses, across nouns and verbs, through their interparaphrasability." (Fillmore & Atkins (1994: 357)) For example, a frame-semantic approach makes it possible to link a specific sense of *to risk* ("to incur the chance of unfortunate consequences by (doing something): *to risk climbing the cliff*" (1994: 352)) with a specific sense of its nominal counterpart ("dangerousness, a dangerous situation: *too much risk attached to it; the risks involved*" (1994: 354)). Similarly, a different sense of *to risk* ("to do something in spite of the possibility of

[15] See Fillmore & Atkins (1994: 369-370) on additional differences between the traditional lexicography approach to describing word meanings and the frame-semantic approach.

(unfortunate consequences): *to risk falling; to risk a fall*" (1994: 352)) can be linked to a different sense of its nominal counterpart ("unpleasant possibility: *the risk of being killed/of an attack*" (1994: 354)) because they evoke the same underlying frame.

3.2.3 The Role of Corpus Examples

The inclusion of corpus examples as a part of frame-semantic descriptions is another important idea developed in the 1990s. This idea has its roots in Fillmore & Atkins' (1992, 1994) discussions about how *risk* is described by a variety of traditional dictionaries. The authors point out that print dictionaries do not agree on how *risk* functions in everyday language. This problem is traced back to the difficulties that traditional dictionaries have with "(a) sense differentiation in the verb and noun; (b) distinction between 'run a risk' and 'take a risk'; (c) patterns of verb complementation." (1994: 363) In order to produce more accurate and comprehensive descriptions of how words are used, Fillmore and Atkins recommend using "sophisticated computational tools for lexical analysis" (1994: 376), in particular a "tagged and parsed corpus" (1994: 377). Building on this idea Atkins (1995: 39) points out that by building on semantic frames as common structuring devices for the lexicon, "the examples constitute an integral part of the description of meaning and appear whenever they are needed to illustrate a point." Including numerous examples has the added benefit that "every fact would be accompanied by an example – or rather, by many examples, for the user would be able to call up more at will; each fact stated formally would be linked to the corpus usages exemplifying it." (Atkins (1995: 40))

Having summarized the basic principles of Frame Semantics as well as their further theoretical developments in lexicology and lexicography in the 1980s and 1990s, I now discuss their practical implementation in the context of FrameNet. Throughout the following section, I also compare how the information provided by FrameNet differs from that provided by WordNet.

4. FrameNet

The FrameNet database represents the product of a project in computational lexicography that is based on the principles of Frame Semantics. It contains lexical entries for words, descriptions of frames (including their frame elements), annotated subcorpora, as well as sense descriptions. During its six-year period, the FrameNet team collected and analyzed lexical descriptions of more than 7,000 lexical units based on more than 130,000 annotated corpus sentences (Fillmore et al. (2003a)). In contrast to other lexical resources such as COMLEX or WordNet, FrameNet considers the lexical unit (Cruse (1986: 23-48)) as the primary unit of analysis whose semantic and syntactic properties are described with respect to a semantic frame. A lexical unit is defined as a pairing of a word with a particular sense that evokes a semantic frame. For example, *cure* as in *They cured the patient* and *They cured the pork* would be analyzed as evoking two different frames ('cure' vs. 'preserving') which means that there are (at least) two distinct lexical units for this verb. This categorization differs from that employed by WordNet in that it defines the different senses of *cure* (as well as semantically related nouns and adjectives) with reference to different underlying semantic frames instead of other verbs belonging to two

different synsets (see section 2.1 above). As the following sections illustrate, the practical application of Frame Semantics to lexical description also structures the workflow of the project and the structure of the database.

4.1 Workflow

The core of FrameNet activities lies in the identification and description of semantic frames as lexical structuring devices. Meanings of words as well as their lexico-syntactic properties are then analyzed by directly appealing to the frames. The workflow roughly proceeds as follows. First, a frame to which a given lexical unit belongs is identified and described. This is done by "characterizing schematically the kind of entity or situation represented by the frame." (Fillmore et al. (2003b: 297)) Each of the entities of the frame is given frame element (FE) names as mnemonic labels. Next, a list is compiled that contains words whose lexical units are semantically related and might therefore be described by using the same frame. Based on the syntactic and collocational contexts in which a prototypical member of the word list occurs in the British National Corpus, a thorough corpus search is conducted automatically for each lexical unit. The resulting subcorpora are then semantically annotated by hand. During this step, annotators carefully choose representative instances of each lexical unit and mark constituents with frame element labels. Subsequently, an automated script adds information about the phrase type (PT) and grammatical function (GF) to the annotation, thereby arriving at an inventory of syntactic valence patterns.[16] The resulting annotation consists of various layers of syntactic and semantic information as the following example shows.

(6) Layered Annotation including frame elements (FE), phrase types (PT), and grammatical functions (GF).

(Text)	*Joe*	*brushed*	*the rolls*	*with butter*
FE	AGENT		GOAL	THEME
PT	NP		NP	PP
GF	Ext		Obj	Comp

The table in (6) illustrates only one combination of frame elements occurring with the lexical unit *brush* as the target in the Filling frame. *Joe* is the AGENT, *the rolls* is the GOAL, and *with butter* is the THEME frame element. As can be seen, each constituent (except for the target) is also annotated with information about its PT (noun phrase, prepositional phrase) and GF (external argument, object, and complement). Besides the syntactic valence pattern in (6), there are at least five other valence patterns exemplifying how the semantics of *brush* may be realized differently. This small example shows that the breadth and depth of information recorded for each lexical unit during this process

[16] Note that FrameNet's workflow is not strictly linear as one might conclude form the brief description here. For example, since the workflow is corpus-driven it may happen that new insights about a lexical unit's distribution emerge during the subcorporation or annotation process. In such cases, frame descriptions or frame elements have to be revised in order to accommodate the discoveries made (see Fillmore et al. (2003b) for more details).

reflects the implementation of the detailed methodological groundwork laid by Fillmore and Atkins (1992, 1994) (cf. section 3.2). More specifically, it reveals how careful considerations of theoretical concepts from Frame Semantics structure the workflow and design of FrameNet.

The work leading to the discovery and description of lexical information is supported by a variety of software tools (the so-called integrated FrameNet desktop). These tools help with defining frames, their frame elements, and relations between frames, conducting corpus searches, annotating example sentences with frame-semantic information, automatically adding syntactic information to the frame-semantic annotation, and storing all of the information in the FrameNet database.[17] This MySQL database consists of two parts, the lexical database and the annotation database. The former contains, among other things, information about frames, frame elements, lemmas, lexemes, parts of speech, lexical units, as well as information about the relations holding between these elements. The latter is linked to the former and contains annotated corpus sentences exemplifying the distribution of lexical units in the British National Corpus.[18]

4.2 The Implementation of Frame Semantics in the FrameNet database

The FrameNet data can either be exported in XML format for further use in other NLP applications (cf. Baker et al. (2003: 292-293)), or they can be accessed via web-based browsing or search interfaces that display the data in human-readable format (see http://www.icsi.berkeley.edu/~framenet). Both interfaces rely on automatically generated data reports based on completed annotations contained in the FrameNet database. They make it possible to display the semantic and syntactic combinatorial possibilities of each lexical unit in many different ways.

4.2.1 Verbs in FrameNet

To illustrate the range of data FrameNet provides for verbs, we return to the verb *cure* whose analysis in WordNet we reviewed in section 2. FrameNet lists two distinct lexical units according to the semantic frames to which they belong, namely Cure and Preserving. Clicking on the "annotation" link for *cure* in the Cure frame creates an annotation report with the Frame Element Table (see Figure 1 in the appendix). It lists all of the frame elements of the Cure frame such as BODYPART, AFFLICTION, PATIENT, and HEALER (among others) marked with different color tags. This catalog is followed by annotated corpus sentences ordered by syntactic frames. Each sentence displays the words that instantiate the frame elements, highlighted with the same colors as the frame elements in the Frame Element Table. This display thus allows for easy identification of frame elements and their syntactic realizations in corpus sentences, just as Fillmore & Atkins (1992, 1994) envisioned it in their programmatic papers on the structure of frame-based lexicons.

This information differs from that in WordNet in (at least) three significant ways. First, the semantic description is frame-based, thereby appealing directly to a collection

[17] For more detailed information on the workflow of FrameNet as well as the tools employed, see Fillmore et al. (2003b).
[18] For a comprehensive review of the structure of the FrameNet database, see Baker et al. (2003).

of knowledge encoded in the description of the Cure frame. In contrast, WordNet's description includes a definition of the verb in the traditional sense in combination with a gloss and a list of other members belonging to the same synset.[19] Second, FrameNet provides an extensive list of corpus-based example sentences illustrating how the semantics of *to cure* are realized by many syntactic patterns. WordNet contains one example sentence illustrating the use of *cure* and provides two "sentence frames" illustrating its syntactic properties. Third, whereas WordNet includes information about frequency, hypernyms, troponyms, or familiarity (among others) of *cure*, FrameNet has little information to offer in comparison. A FrameNet concept somewhat comparable to that of hyponymy in WordNet is that of Frame Inheritance where frames may inherit frame-semantic information from more abstract frames (cf. Baker et al. (2003: 286)). In this case, the Cure frame inherits information from the Intentionally_affect frame.[20] As one of FrameNet's primary goals is to document all of a word's lexicographically relevant uses found in the corpus, it records them, but without inclusion of frequency information.

Another feature of frame-based dictionaries proposed by Fillmore & Atkins (1992) is one that allows "the user who wishes to be reminded of the properties of the frame associated with a given word to open an additional window that presents information about it." (1992: 75) This theoretical proposal, like the others discussed so far, is realized in FrameNet by providing users with detailed information about underlying frames. For example, for *cure* users may click on the "frame" link for the Cure frame, which opens a new window with its frame description ("This frame deals with a Healer treating and curing an Affliction (the injuries, disease, or pain) of the Patient, sometimes also mentioning the use of a particular Treatment or Medication.") (cf. Figure 2 in the appendix). Following the frame description we find a detailed list of frame elements, their descriptions, and example sentences showing how these frame elements are realized in context. Each frame element contained in these descriptions is highlighted with the same colors that are used to mark frame elements in the annotation report. In addition, each frame description provides information about how the frame is related to other frames. These relations include Frame Inheritance, The Subframe Relation, The Uses Relation, and the 'See also' Relation (for details, see Baker et al. (2003: 286-287)). Finally, each frame description includes a list of all verbs, nouns, and adjectives that evoke the frame. As already pointed out in the previous paragraph, it is this reference to semantic frames that sets lexical descriptions in FrameNet apart from those in WordNet.

Another feature particular to FrameNet, and already addressed briefly above, concerns the representation of syntactic information in FrameNet. This feature, too, was already conceived of by Fillmore & Atkins (1992) in their work on applying Frame Semantics to applied computational lexicography. They proposed that a frame-based dictionary should "display relationships between particular lexical meanings and specific

[19] This structural difference also explains why the intransitive and transitive versions of *cure* in its food preserving sense have two separate entries in WordNet. In FrameNet, both versions are described in terms of one lexical unit evoking the same Cure frame. The syntactic differences are thus expressed in terms of different valence patterns.

[20] This frame describes situations in which an AGENT causes a PATIENT to be affected, sometimes by a particular MEANS or by use of an INSTRUMENT.

lexico-syntactic patterns. Each of these lexico-syntactic patterns will have its components indexed with specific parts of aspects of the associated frame." (1992: 75) The practical implementation of this proposal reveals how tightly FrameNet's design is based on the early ideas about the practical applications of Frame Semantics. Consider, for example, the types of syntactic information offered by FrameNet for verbs.

The first set of syntactic information is contained in a so-called Realization Table (see Figure 3 in the appendix). It presents a definition from the *Concise Oxford Dictionary* and lists the different syntactic realizations of the core frame elements (including the number of annotated example sentences) in a table. For example, for *cure* the Realization Table shows that FrameNet contains 16 annotated example sentences for the frame element PATIENT. It also informs the user that PATIENT has four different syntactic realizations, namely as NPs (occurring either as an external argument or as a direct object) or as various types of PPs.

The second set of syntactic information is presented in a Valence Table (see Figure 4 in the appendix). It displays the core frame elements and provides an exhaustive list of how they are realized syntactically (in terms of Phrase Type and Grammatical Function) in corpus sentences. For the combination of AFFLICTION, HEALER, and TREATMENT, there are four different syntactic realizations with *cure*: (1) [NP, NP, NP]; (2) [NP, NP, PP[with]]; (3) [NP, NP, PP[without]]; (4) [NP, Sfin, PPing[by]].[21] As with the other displays of lexical information, FE names in the Valence Table are highlighted with the same colors for better readability.[22] Keeping in mind the lexical description offered by FrameNet for the verb *cure*, we now turn to the description of its nominal counterpart to see how it differs from that offered by WordNet.

4.2.2 Nouns and Adjectives in FrameNet

One of the major points proposed in Fillmore & Atkins' (1992) seminal paper concerns the organization of the lexicon as a whole. They suggest that the "frame descriptions will themselves contain pointers allowing access to other expressions in the language whose meanings are founded on the same schema." (1992: 76) This programmatic suggestion has been implemented in FrameNet not only for words that belong to the same part of speech (e.g., such as *cure*, *alleviate*, *heal*, and *palliate*). It has also been implemented for other parts of speech such as nouns and adjectives. For example, the noun *cure* is described with respect to the same Cure frame, because it evokes the same types of knowledge necessary to interpret its verbal counterpart. FrameNet provides a complete semantic and syntactic description of the noun *cure* in terms of separate Frame Element, Frame Realization, and Valence Tables (see Figure 5 in the appendix) which all refer to

[21] Because of limited space, I have left out the Grammatical Function information. For details, please see the FrameNet website at http://www.icsi.berkeley.edu/~framenet.

[22] Other types of syntactic information captured by FrameNet are so-called null instantiations and 'Gov-X' annotations. The former refer to "frame elements that are conceptually necessary but do not occur as lexical or phrasal material." These missing elements are annotated and occur as parts of lexical entries. The latter refer to nouns that name artifacts and natural kinds that "serve as slot fillers for frames evoked by verbs, adjectives, or event nouns." By annotating these special nouns with respect to the types of predicates that govern them (typically verbs), it becomes possible to establish principled semantic generalizations between governing and governed words belonging to different frames (for more details, see Fillmore et al. (2003b: 320-321)).

the same Cure frame description discussed above. The same holds for adjectives such as *curable*, *curative*, *palliative*, and *therapeutic*, which evoke the same semantic frame (see Figure 6 as an example). These examples illustrate that FrameNet differs in its organization from WordNet in that it describes verbs, nouns, and adjectives with respect to the same underlying semantic frame. In other words, all three parts of speech have the capacity to evoke the same types of semantic knowledge. In contrast, lexical descriptions in WordNet are defined via relations in semantic-conceptual networks that are different for each part of speech.[23] This difference in lexical organization is also reflected in the treatment of polysemy in the two lexical databases.

4.2.3 Polysemy

Similar to other design features of FrameNet, its treatment of polysemy was first envisioned by Fillmore & Atkins: "We need the means of associating a word (or a group of words, or a group of word uses) with particular semantic frames, and then to describe the varying ways in which the elements of the frame are given syntactic realization." (1992: 101) This proposal led the authors to suggest that one "should not have to regard each of these varying mappings as different senses of the word." (ibid.) Our discussion of the verb *cure* has shown that the theoretical proposals put forward by Fillmore & Atkins (1992) have been applied straightforwardly to its description in FrameNet. That is, besides the Cure frame, *cure* also evokes a different frame, namely the Preserving frame.

Describing polysemous structures by appealing to semantic frames has two advantages. First, it makes it possible to describe different syntactic frames occurring with the same verb as belonging to the same semantic frame. This treatment is in contrast to WordNet, which provides two distinct senses for the preserving sense of *cure*. Similar to treatments of polysemy in traditional dictionaries, this distinction seems to be motivated by the fact that this sense occurs with two distinct syntactic frames, i.e., an intransitive and a transitive frame (cf. *They cured the apricots* vs. *The apricots cured in the sun*). In FrameNet, both usages are described with respect to the same underlying Preserving frame, thereby adhering to the requirement that "usage differences that need to be reported are best described, not in terms of lexical semantic differences as such, but as differences in the manner of syntactic realization of the elements of their common frame." (Fillmore & Atkins (1992: 101))

The second advantage of treating polysemy by appealing to differences between semantic frames has to do with word sense disambiguation. For example, in cases in which an NLP application needs to determine the sense of a verb, sparse semantic and syntactic information is not sufficient. For the proper sense disambiguation of *cure* in sentences such as *Nancy cured {the ham/the patient}*, it is necessary to know more about the frame semantics of the postverbal NP because it helps to determine whether the Cure or the Preserving frame is evoked. Without this knowledge, it is very difficult to resolve this ambiguity.

[23] WordNet provides some information about the relationships between semantically related words belonging to different parts of speech by including the "derivationally related forms" relation. Note, however, that this function relies primarily on syntactic derivation, and not semantic similarity.

5. Conclusions and Outlook

In this paper, I have compared the different theoretical approaches underlying the WordNet and FrameNet databases in order to show how they have been applied in practice. Whereas the architecture of WordNet relies on many "traditional" conceptual-semantic relations such as synonymy, hyponymy, meronymy, and troponymy, FrameNet is built according to frame semantic concepts. As such, FrameNet represents the practical application of a particular semantic theory (Frame Semantics). Because the two databases were conceived with different goals in mind – WordNet primarily as a semantic database, FrameNet as a resource for computational lexicography – their organizational principles as well as the types of semantic and syntactic information differ from each other. The following table summarizes some of the key differences in the organization of the two databases.

Table 2: Key differences between WordNet and FrameNet

	WordNet	**FrameNet**
Theoretical background	Traditional lexical semantic relations and psycholinguistic principles	Frame Semantics
Organizational units	Words, collocations, multi-word expressions	Lexical units
Independent organizational units larger than words	n.a.	Semantic frames
Semantic relations between words	Synonymy, antonymy, polysemy, hyponymy, hypernymy, troponymy, meronymy, etc.	Polysemy, ability of a lexical unit to evoke the same semantic frame as other lexical units
Analysis of different parts of speech	In terms of different lexical hierarchies and conceptual-semantic relations	With respect to the same semantic frame
Hierarchical relations between organizational units	Multitude of different levels depending on the part of speech (e.g., troponymy, hyponymy)	Frame Inheritance, Subframe Relation, Uses Relation, 'See also' relation
Frequency information	Senses ordered by estimated frequency	n.a.
Treatment of polysemy	Influenced by syntactic properties and traditional lexicographic practice	Based on semantic frames
Syntactic information	Limited number of "sentence frames"	Exhaustive list of lexico-syntactic patterns linked to semantic information
Use of example	Limited number of example	Corpus example(s) for

sentences	sentences	each attested lexico-syntactic pattern

The comparison of the two databases illustrates that each has its strengths in different areas of lexical description. This point brings us back to the three questions posed at the beginning of this paper. With respect to the first question (*How should the lexicon be constructed – by hand or automatically?*), our discussion of the two databases has shown that a hybrid approach appears to be most effective. That is, without carefully planning and examining of the data, lexical idiosyncrasies as well as generalizations might be overlooked. Applying traditional semantic and lexicographic methodology to the construction of WordNet, or, applying frame semantic principles to the design of FrameNet has proven to be effective for constructing these databases. Similarly, without automation, processing large amounts of data (e.g., when searching in large electronic corpora) would present a major bottleneck in the construction of lexical databases.

Finding an answer to the second question (*What kind of information should the lexicon contain?*) is more difficult. As I have shown, WordNet and FrameNet differ both in their theoretical backgrounds as well as in the types of lexical information they aim to provide. The former was conceived as a semantic database, whereas the latter was conceived as a resource for computational lexicography. As such, the two resources vary as summarized in Table 2. An answer to the second question can therefore be found only in the context of applications that require different types of lexical information.

The same holds for finding answers to the third question (*What should the design of the lexicon be – that is, how should its contents be organized and made accessible?*). Comparing the designs of WordNet and FrameNet I have demonstrated that both differ significantly in their organization because of the different theoretical approaches taken towards lexical description. Here, too, an answer can only be found in the context of applications that require different types of semantic and syntactic information. Regarding accessibility requirements, both databases can be downloaded and implemented in other NLP applications, or, they can be accessed via the different web interfaces discussed in this paper. Depending on the scope of lexical information required, a combination of data provided by both WordNet and FrameNet might prove to be most useful. Such a feature is offered by FrameNet's web-based search interface that links to the WordNet database, making it possible to combine the best of both worlds. Clearly, the richness of organizational possibilities holds promise of further improvements in the future design of lexical databases.

References:

Al-Halami, R., and R. Kazman (1998): "Temporal Indexing through Lexical Chaining," in: Fellbaum, C. (ed.), *WordNet: An Electronic Lexical Database,* 333-352. Cambridge, Mass.: MIT Press.

Atkins, B.T.S. (1994): "Analyzing the Verbs of Seeing: A Frame Semantic Approach to Corpus Lexicography," in: Johnson, C. et al. (eds.), *Proceedings of the Twentieth Annual*

Meeting of the Berkeley Linguistics Society, 42-56. Berkeley: Berkeley Linguistics Society.

Atkins, B.T.S. (1995): "The Role of the Example in a Frame Semantics Dictionary," in: Shibatani, M., and S. Thompson (eds*.), Essays in Semantics and Pragmatics – In Honor of Charles J. Fillmore*, 25-42. Amsterdam/Philadelphia: Benjamins.

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

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

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

Burg, J.F.M., and R.P. van de Ried (1998): "COLOR-X: Using Knowledge from WordNet for Conceptual Modeling," in: Fellbaum, C. (ed.), *WordNet: An Electronic Lexical Database,* 333-352. Cambridge, Mass.: MIT Press.

Cruse, A. (1975): "Hyponymy and Lexical Hierarchies," in: *Archivum Linguisticum* 6: 26-31.

Cruse, A. (1977): "The Pragmatics of Lexical Specificity," in: *Journal of Linguistics* 13: 153-164.

Cruse, A. (1986): *Lexical Semantics*. Cambridge: Cambridge University Press.

Dowty, D. (1991): "Thematic Proto-roles and Argument Selection," in: *Language* 67: 547–619.

Fellbaum, C. (1998a): *WordNet: An Electronic Lexical Database*. Cambridge, Mass.: MIT Press.

Fellbaum, C. (1998b): "A Semantic Network of English: The Mother of all WordNets," in: *Computers and the Humanities* 32: 209-220. .

Fellbaum, C. (1998c): "Introduction," in: Fellbaum, C. (ed.), *WordNet: An Electronic Lexical Database*, 1-20. Cambridge, Mass.: MIT Press.

Fellbaum, C. (1998d): "A Semantic Network for English Verbs," in: Fellbaum, C. (ed.), *WordNet: An Electronic Lexical Database*, 333-352. Cambridge, Mass.: MIT Press.

Fillmore, C.J. (1968): "The Case for Case," in: *Universals in Linguistic Theory*, eds. E. Bach and R. Harms, 1–88. New York: Holt, Rinehart, and Winston.

Fillmore, C.J. (1970): “The Grammar of Hitting and Breaking,” in: Jacobs, R.A. and P.S. Rosenbaum (eds.), *Readings in English Transformational Grammar*, 120-133. Ginn and Company.

Fillmore, C.J. (1975): “An Alternative to Checklist Theories of Meaning,” in: Cogen, C. et al. (eds.), *Proceedings of the First Annual Meeting of the Berkeley Linguistics Society*, 123–31. Berkeley: Berkeley Linguistics Society.

Fillmore, C.J. (1976): “Frame Semantics and the Nature of Language,” in: Harnad, S.R. et al. (eds.), *Origins and Evolution of Language and Speech*, 20-32. New York: New York Academy of Sciences.

Fillmore, C. J. (1977a): “Scenes-and-Frames Semantics,” in: Zampolli, A. (ed.), *Linguistics Structures Processing*, 55-81. Dordrecht: North Holland Publishing Company.

Fillmore, C.J. (1977b): “Topics in Lexical Semantics,” in: Cole, R.W. (eds.), *Current Issues in Linguistic Theory*, 76-138. Indiana: Indiana University Press.

Fillmore, C.J. (1978): “On the Organization of Semantic Information in the Lexicon,” in: *Papers from the Parasession on the Lexicon*, eds. D. Frakas, et al., 148-73. Chicago: Chicago Linguistic Society.

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

Fillmore, C. J. (1985a): “Frames and the Semantics of Understanding,” in: *Quadernie di Semantica* 6: 222–254.

Fillmore, C.J. (1985b): “Syntactic Intrusions and the Notion of Grammatical Construction,” in: Niepokuj et al. (eds.), *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society*, 73-86. Berkeley: Berkeley Linguistics Society.

Fillmore, C.J. (1988): “The Mechanisms of Construction Grammar,” in: Axmaker, S. et al. (eds.), *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistic Society*, 35–55. Berkeley: Berkeley Linguistics Society.

Fillmore, C.J. (1994): “The Hard road from Verbs to Nouns,” in: Chen, M.Y., O.J.L. Tzeng, and W.S.Y. Wang (eds.), *In Honor of William S.Y. Wang: Interdisciplinary Studies on Language and Language Change*, 105-129. Pyramid Press.

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

Fillmore, C.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*, 349-393. Oxford: Oxford University Press.

Fillmore, C.J., and B.T.S. Atkins (2000): "Describing Polysemy: The Case of 'Crawl'," in: Ravin, Y., and C. Leacock (eds.), *Polysemy*, 91–110. Oxford: Oxford University Press.

Fillmore, C.J., and P. Kay (1993): *Construction Grammar*. Manuscript. University of California at Berkeley.

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

Fillmore, C.J., M.R.L. Petruck, J. Ruppenhofer, and A. Wright (2003b): "FrameNet in Action: The Case of Attaching," in: *International Journal of Lexicography* 16: 297-333.

Goldberg, A. (1995): *Constructions: a Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press

Gross, D., U. Fischer, and K.J. Miller (1989): "The Organization of Adjectival Meanings," in: *Journal of Memory and Language* 28: 92-106.

Jackendoff, R. (1990): *Semantic Structures*. Cambridge: MIT Press.

Langacker, R. (1990): "The Rule Controversy: A Cognitive Grammar Perspective," in: *Center for Research in Language Newsletter* 4: 4–15.

Leacock, C., and M. Chodorow (1998): "Combining Local Context and WordNet Similarity for Word Sense Identification," in: Fellbaum, C. (ed.), *WordNet: An Electronic Lexical Database*, 265-284. Cambridge, Mass.: MIT Press.

Lowe, J.B., C.F. Baker, and C.J. Fillmore (1997): "A Frame-Semantic Approach to Semantic Annotation," in: *Proceedings of the SIGLEX Workshop,* ed. M. Light. Held April 4-5 in Washington, D.C. in conjunction with ANLP-97.

Lyons, J. (1977): *Semantics*. Cambridge: Cambridge University Press.

McLeod, C. (1998): "COMLEX Syntax," in: *Computers and the Humanities* 31: 459-481.

Miller, G. A. (1990)(ed.): "WordNet: An on-line lexical database." Special issue of *International Journal of Lexicography* 3(4).

Miller, G.A. (1998a): "Foreword by George A. Miller," in: Fellbaum, C. (ed.), *WordNet: An Electronic Lexical Database*, xv-xxii. Cambridge, Mass.: MIT Press.

Miller, G.A. (1998b): "Nouns in WordNet," in: Fellbaum, C. (ed.), *WordNet: An Electronic Lexical Database*, 23-46. Cambridge, Mass.: MIT Press.

Miller, K.J. (1998): "Modifiers in WordNet," in: Fellbaum, C. (ed.), *WordNet: An Electronic Lexical Database*, 47-68. Cambridge, Mass.: MIT Press.

Minsky, (1975): "A Framework for Representing Knowledge," in: *The Psychology of Computer Vision*, ed. P. H. Winston, 211–77. New York: McGraw-Hill.

Petruck, M. (1996): "Frame Semantics," in: Verschueren, J., J-O Östman, J. Blommaert and C. Bulcaen (eds.), *Handbook of Pragmatics*, 1–13. Amsterdam: Benjamins.

Ravin, Y. (1990): *Lexical Semantics without Thematic Roles*. Oxford: Oxford University Press.

Ravin, Y., and C. Leacock (2000): "Polysemy: An Overview," in: *Polysemy*, eds. Y. Ravin and C. Leacock, 1–29. Oxford: Oxford University Press.

Saeed, J.I. (1997): *Semantics*. Oxford: Blackwell Publishers.

Schank, R.C., and R.P. Abelson (1975): *Scripts, Plans, Goals, and Understanding*. Hillsdale: Lawrence Erlbaum.

Tengi, R.I. (1998): "Design and Implementation of the WordNet Lexical Database and Searching Software," in: Fellbaum, C. (1998): *WordNet: An Electronic Lexical Database*, 105-128. Cambridge, Mass.: MIT Press.

Vorhees, E.M. (1998): "Using WordNet for Text Retrieval," in: Fellbaum, C. (eds.) *WordNet: An Electronic Lexical Database*, 285-304. Cambridge, Mass.: MIT Press.

Van Valin, R.D. Jr., and D.P. Wilkins (1996): "The Case for 'Effector': Case Roles, Agents, and Agency Revisited," in: Shibatani, M., and S. Thompson (eds.), *Grammatical Constructions*, 289-322.Oxford: Oxford University Press.

Vossen, P. (1998): "Introduction to EuroWordNet," in: *Computers and Humanities* 32: 73-89.

Appendix

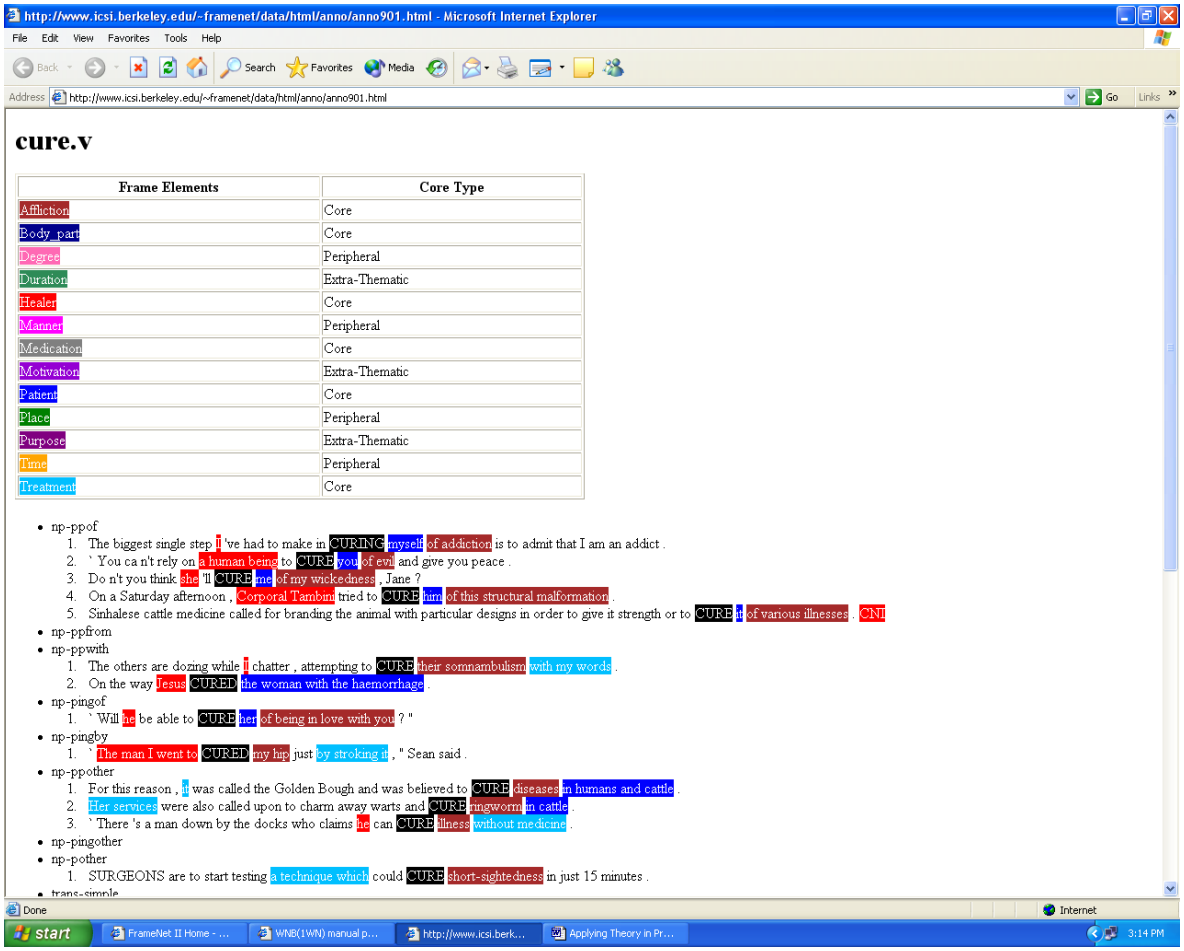


Figure 1: Annotation Report for the verb *to cure* in the CURE frame.

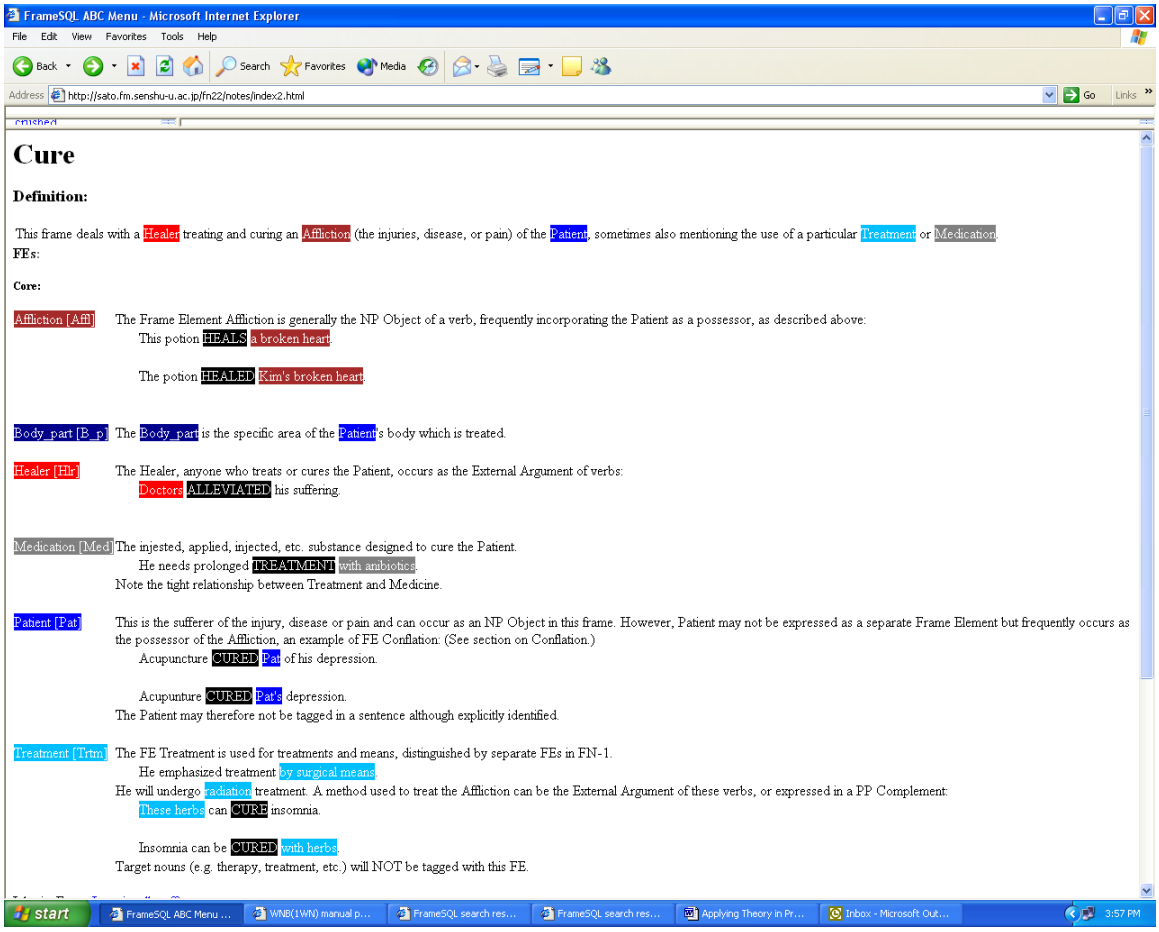


Figure 2: Frame Description of Cure frame

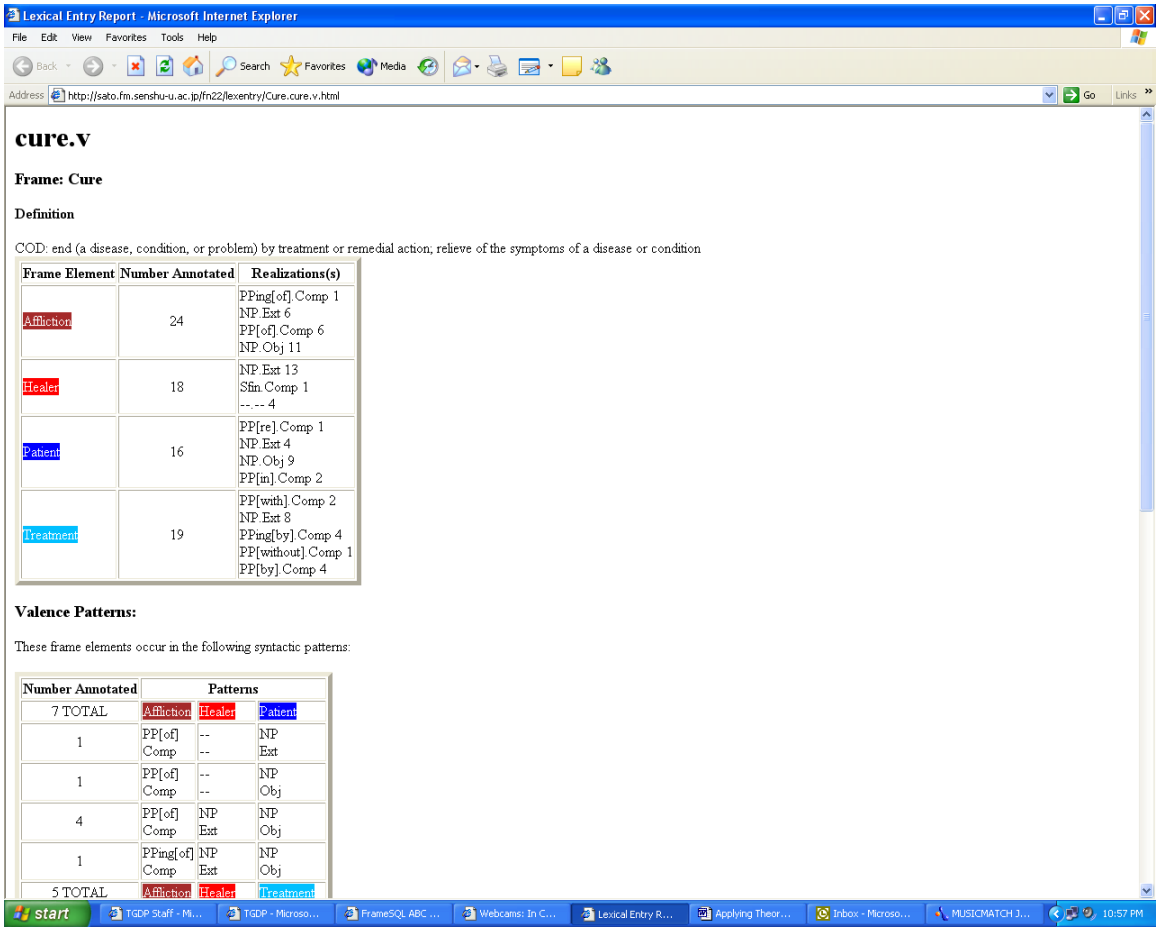
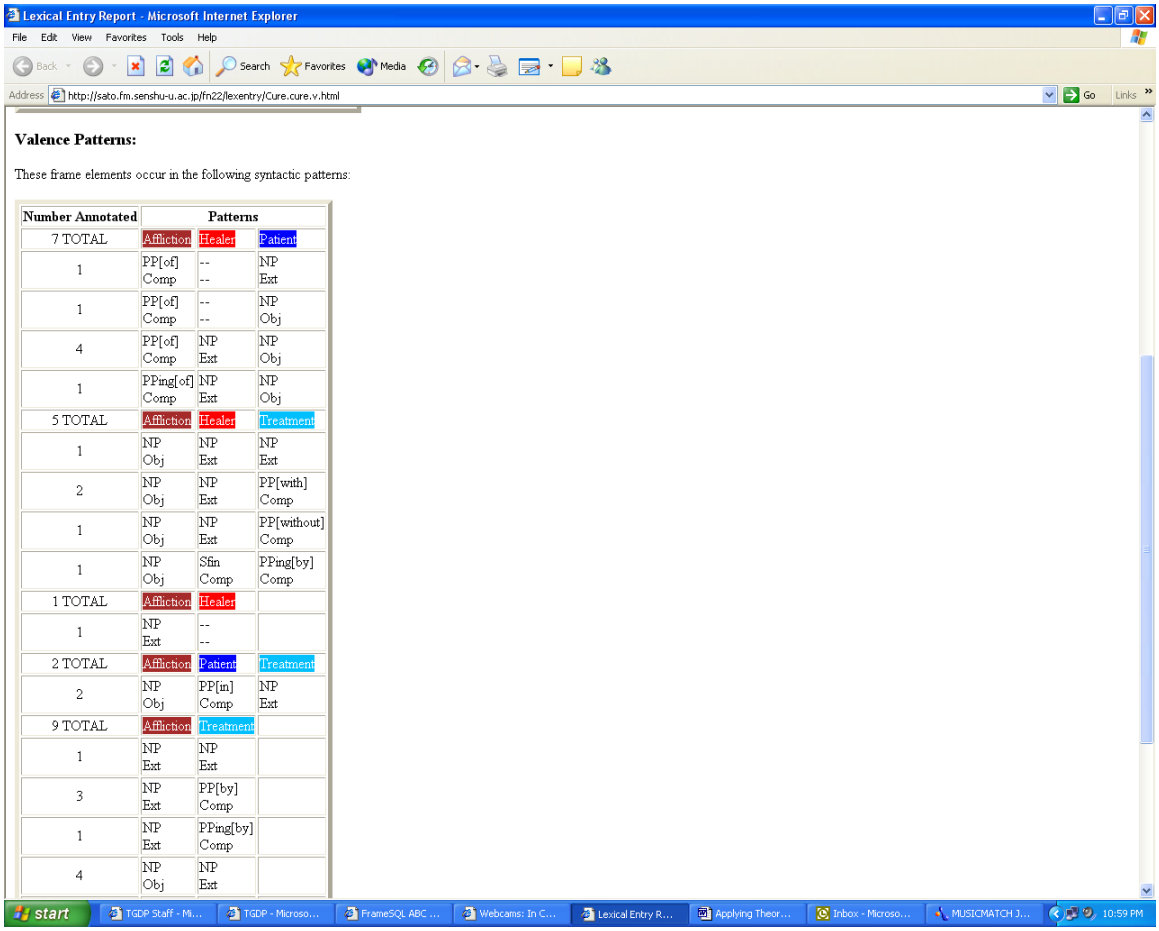


Figure 3: Realization Table for the verb *cure*



Number Annotated	Patterns		
7 TOTAL	Affliction	Healer	Patient
1	PP[of] Comp	-- --	NP Ext
1	PP[of] Comp	-- --	NP Obj
4	PP[of] Comp	NP Ext	NP Obj
1	PPing[of] Comp	NP Ext	NP Obj
5 TOTAL	Affliction	Healer	Treatment
1	NP Obj	NP Ext	NP Ext
2	NP Obj	NP Ext	PP[with] Comp
1	NP Obj	NP Ext	PP[without] Comp
1	NP Obj	Sfin Comp	PPing[by] Comp
1 TOTAL	Affliction	Healer	
1	NP Ext	-- --	
2 TOTAL	Affliction	Patient	Treatment
2	NP Obj	PP[in] Comp	NP Ext
9 TOTAL	Affliction	Treatment	
1	NP Ext	NP Ext	
3	NP Ext	PP[by] Comp	
1	NP Ext	PPing[by] Comp	
4	NP Obj	NP Ext	

Figure 4: Valence Table for the verb *cure*

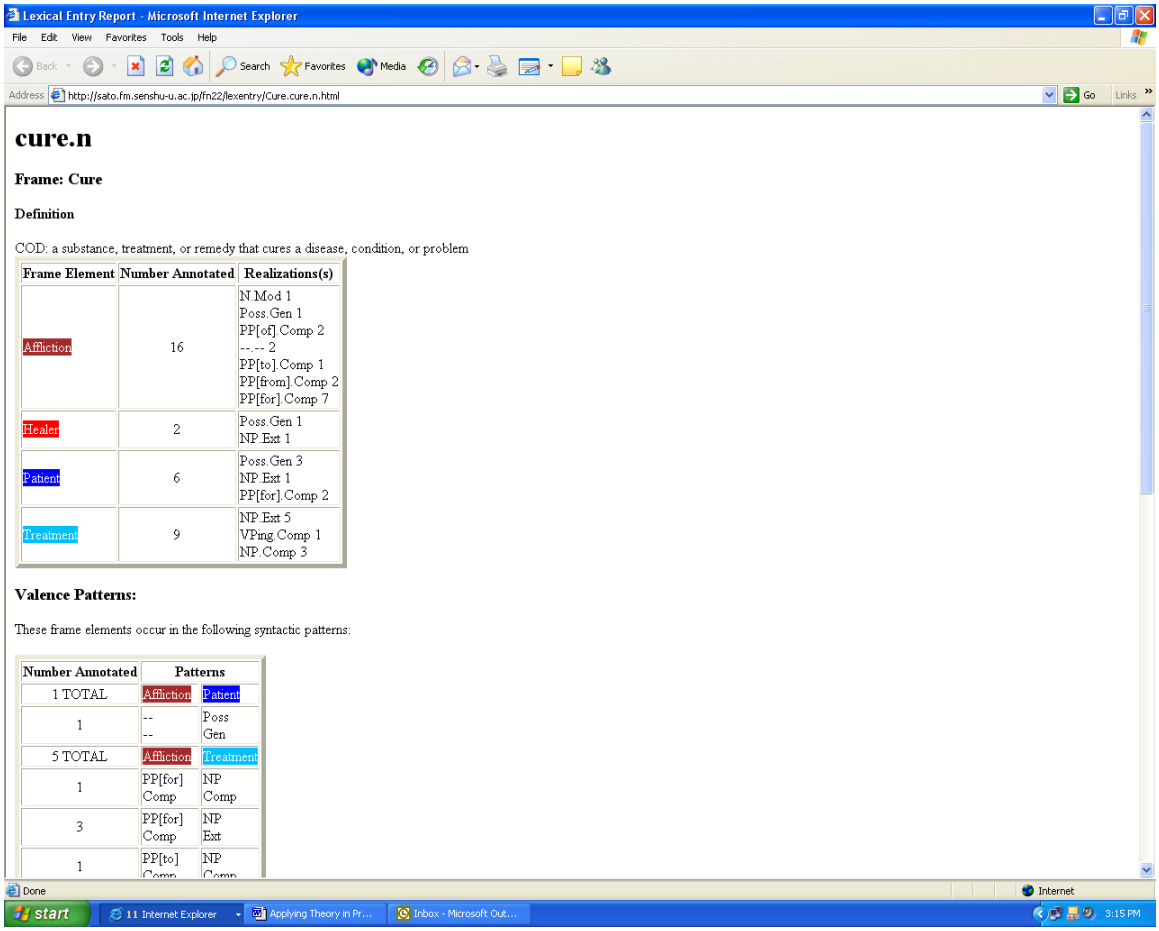


Figure 5: Frame Realization and Valence Tables for the noun *cure*

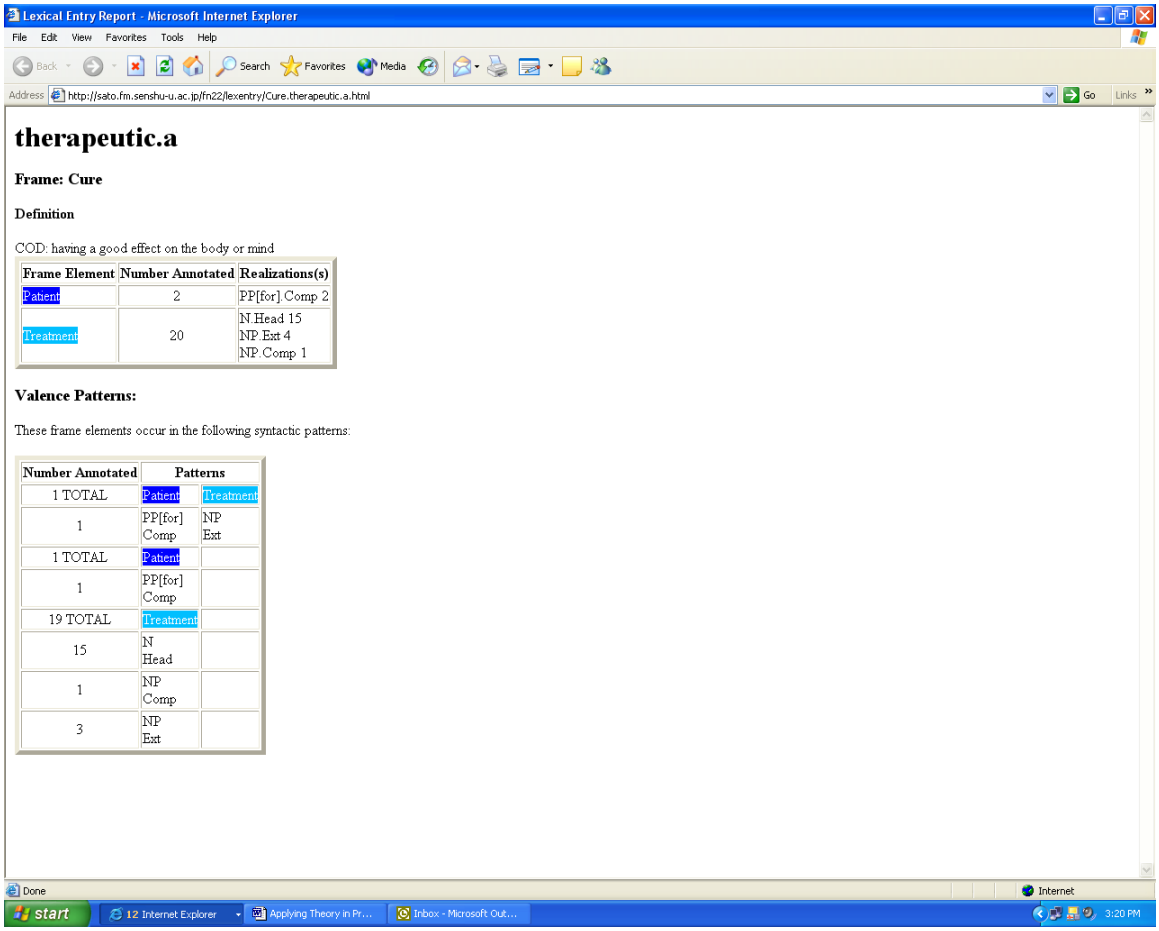


Figure 6: Frame Element Realization and Valence Tables for *therapeutic*