

# Representing Support Verbs in FrameNet

**Miriam R. L. Petruck**

1249 Center St, Suite #600  
Berkeley, CA 94704

miriamp@icsi.berkeley.edu

**Michael Ellsworth**

1249 Center St, Suite #600  
Berkeley, CA 94704

infinity@icsi.berkeley.edu

## Abstract

This paper presents FrameNet’s approach to the representation of Support Verbs, as but one type of multiword expression (MWE) included in the database. In addition to motivating and illustrating FrameNet’s newly consistent annotation practice for Support Verb constructions, the present work advocates adopting a broad view of what constitutes a multiword expression.

## 1 Introduction

Natural Language Processing (NLP) research has been interested in the automatic processing of multiword expressions, with reports on and tasks relating to such efforts presented at workshops and conferences for over ten years (e.g. ACL 2003, LREC 2008, COLING 2010, EACL 2014, NAACL 2015). Overcoming the challenge of automatically processing MWEs remains quite elusive because of the difficulty in recognizing and interpreting such forms. Primarily concerned with the mapping of meaning to form via the theory of Frame Semantics (Fillmore 1985, 2012), FrameNet represents MWEs from the perspective of their semantic heads.

Existing statistical approaches to acquiring MWEs (e.g. Villavicencio et al. 2007, Bannard 2005, Nakov 2013) only offer partial solutions to the problem of MWEs. Many, if not most, such approaches focus on identifying MWEs, and do not address the meaning of the MWEs. In the specific case of noun compounds, Nakov (2013) addressed meaning with a fixed set of relationships between members of the compound or by specifying a more explicit paraphrase (Nakov and Hearst 2013). Other efforts have focused on the meaning of verb particle constructions, by distinguishing between meaning classes of parti-

cles (Cook and Stephenson 2006). Salehi et al. (2015) tested newer methods using word embeddings in English and German for compound nouns and verb particle combinations. These studies focused on predicting MWEs, and have not been assessed for the method’s utility on vector meaning representations for MWEs. In contrast, the FrameNet analysis of MWEs treats all known kinds of multi-word expressions in English and offers a description of their meaning with the same powerful Frame Semantics system that FN uses for single words.

The rest of this paper is structured as follows. Section 2 describes FrameNet briefly; Section 3 provides background to MWEs, also discussing MWEs in FrameNet and specifically support verb constructions. Section 4 presents the terminology that FN uses in its representation of support verbs, and includes an example. Finally, Section 5 offers concluding remarks.

## 2 Background to FrameNet

FrameNet (FN) is a knowledge base with unique information about the mapping of meaning to form in the vocabulary of contemporary English through the theory of Frame Semantics (Fillmore 1985, 2012, Fillmore and Baker 2010).

At the heart of Frame Semantics is the **semantic frame**, i.e. an experience-based schematization of the language user’s world that allows inferences about participants and objects in and across events, situations, and states of affairs. To date, FN has characterized more than 1,200 frames, nearly 13,500 lexical units (LUs), defined as a pairing of a lemma and a frame, and over 202,000 manually annotated sentences that illustrate the use of each.

A FN frame definition includes a description of a prototypical situation, along with a specification of the frame elements (FEs), or semantic roles, that uniquely characterize that situation. FN distinguishes three types of FEs, core, pe-

ripheral, and extrathematic, where core FEs uniquely define a frame. Thus, FrameNet defines the *Revenge*<sup>1</sup> frame as an AVENGER performing a PUNISHMENT on an OFFENDER as a response to an INJURY, inflicted on an INJURED\_PARTY. These five are core FEs in the *Revenge* frame. Peripheral FEs, such as TIME and PLACE, capture aspects of events more generally. Extrathematic FEs situate an event against the backdrop of another state of affairs; conceptually these FEs do not belong to the frame in which they occur. Example (1) shows a FN analysis of verb *avenge.v*; beyond FEs, FN annotates phrase type and grammatical information.<sup>2</sup>

1. [Sam<sub>AVENGER/NP/External</sub>] **avenged** [his brother<sub>INJURED\_PARTY/NP/Object</sub>] [after the incident<sub>TIME/PP/Dependent</sub>].

In (1), *Sam*, the AVENGER, is a NP and functions as the external; *his brother*, the INJURED\_PARTY, is a NP and serves the grammatical function object; *after the incident*, the TIME, is a PP dependent. FN lexical entry reports include a table of valence patterns that displays the automatically summarized results of FE, grammatical function and phrase type annotation, as given in Figure 1.

Avenger	Injured_party	Time	
NP	NP	AVP	
Ext	Obj	Dep	
Avenger	Injury	Offender	Punishment
CNI	NP	DNI	INI
--	Ext	--	--
CNI	NP	INI	PP[by]
--	Ext	--	Dep

Figure 1: Partial Valence Table for *revenge.v*

Note that the red arrow in Figure 1 indicates the valence pattern of Example (1).

The FN hierarchy of frames links frames to each other via a number of frame-to-frame relations. For example, **inheritance** is a relation where for each FE, frame relation, and semantic characteristic in the parent, the same or a more specific analogous entity exists in the child. Thus, to illustrate, *Revenge* inherits *Re-*

<sup>1</sup> The names of FN frames appear in *Courier New* typeface. An underscore appears between each word of a frame name of more than one word; FN only capitalizes the first word of the name.

<sup>2</sup> FN uses *external* for subjects, including of raising Vs, and a limited set of grammatical functions.

*wards\_and\_punishments*, which in turn inherits *Response*, as well as *Intentionally\_affect* as Figure 2 depicts.<sup>3</sup>

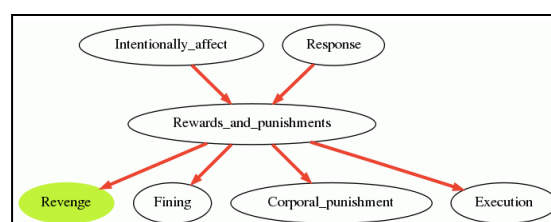


Figure 2: Inheritance Relations in FN

While FN provides frame-specific semantic annotation, its powerful nature becomes most evident when leveraging the larger frame hierarchy, linked through its frame-to-frame relations, of which the local frame is a part.

Not surprisingly, the *Revenge* frame provides the background knowledge structure for defining and understanding a number of MWEs. The following support verb constructions are instances of MWEs defined in terms of *Revenge*: *get even.v*, *get back.v*, *take revenge.v*, and *exact revenge.v*; details appear in Section 3.

## 3 Multiword Expressions

### 3.1 Background

Multiword expressions manifest in a range of linguistic forms (as Sag et al. (2002), among many others, have documented), including: noun + noun compounds (e.g. *fish knife*, *health hazard* etc.); adjective + noun compounds (e.g. *political agenda*, *national interest*, etc.); particle verbs (*shut up*, *take out*, etc.); prepositional verbs (e.g. *look into*, *talk into*, etc.); VP idioms, such as *kick the bucket*, and *pull someone's leg*, along with less obviously idiomatic forms like *answer the door*, *mention someone's name*, etc.; expressions that have their own mini-grammars, such as names with honorifics and terms of address (e.g. *Rabbi Lord Jonathan Sacks*), kinship terms (e.g. *second cousin once removed*), and time expressions (e.g. *August 9, 1929*); support verb constructions (e.g. verbs: *take a bath*, *make a promise*, etc; and prepositions: *in doubt*, *under review*, etc.). Linguists address issues of polysemy, compositionality, idiomaticity, and continuity for each type of MWE mentioned here.

While native speakers use MWEs with ease, their treatment and interpretation in computa-

<sup>3</sup> See, for instance, Petruck and de Melo (2012) for an overview of all of FN's frame-to-frame relations.

tional systems requires considerable effort due to the very issues that concern linguists.

### 3.2 Multiword Expressions in FrameNet

Although not stated explicitly, Fillmore (2006) suggests that linguists and NLP researchers must consider a very broad view of MWEs instead of limiting the scope of their study to those that fit some analytic or classificatory definition.

While FrameNet includes MWEs in its lexicon, it does not analyze any of them internally. For example, given the (noun + noun) MWE *fish bowl*, FN does not offer an analysis of the relationship between *fish* and *bowl*, the two nouns in the compound when that compound is the focus of the analysis. However, FN does provide semantico-syntactic information about the use of MWEs. Consider the sentence *Smedlap bought a large **fish bowl***, where the (bold-faced) target, i.e. the compound noun, evokes the `Containers` frame, with the core FE `CONTAINER` and several peripheral FEs, including `TYPE`. A FN analysis of the sentence indicates that the adjective phrase *a large* is a grammatical dependent of the LU ***fish bowl***, and is annotated with the FE `TYPE`.

2. Smedlap bought [a large `TYPE`]  
[**fish bowl** `CONTAINER`].

In contrast, if the target LU is the head noun of a noun + noun compound, as in *fish bowl*, FN annotates the modifier of that compound with the FE that the modifier instantiates, here `USE`, thus yielding the analysis in (3). Note that both *bowl* and *fish bowl* evoke `Containers`, with analysis of each employing the same set of FEs.

3. Smedlap bought a [*fish* `USE`]  
[**bowl** `CONTAINER`].

The different analyses in (2) and (3) are a result of the different target LUs in each example.

Table 1, below, lists the types of MWEs found in FrameNet, and provides examples of each.<sup>4</sup>

	Examples
<b>MWE Type</b>	
<b>Compound Noun</b>	fish bowl
<b>Verb + Particle</b>	take off bring out put on
<b>Support Constructions</b>	
<b>Support Verb</b>	make a decision say a prayer find solace
<b>Support Preposition</b>	in possession under attack at large

Table 1: MWE Types in FrameNet

### 3.3 Support Verbs in FrameNet

This section briefly describes support verbs, very broadly defined (e.g. *give advice*, *find solace*, *make a decision*), including plain support verbs, as well as Melcuk's (1996) lexical functions (e.g., causatives), and the discrepancy between the syntactic heads and semantic heads of such forms. Since FN has included

Both Meaning Text Theory (MTT) and Frame Semantics (FS) are interested in characterizing the lexical structure of support verb constructions (as in Table 1), despite the different approaches of each theory. In MTT, lexical functions describe collocations that serve a range of purposes, including, for instance, `MAGN`, for collocations that emphasize the extremeness of another word (e.g. *red* hot) and `CAUS` for collocations that express the causation of a word (e.g. *give* a heart attack). Both theories want to describe (a) the verb and the nominal syntactic head of the verb's dependent; (b) the way that the situation or frame that the noun evokes receives expression in the support construction; and (c) how the syntactic dependents of the verb match the semantic roles in the frame that the noun evokes. Some of the shared goals for analyzing support verb constructions motivated exploring the possibility of aligning them (Bouveret and Fillmore 2008), but numerous practical matters, such as different sets of terminology and methodology, precluded any such alignment.

Still, a brief overview of the key differences in the two approaches will illuminate the flexibility of the FrameNet approach. MTT models a limited set of syntactic and semantic relationships between parts of a MWE. Though MTT allows for some multiword expressions involving syn-

<sup>4</sup> FrameNet also includes support nouns and support adjectives, which the authors believe to be of limited interest to the larger computational linguistics audience.

tactic and semantic relations beyond these relationships, they do not form part of the larger system. In contrast FrameNet handles all types of meaning relations through its use of frames. The two approaches are complementary in that FN does not model the syntactic relation between the parts of MWEs in a general way, other than the annotation of the syntactic head of the MWE and its part-of-speech.

The support verb construction considered here is but one linguistic form that shows the discrepancy between a syntactic and a semantic head. For example, consider *bottle of champagne*, where *bottle* may refer to a measure (e.g. *They drank a bottle of champagne to celebrate*), or it may indicate a container (e.g. *He broke the bottle of champagne over the newly painted boat*). Regardless of linguistic form, such discrepancies present a challenge to NLP, specifically natural language understanding (NLU). NLU systems must know that *breaking a bottle* is possible, but *breaking champagne* is not. Thus, success in NLP depends, in part, on systems that include the means to resolve the discrepancy between syntactic and semantic heads.

#### 4 Representing Support Vs in FrameNet

This section motivates FrameNet’s approach to the representation of support verbs, introduces the terminology that FN uses in their representation, illustrating each, and providing an example that shows the advantage of exploiting FN information for these constructions.

##### 4.1 Motivating FrameNet’s Approach

FrameNet began as a lexicography project, and to a large extent remains such, with more attention to the needs of NLP recently than in early phases of the project. As such, FN considered lexicographic factors to determine its approach to representing support verb construction. Nevertheless, FN views the features that it uses in its annotation as showing promise for NLP.

##### 4.2 Terminology

Table 2 displays all possible combinations of the three features that characterize different types of lexically relevant governors, be they supports (as defined in FN), or not. What follows first is a list of features that characterize the relationship between governing and governed words in general: specifically, we define *Bleached*, *FE-supplying*, and *Idiosyncratic*. Then, this section provides a description of the labels that FrameNet uses for

particular combinations of these features, i.e. *Support*, *Copula*, *Controller*, and *Governor*. In the examples that follow, underlining identifies the dependent word with annotation to discuss.

	Bleached		Non-Bleached	
	+FE	-FE	+FE	-FE
+ Idio.	Supp		Supp	
- Idio.	Cop		Ctrlr	Gov

**Table 2: Terminology for Lexically Relevant Governors**

- **Bleached:** Bleached indicates that the governor does not contribute significant content semantics to the combination of governor and governed word (e.g. *she took revenge, there was rain*). In Non-Bleached cases, added frame annotation models the added meaning from the governor.
- **FE:** FE-supplying (or not) indicates that syntactic dependents of the governing word fill semantic roles of the governed word (e.g. *they gave me a shock*).
- **Idio:** Idiosyncratic covers lexical material whose combination is not predictable from the general meaning of the individual words (e.g. *the US lifted the sanctions*).

These three features underlie the annotation labels that FN employs:

- **Cop:** Copula is for annotating BE, and copula-like verbs (e.g. *seem happy, appear smart*).
- **Ctrlr:** Controller identifies the verb whose subject or object is also the subject or object in the dependent clause (e.g. *attempt a rescue*).
- **Gov:** Governor identifies a word that is used in a prototypical way with a dependent, but without any unusual meaning or any supplying of an FE to its dependent (e.g. *wear boots*).
- **Supp:** Support identifies words that would not mean the same thing without their syntactic dependent (e.g. *take a bath*).

In Table 2, above, the highlighted cell indicates the situation where FrameNet annotates the support item (here, a verb or a preposition) as a separate target, and the combination of Supp + Target is not quite equivalent semantically to the Target alone. (See the example (5).) Regular supports (*exact in exact revenge*) need no further analysis and FN does not annotate them further.

### 4.3 Example

Consider example (4), where the analysis focuses on the support verb expression *took a dirt nap*.<sup>5</sup>

4. Horatio<sub>PROTAGONIST</sub> [*took*<sup>Supp</sup> a **dirt nap**].

FN characterizes *dirt nap*, the target of analysis, in terms of the *Dead\_or\_alive* frame, defined as a situation in which a *PROTAGONIST* is in the dynamic, maintained state of being alive or has exited that state. FN records *Horatio*, the syntactic subject of the verb *took* as the *PROTAGONIST*, and marks *took* with the label *Supp*. By characterizing *took a dirt nap* in terms of its semantic head, *dirt nap*, FN provides needed information about the participants in the event that the support verb expression describes. Independent of the task, e.g. translation, summarization, search, etc., any NLP system must know that *Horatio* is the participant who is dead. FrameNet provides that information.

Characterizing MWEs for identification and representation in NLP requires systematizing the kinds of combinations that exist. FN provides an elaborate classification system that informs downstream tasks whether the syntactic head or a syntactic dependent is the most important part of a MWE semantically. Importantly, FN provides a unified way to represent the meaning of all types of combinations. This approach includes partially compositional cases, as in (5), where the curly brackets identify the support verb construction.

5. Officials {*lifted*<sup>Supp</sup> [Oslo's<sub>EVALUEE</sub>]  
**penalty**}.

In example (5), *Oslo's* fills the *EVALUEE* role of *Rewards\_and\_Punishments*, which the noun *penalty* evokes. Additionally, the support verb *lift* evokes another frame, i.e. the *Cause\_to\_end* frame, with two core FEs, *AGENT* and *PROCESS*. In this second frame, the noun *officials* fills the *AGENT* role and the NP *Oslo's penalty* fills the *PROCESS* role, shown in example (6), below.

6. [Officials<sub>AGENT</sub>] {**lifted**<sup>Supp</sup> [Oslo's  
*penalty*<sub>PROCESS</sub>]}

---

<sup>5</sup> Given the goal of this work, while recognizing the metaphor, we provide no analysis for *dirt nap*, or further information about its being a compound noun.

Also, in the definitions of LUs that only evoke the frame with certain dependents, e.g. *lift.v* here, FN records the semantic type *Support\_only\_LU*. At present, no automatic NLP method captures the complexity of information that FN characterizes. As such, in conjunction with automated semantic parsing (Das et al. 2014), FN holds great potential for use in NLP tasks that depend on processing support verb constructions, as one type of MWE.

## 5 Conclusion

This paper has provided a brief overview of multiword expressions in FrameNet focusing on one type of such expression, namely support verb constructions. In addition, the present work has achieved its goals of motivating, presenting, and illustrating FrameNet's current policy and newly consistent practice of representing support verb construction. Importantly, the paper also shows that FrameNet offers crucial information about the **meaning** of support verb constructions. Statistical approaches, which tend to focus on the identification of MWEs in text, do not provide such information.

## Acknowledgments

FrameNet's treatment of multiword expressions, including support verbs, draws significantly upon Fillmore (2006).

The authors are grateful to the FrameNet team for the numerous discussions about the ideas and annotation conventions presented in this work.

## References

- C. Bannard. 2005. Learning about the meaning of verb particle constructions from corpora. *Computer Speech and Language Special Issue on MWEs* 19(4): 467–478.
- M. Bouveret and C. J. Fillmore. 2008. Matching Verbo-nominal Constructions in FrameNet with Lexical Functions in MTT. In E. Bernal and J. De-Cesaris (Eds.). *Proceedings of the 13th EURALEX International Congress*, pp. 297–308.
- P. Cook and S. Stevenson. 2006. Classifying Particle Semantics in English Verb-Particle Constructions. In *Proceedings of the Workshop on Multiword Expressions: Identifying and Exploiting Underlying Properties*. ACL: Sydney, 45–53.
- D. Das, D. Chen, A. F. T. Martins, N. Schneider, and N. A. Smith. 2014. Frame-semantic parsing. *Computational Linguistics* 40.1: 9–56

- C. J. Fillmore. 1985. Frames and the semantics of understanding. *Quaderni di Semantica*, 6.2: 222-254.
- C. J. Fillmore. 2006. Multiword Expressions: An Extremist Approach, Unpublished PPT. ICSI and University of California, Berkeley, CA.
- C. J. Fillmore. 2012. Encounters with Language. *Computational Linguistics*. 38.4: 701-718.
- C. J. Fillmore and C. Baker. 2010. A Frames Approach to Semantic Analysis. In B. Heine and H. Narrog (eds.), *The Oxford Handbook of Linguistic Analysis*. Oxford: OUP, pp. 791-816.
- I. A. Mel'čuk, I. 1996. Lexical Functions: A Tool for the Description of Lexical Relations in a Lexicon. In L. Wanner (ed.), *Lexical functions in lexicography and natural language processing*, Amsterdam: John Benjamins, pp. 37-102.
- P. Nakov. 2013. On the interpretation of noun compounds: Syntax, semantics, and entailment, *Natural Language Engineering* 19.3: 291-330.
- P. Nakov and M. A. Hearst. 2013. Semantic Interpretation of Noun Compounds Using Verbal and Other Paraphrases, *ACM Transactions on Speech and Language Processing*, special issue on Multiword Expressions, 10.3, Article 13.
- M. R. L. Petruck and G. de Melo. 2012. Precedes: A Semantic Relation in FrameNet, *Proceedings of the Workshop on Language Resources for Public Security Applications at the 8<sup>th</sup> Conference on International Language Resources and Evaluation (LREC)*, Istanbul, Turkey, pp. 45-49.
- J. Ruppenhofer, M. Ellsworth, M. R. L. Petruck, C. R. Johnson, and J. Scheffczyk. 2010. *FrameNet II: Extended Theory and Practice*. Web Publication: [framenet.icsi.berkeley.edu/book](http://framenet.icsi.berkeley.edu/book).
- I. A. Sag, T. Baldwin, F. Bond, A. Copestake, and D. Flickinger. 2002. Multiword expressions: A pain in the neck for NLP. In *Proceedings of the 3rd International Conference on Intelligent Text Processing and Computational Linguistics (CICLing-2002)*, Berlin: Springer, pp.1-15.
- B. Salehi, P. Cook, and T. Baldwin. 2015. A Word Embedding Approach to Predicting the Compositionality of Multiword Expressions. In *Proceedings of the Annual Conference of the North American Association for Computational Linguistics*, Denver, pp. 977-983.
- A. Villavicencio, V. Kordoni, Y. Zhang, M. Idiart, and C. Ramisch. 2007. Validation and evaluation of automatically acquired multiword expressions for grammar engineering. In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pp. 1034-1043, Prague: ACL.