

FrameNet’s *Using* Relation as a Source of Concept-based Paraphrases

Jen Sikos and Sebastian Padó

April 24, 2018

Abstract Paraphrases offer different ways of describing the same state of affairs, but formally characterizing paraphrases has proved a challenging task. Hasegawa et al. (2011) pointed out the usefulness of FrameNet for paraphrase research, focussing on paraphrases which are backed by underlying classical linguistic relationships such as synonymy or voice alternations. This article proposes that other frame-to-frame-relations, notably ‘Using’, can serve as a source for concept-based paraphrases – that is, paraphrases that are backed by common sense knowledge, as in “he called him a hero” – “he praised him for being a hero”. While the predicates in these sentences are not synonymous, we would argue that the sentences are paraphrases – albeit of a kind that involves world knowledge about the relationship between different event classes. We make two main contributions: First, we propose a shallow taxonomy for the frame pairs that instantiate ‘Using’ that is motivated by their ability to form concept-based paraphrases. Second, we zoom in on the subclass of ‘Using’ instances which supports concept-based paraphrasing and provide a formalization of some prominent types of side conditions that are necessary to produce felicitous paraphrases.

1 Introduction

Paraphrases are alternative ways of describing the same state of affairs (Leech, 1974), that is, they are sentence-level or phrase-level synonyms. For language users, paraphrases are a fundamental instrument to vary the form of the content they want to express. For linguists, checking the felicity of certain paraphrases is a widely used diagnostic. An example is the influential work by Levin (1993) which defines verb classes in terms of the diathesis alternations that they support, such as the dative alternation (*give the policeman a book – give a book to the policeman*).

Explaining paraphrase relations between sentences was recognized early on as a fundamental goal of semantics (Fodor, 1961; Katz and Fodor, 1963; Leech, 1974). It also forms an integral part of the program of Fillmore’s “semantics of understanding” (Fillmore, 1985). Beyond surface-level alternations, Fillmore considers paraphrases based on different predicates, such as the following example:

- (1) a. [He_{WINNER}] **defeated** [the former champion_{LOSER}].
b. [The former champion_{LOSER}] was **beaten** by [him_{WINNER}].

Fillmore explains the paraphrase relationship between (1a) and (1b) by asserting that both predicates (*frame-evoking elements*) make reference to the same scenario or scene (*frame*), namely BEAT_OPPONENT. This frame evokes a situation in which there is a competition where someone lost and someone was a winner. In combination with knowledge about the grammatical realization of semantic roles (*frame elements*), this allows the reader infer that the two sentences express the same meaning.

Frame Semantics is in the fairly unique position of associating predicates with both linguistic and conceptual information. The linguistic information consists of the frames as predicate classes, the sets of semantic roles associated with them, and the roles’ syntactic realization patterns. The conceptual information consists of the characterization of situations and their participants, as well as the relationships that are encoded between frames. The theory of Frame

Semantics has shown to be feasible for large-scale, lexical-semantic analysis by the construction of FrameNet¹ (Fillmore et al., 2003). FrameNet is a frame-semantic dictionary that currently covers more than 1,000 frames with 13,639 predicates. This makes it an attractive resource for paraphrase research in general (Hasegawa et al., 2011) and has been used in this capacity in the natural language processing community (Ellsworth and Janin, 2007; Coyne and Rambow, 2009; Aharon et al., 2010).

There has been considerable debate on what constitutes a paraphrase, in particular with respect to the amount of conceptual knowledge that may be involved – see Schreyer (1978) for a general discussion and Chang et al. (2002) and Narayanan (2014) for FrameNet-specific discussions. Hasegawa et al. (2011) take a relatively conservative approach and restrict themselves mostly to what they call *language-based* paraphrases. These include any paraphrases that involve linguistic variation, such as antonymy (*we continued doing it – we didn't stop doing it*), support verb constructions (*they discussed it – they had a long discussion about it*), and voice alternations (*the management rewarded Susan – Susan was rewarded by management*). In contrast, this article considers a wider range of *concept-based paraphrases* that cannot be readily reduced to a linguistic relation but instead involve common sense knowledge. The concept-based paraphrases we find apply to full predicate-argument structures, where the reader must draw upon world knowledge for their interpretation. These types of paraphrases have been found in large numbers in parallel and comparable documents, both multilingually and monolingually (Baker, 1992; Padó, 2007), but have not received much attention in the linguistic literature as they are difficult to define.

Consider the following example:

- (2) a. [He SPEAKER] **called** [him ENTITY] [a hero LABEL].
b. [He COMMUNICATOR] **praised** [him EVALUÉE] [for being a hero REASON].

The predicates in these sentences are not strictly synonymous. While “call”

¹<https://framenet.icsi.berkeley.edu>

refers to a communication event and evokes the LABELING frame in FrameNet, “praise” necessarily expresses a value judgment and is therefore assigned to the JUDGMENT_COMMUNICATION frame. Nevertheless, we would argue that (2a) and (2b) are in fact paraphrases – albeit a kind of paraphrase that involves world knowledge about the relationships between labeling and judging events. We refer to these cases as concept-based. The structure of FrameNet already furnishes most of the knowledge necessary to begin describing these relationships, which makes it possible to capture a wider range of paraphrases than previously considered.

A deeper formalization of concept-based paraphrases could be conceived via the (embodied) schemas from Embodied Construction Grammar (ECG, Bergen and Chang (2005)). Schemas provide an interface between language and mental simulations of the world and therefore describe paraphrastic relationships via reference to the same underlying schema. However, in the absence of a comprehensive resource of schemas, we provide a more shallow account of paraphrases at the FrameNet level.

More specifically, we concentrate on FrameNet’s frame-to-frame relations (Fillmore et al., 2004) which define a set of relationships between frames. These differ in their status regarding the linguistic-conceptual divide (see Section 2 for details). This article considers the most frequently applied, conceptually motivated frame-to-frame relation, **Using**. **Using** links a more concrete frame to a more abstract but conceptually related frame (see Section 2.1 for details). For instance, **Using** relates the frames LABELING and JUDGMENT_COMMUNICATION in Example (2), and it is a promising candidate in general for explaining a range of conceptually motivated paraphrases. As we describe in further detail below, **Using** is, arguably by necessity, a somewhat imprecise relation. In this work, we ask to what extent do the frame pairs related by **Using** lead to felicitous paraphrases, and we find that paraphrasing is often only possible when certain semantic constraints (*side conditions*) are fulfilled. This includes Example (2a), where the LABEL that is predicated in the LABELING frame must carry a positive or negative sentiment in order to be amenable for paraphrasing with

the JUDGMENT_COMMUNICATION frame in (2b). This information is currently not represented in the **Using** relation.

While the central focus of this work is **Using** as a source of concept-based paraphrases, it is also worth noting that some conceptual paraphrases can be produced within a single frame. One such example is FrameNet’s *incorporated frame element*, in which a role can be expressed externally to the predicate or as part of the semantics of the frame-evoking element itself (*she spread butter on the toast – she buttered the toast*). While these within-frame paraphrases also build on a certain amount of world knowledge, the patterns that form them are often lexically specific – e.g., the paraphrase above only works for the predicate *to butter* (Jackendoff, 1990). Instead, we focus on frame-to-frame relations which we expect will provide a more general source of concept-based paraphrases.

The two main contributions of this paper address the range of relations that can be found within the **Using** frame-to-frame relation. First, we present a comprehensive analysis of the frame pairs that instantiate **Using** in FrameNet 1.5² motivated by their suitability for paraphrasing. This analysis suggests a coarse-grained taxonomy of the frames that instantiate **Using**, presented in Section 3, which introduces four subclasses of the **Using** relation. Our second contribution is the selection of one subclass which supports concept-based paraphrasing and a formalization of some prominent types of side conditions associated with this subclass, presented in Section 4.

2 Frame-to-frame relations in FrameNet

Frame-to-frame relations connect a single frame to another and are FrameNet’s prime mechanism to account for conceptual relationships beyond individual frames. The most current version of FrameNet has 14 frame-to-frame relations, including **Inheritance**, **Using**, **Perspective_on**, **Subframe**, **Precedes**, **Inchoative_of**, **Causative_of**, and **Metaphor**.

²The appendix notes the frames that have changed from FrameNet v1.5 to the latest version, FrameNet v1.7.

As mentioned above, some frame-to-frame relations are based on linguistic and/or ontologically-based properties while others are more conceptual in nature. The paraphrasing study by Hasegawa et al. (2011) focuses on the former. **Inheritance** belongs to this first type of frame-to-frame relations, as it is typically represented as “is-a” in ontological frameworks (Scheffczyk et al., 2006; Chow and Webster, 2007). In **Inheritance**, the child frame gives further specification to the parent frame, but core semantic roles are mappable from the parent to the child. Paraphrasing with the **Inheritance** relation can often be a case of hyponym/hypernym swapping, as Example (3) which uses the COMMERCE_BUY frame (3a) that inherits from the GETTING frame (3b).

- (3) a. [Myeloski_{BUYER}] had insisted on **buying** [Duncan_{RECIPIENT}] [a pizza_{GOODS}] at Pizza Hut.
 b. [Myeloski_{SOURCE}] had insisted on **getting** [Duncan_{RECIPIENT}] [a pizza_{THEME}] at Pizza Hut.

Other frame-to-frame relations in this group reflect well-studied linguistic phenomena, such as the **Causative_of/Inchoative_of** relations (Petrucci et al., 2004). Frame pairs in the **Causative_of** relation can produce paraphrases where one frame expresses a CAUSE and the other is a stative event, such as the CAUSE_TO_FRAGMENT frame which is linked via **Causative_of** to the BREAKING_APART frame. The CAUSE_TO_FRAGMENT frame in (4a) takes an AGENT as part of its core conceptual structure while the BREAKING_APART frame in (4b) does not and instead needs to express the AGENT as an oblique.

- (4) a. [Mark_{AGENT}] **broke** [the windscreen_{WHOLE_PATIENT}] [into pieces_{PIECES}].
 b. [The windscreen_{WHOLE}] **broke** [into pieces_{PIECES}] because of Mark.

Among the relations considered by Hasegawa et al. (2011), the least clearly linguistic one is **Perspective_on**, which connects frames that present different perspectives of the same underlying situation. For example, GET_A_JOB and HIRING are both connected to EMPLOYMENT_START via this relation, giving rise the following paraphrase:

- (5) a. [I_{EMPLOYEE}] **signed on** [with YouTube_{EMPLOYER}] [to make them a new GUI_{TASK}].
- b. [YouTube_{EMPLOYER}] **hired** [me_{EMPLOYEE}] [to make them a new GUI_{TASK}].

The relationship between “sign on” and “hire” does not correspond to any classical lexical-semantic or ontological relation, but there is arguably a strong mutual *inference relation* between them, in the sense spelled out by Norvig (1987):

[An inference is] any assertion which the reader comes to believe to be true as a result of reading the text, but which was not previously believed by the reader, and was not explicitly stated in the text.

In fact, inference, in the absence of linguistic and ontological explanation, appears to form the basis for the concept-based frame-to-frame relations in FrameNet. One prominent relation in this concept-based group of frame-to-frame relations is the **Subframe** relation which models *script* or *schema* knowledge (Schank and Abelson, 1977; Abelson, 1981):

- (6) a. [They_{SUSPECT}] were **arrested** [for robbery_{CHARGES}].
- b. [They_{DEFENDANT}] were put on **trial** [for robbery_{CHARGES}].

In this example, the individual sentences activate specific subframes of the general CRIMINAL_PROCESS schema which binds together a sequence of events.

In NLP, recent work has sought to automatically predict frame-to-frame relations in text with modest success (Botschen et al., 2017), and such relations have also been used to generate paraphrases automatically (Coyne and Rambow, 2009). In this study, the authors targeted the **Perspective_on** relation from Example (5). Predicates were taken from a source frame along with the syntactic valence patterns associated with those predicates. The same procedure was applied to the target frame, which is expected to have different syntactic patterns associated with its predicates. The syntactic patterns from the source and target predicates are then treated as the rules that authors used to produce a paraphrase

from the source frame to the target frame. As we describe in Section 3, we adopt a similar approach and manually generate paraphrases across **Using** frames by paraphrasing predicates.

2.1 The Using relation

Despite being one of the largest frame-to-frame relations in FrameNet, **Using** has thus far not been taken up as a serious object of study in paraphrase research. Somewhat surprisingly, there are even two definitions of **Using** that circulate. The definition in the FrameNet documentation is given as (Ruppenhofer et al., 2010, p. 83):

... a particular frame [that] makes reference in a very general kind of way to the structure of a more abstract, schematic frame [...]

Another study, Petruck and de Melo (2012), defines **Using** as a

... relationship between a child frame and parent frame in which only some of the [frame elements] in the parent have a corresponding entity in the child.

These two definitions are not contradictory, but they are not equivalent either. The first focuses on the difference in specificity between two frames without prescribing any specific relation between their respective frame elements. The second concentrates on the presence of a partial mapping between the frame elements without imposing specific constraints on the two frames' relative specificity.

Thus, it is not surprising that the frame pairs in this relation form a somewhat inhomogeneous group. Some of them appear to be a better fit with Petruck and De Melo's definition. An example is the JUDGMENT_COMMUNICATION – LABELING frame pair from Example (2), revisited in Figure 1; the LABELING frame includes predicates such as “call”, “brand”, and “term”, and refers to a speaker using a label to characterize an entity. In the second sentence, “praise” evokes JUDGMENT_COMMUNICATION, a frame in which a

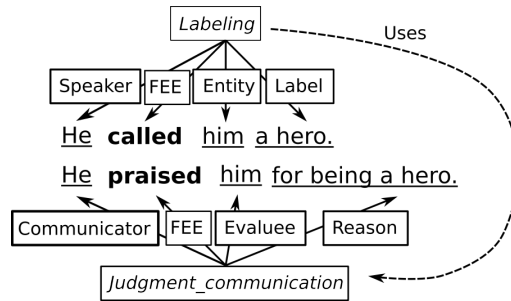


Figure 1: Visualization of the paraphrase relation in Example (2).

communicator expresses an opinion about another person or phenomena. JUDGMENT_COMMUNICATION includes predicates such as “condemn” and “praise,” where each conveys a sentiment that can be either positive or negative. These LABELING and JUDGMENT_COMMUNICATION frames differ in their semantic roles: JUDGMENT_COMMUNICATION includes a frame element ADDRESSEE that is absent from LABELING. Conversely, the LABEL frame element of LABELING is optional in JUDGMENT_COMMUNICATION. Therefore, there is no proper mapping across rolesets, and no **Inheritance** relation can be established between the two frames. At the same time, there is a considerable intersection between frames in terms of the states of affairs that can be verbalized within either frame. Thus, FrameNet falls back to **Using** (cf. Figure 1) to capture this relationship.

Other pairs in **Using** correspond very well to the definition provided by Ruppenhofer et al. (2010). An example is TRANSLATING – MENTAL_ACTIVITY, where MENTAL_ACTIVITY is an abstract, non-lexicalized frame covering all kinds of “activity of the mind” by a sentient entity.

One way to understand the relationship between the two definitions is as a reflection of FrameNet as a growing resource; in an ideal FrameNet lexicon, shared conceptual structures between two concrete frames would be factored out into an abstract frame that both concrete frames use. Since this is infeasible in practice, **Using** relations are a pragmatic device to indicate partially-shared

conceptual structures without having to specify a new, abstract frame.³

For our purposes, we need to ask to what extent a **Using** relation between frame pairs is sufficient to give rise to valid paraphrases. Ruppenhofer’s definition leaves open the potential problem that the general frame is too abstract to express the meaning of the more specific frame, while Petruck and de Melo’s definition may also cover frame pairs that are too dissimilar to paraphrase one another. In what follows, we investigate this question by classifying **Using** frame pairs (Section 3) and discussing side conditions for concept-based paraphrases (Section 4).

3 Analyzing and Classifying Instances of Using

To better understand the heterogeneity among the frame pairs in **Using**, we created a coarse-grained subclassification of the relation. Our overarching goal was to assess the usefulness of FrameNet’s frame-to-frame relations for concept-based paraphrasing. Thus, if Frame 1 uses Frame 2, the guiding principle of our classification was to decide whether Frame 2 could then be used to paraphrase Frame 1. We developed our classification in a bottom-up, data-driven manner by compiling all 490 frame pairs listed for **Using**. The full classification can be downloaded in text and PDF format from <http://www.ims.uni-stuttgart.de/forschung/ressourcen/lexika/FN-using.html>.

For each frame pair where Frame 1 uses Frame 2, we performed the following steps:

1. Sample a predicate (frame-evoking element) p from the frame-evoking elements (FEE) in Frame 1.
2. Sample a sentence s from the example sentences for p provided by FrameNet, which are mostly drawn from the British National Corpus (BNC)⁴.

³“We have defined this relation mainly because working out full inheritance lattices among frames has proven to be extremely contentious and time-consuming. The Uses [sic] relation allows us to define a similar relation without getting bogged down in details.” (Baker et al., 2003)

⁴<http://www.natcorp.ox.ac.uk/>

3. Test whether s can be paraphrased with Frame 2 by manually generating paraphrases. If this is possible without introducing additional frames, we call this a *minimal paraphrase*. Nearly all of the examples in Sections 1 and 2 are minimal paraphrases, with one possible exception⁵.
 - If there are minimal paraphrases, record semantic properties that must be met to make the paraphrase felicitous, if there are any (“side conditions”)
 - If there are no minimal paraphrases, but paraphrasing is nonetheless possible by introducing additional frames, record additional frames that can be used to produce a felicitous paraphrase.
 - If no paraphrases are possible, record why not.
4. Repeat process for other s and p to obtain a comprehensive understanding of the relation between Frame 1 and Frame 2.

Consider the following example. If the frame pair of interest were JUSTIFYING and COMMUNICATION, where JUSTIFYING is Frame 1 and COMMUNICATION is Frame 2, we start by selecting a sentence for JUSTIFYING from FrameNet:

(7) [Kim_{AGENT}] failed to [fully_{DEGREE}] **justify** [his decision_{ACT}].

We then generate a paraphrase by hand using a predicate from the COMMUNICATION frame:

(8) [Kim_{COMMUNICATOR}] failed to [convincingly_{MANNER}] **convey** [his decision_{MESSAGE}].

In this example, a paraphrase is possible. The paraphrase is not minimal, though, since COMMUNICATION frame alone does not express the argumentation/persuasive aspect of a JUSTIFYING event – according to our findings, this disparity holds generally across predicates in COMMUNICATION. As a solution, we can introduce the argumentation aspect through a MANNER adverb (here,

⁵Example 6 is minimal if we assume that the support verb construction *put on trial* is a single FEE.

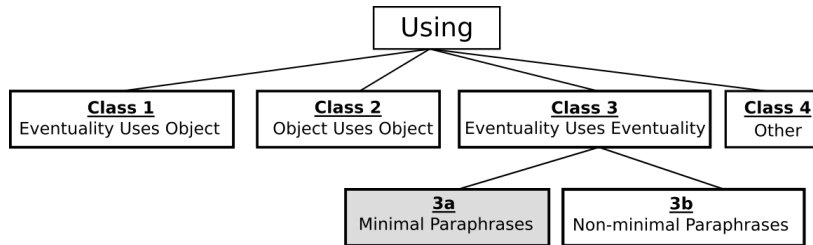


Figure 2: The **Using** relation classification. In Section 4, we focus on class 3a (minimal paraphrases) where one frame can directly paraphrase the other.

“convincingly”) which introduces a **SUASION** frame. Thus, we conclude that **JUSTIFYING** can be paraphrased by a combination of **COMMUNICATION** and **SUASION**.

A subclassification for Using. When looking for regularities that could explain why some frame pairs in **Using** led to paraphrases while others did not, the ontological types of the frame-evoking elements emerged as the main determinant. FrameNet is very inclusive in that it allows all open word classes to evoke frames; however, they clearly differ in the ontological types that they express. For our purposes, the crucial distinction turned out to be between *eventualities* (events and states, Bach (1981)) on the one side, expressed by verbs, deverbal nominalizations, adjectives, or adverbs, and *objects* (in the ontological sense) on the other side, expressed by common nouns. We also introduced a type *other* for frames that are not lexicalized (i.e., frames which do not come with frame-evoking elements) and exist purely as abstract conceptual structures. If a frame contained frame-evoking elements for both eventualities and objects, we assigned it to the class of the majority of frame-evoking elements.

The combinations of these three frame types (eventualities, objects, and other) explain the paraphrasing behavior of the frames in **Using** and give rise to four classes that form the first tier of our proposed classification, as shown in Figure 2.

Class 1 consists of frame pairs where Frame 1 is an eventuality and Frame 2 is

an object frame, and this subclass does not admit minimal paraphrases. Class 2 contains frame pairs where both Frame 1 and Frame 2 are objects, and the paraphrases across frames correspond to traditional lexical entailments. Class 3 consists of frame pairs where Frame 1 and Frame 2 are both eventualities. It is the most interesting class from the perspective of concept-based paraphrasing, since only Class 3 enables the paraphrasing of full predicate-argument structures. As the description of predicate-argument structures is the main motivation for FrameNet, in contrast to purely lexically-oriented resources like WordNet (Fellbaum, 1998), we find that this subclass can be used to exploit the expressivity of FrameNet for concept-based paraphrasing. To further discriminate the types of paraphrases that can be produced by class 3 more clearly, we distinguish between minimal paraphrases (class 3a) and non-minimal paraphrases (class 3b). Finally, class 4 consists of frame pairs where at least one frame is non-lexicalized, and therefore no paraphrasing is possible. We now discuss these classes in detail.

3.1 Class 1: EVENTUALITY Uses OBJECT

In this category, an object (expressed by Frame 2) fills a semantic role and is a core component of the eventuality expressed by Frame 1. An example of this subtype is the frame DRESSING which uses ACCOUTREMENTS. The DRESSING event (“wear”, “don”, “adorn”) takes an ACCOUTREMENT (“anklet”, “watch”, “helmet”) as its CLOTHING frame element.

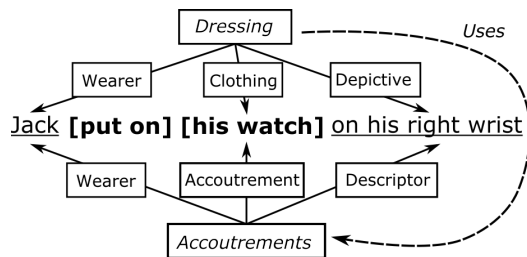


Figure 3: DRESSING uses the ACCOUTREMENTS frame, where the ACCOUTREMENT is a core role in the DRESSING event.

In this example, one frame subsumes the meaning of another frame as part of its conceptual framework. In fact, the object frame is often a core frame element and is a conceptually necessary part of Frame 1. In Figure (3), the DRESSING frame uses the ACCOUTREMENTS frame, where DRESSING requires a CLOTHING frame element. As the ACCOUTREMENTS frame expresses the CLOTHING, it is a requirement for the DRESSING event. Other frame pairs that participate in this sub-type are: TASTING uses FOOD, UNDESSING uses CLOTHING, RECOVERY uses MEDICAL_CONDITIONS, and CREATE_PHYSICAL_ARTWORK uses PHYSICAL_ARTWORKS.

Paraphrasing an eventuality frame (Frame 1) with an object frame (Frame 2) is possible when the eventuality is a state with a relatively generic meaning. For example, in Figure 4, the meaning of the stative predicate “dressed” appears to be missing from its paraphrase in the ACCOUTREMENTS frame. However, we would argue that what we see in Figure 4 is that the DRESSING frame in the stative passive construction can be paraphrased by a combination of ACCOUTREMENT and the preposition “in”. In this instance, “in” is a frame-evoking element for the WEARING frame which expresses the stative dressing event. Informally, then, “being dressed in X” can be paraphrased by “wearing X”.

The example in Figure 3 is typical of a range of cases where an eventuality frame presupposes an object via a **Using** relation, but where the object frame cannot provide a paraphrase on its own. Other examples include the TASTING frame, using the FOOD frame which necessarily evokes an INGESTION in its paraphrase (*she tried the sausage – she had the sausage*), or the CREATE_PHYSICAL_ARTWORK uses PHYSICAL_ARTWORKS which evokes the INTENTIONALLY_CREATE frame in its paraphrase (*Giotto painted a notable series of frescoes – Giotto made a notable series of frescoes*). In all of these cases, FrameNet could alternatively decide to drop the **Using** relation between the eventuality and the object frame and replace it by a more specific relation between two eventuality frames (e.g., DRESSING and WEARING), such as **Causation**. However, in the absence of such relations, creating a paraphrase requires an additional frame in the sentence with the object frame.

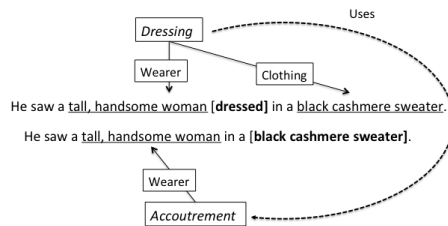


Figure 4: DRESSING uses the ACCOUTREMENTS frame.

3.2 Class 2: OBJECT Uses OBJECT

This subclass consists of two frames that both have nominal, object-denoting predicates. Most of these frame pairs are connected via a relationship that resembles the types of relations that can be found in other ontologies, especially WordNet (Fellbaum, 1998). These relations include meronymy (part-of relation), where the predicates in one frame are components of the predicates in another frame. An example is the frame CLOTHING_PARTS which is composed of objects like “shoelace,” “pocket,” and “sleeve” that uses the CLOTHING frame, which has objects such as “shoe,” “shirt,” and “suit.” Other ontological relationships between nominals can be found in this category as well, including the hyponymy relation. For instance, we find that the frame GIZMO, with predicates such as “tool,” “apparatus,” and “contraption” uses the frame ARTIFACT, which has general nominals such as “technology” and “artifact”.

In the case of meronymy, the second frame can technically minimally paraphrase the first. For instance, in Figure 5, the paraphrase is possible since a person tying his or her shoes entails that it is his or her shoelaces that are actually being tied. That being said, these paraphrases amount to simple lexical entailments which have found much attention in Natural Language Processing (Aharon et al., 2010; Castillo, 2011)

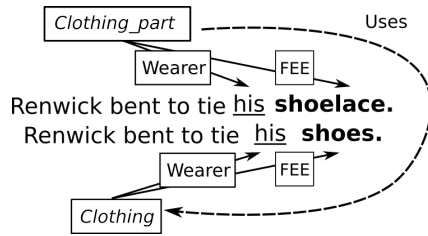


Figure 5: CLOTHING_PART frame uses the CLOTHING frame, where paraphrasing is a lexical substitution.

3.3 Class 3: EVENTUALITY Uses EVENTUALITY

The largest **Using** relation type captures frame pairs in which both frames contain predicates that denote eventualities. While the presence of **Using** generally captures the conceptual similarity between frames, we found that the linguistic structures of the two frames, notably their semantic role inventories, often differ substantially. Consequently, if Frame 1 has a core semantic role whose realization Frame 2 does not support, there cannot be a minimal correspondence across the Frame 1 and Frame 2 paraphrases. This problem gives rise to the two subclasses 3a and 3b: the linguistic structures of the Frame 1 and Frame 2 match in 3a, but not in 3b.

3.3.1 Class 3a: Minimal Paraphrases

The members of this subclass of **Using** are frame pairs where Frame 2 can minimally paraphrase Frame 1 since both their conceptual and linguistic structures match. Empirically, the frames in 3a tend to fit the Ruppenhofer et al. definition (see Section 2.1) of the **Using** relation: the frame pairs are semantically close but differ in their level of abstractness. Thus, the paraphrases that are generated by this sub-category are concept-based, where common world knowledge is necessary to infer one frame as being equivalent to the next.

A maybe somewhat surprising finding is that for most frame pairs in this subclass, paraphrasing is subject to certain side conditions. In Figure 1, the LABELING frame (“call”, “label”, etc.) uses the JUDGMENT_COMMUNICATION

(“accuse”, “praise”, etc.) frame. Not all LABELING events can be paraphrased by JUDGMENT_COMMUNICATION but rather only those which involve a positive or negative judgment that can be expressed by a specific predicate of JUDGMENT_COMMUNICATION. The following gives an example of a “neutral” LABELING event where the paraphrase is not possible:

- (9) a. Back in the olden days, [they SPEAKER] **called** [it ENTITY] [“Bohemian people” LABEL].
 b. *Back in the olden days, [they COMMUNICATOR] **accused/condemned/praised** [it EVALUÉE] [as “Bohemian people” REASON].

This circumstance changes once the speaker attaches a sentiment to the LABEL role. As soon as the LABEL expresses a positively or negatively-valenced term, such as “alcoholic,” “hero,” or “saint,” the relation to the JUDGMENT_COMMUNICATION frame becomes clear:

- (10) a. [They SPEAKER] **called** [her ENTITY] [a saint LABEL].
 b. [They COMMUNICATOR] **praised** [her EVALUÉE] [for being a saint REASON].

In Section 4 below, we concentrate on Class 3a since its minimal paraphrases, which cover about 20% of the frame pairs in **Using**, are both amenable to paraphrasing across eventualities and are simple in that no additional frames are required for successful paraphrasing from Frame 1 to Frame 2. Section 4 will outline how we can add the missing information to **Using**, re-using descriptive devices that already form part of FrameNet.

3.3.2 Class 3b: Non-Minimal Paraphrases

In this class, the semantics of Frame 1 cannot be expressed by Frame 2 alone. In the example in Figure 6, the PRAISEWORTHINESS frame, with adjectival predicates such as “commendable,” “laudable,” and “admirable,” expresses an opinion of an EVALUÉE with respect to some BEHAVIOR. This PRAISEWORTHINESS frame uses the JUDGMENT frame, where the JUDGMENT frame is composed of verbal predicates such as “approve,” “disdain,” and “admire.” JUDGMENT requires a

COGNIZER, which is the sentient person who holds the opinion, and thus the JUDGMENT frame focuses on the act of judgment that this COGNIZER is making. By contrast, the PRAISEWORTHINESS frame focuses on the state of the EVALUEE that has been judged.

Despite a clear conceptual relationship between the PRAISEWORTHINESS and JUDGMENT frames, the realization constraints differ, where JUDGMENT requires a COGNIZER which cannot be provided by PRAISEWORTHINESS alone. A matrix clause like “I believe that ...” could provide the COGNIZER role, which would introduce an additional OPINION frame. Thus, the paraphrase would not be minimal anymore, as two frames would be necessary for the Frame 2 sentence.

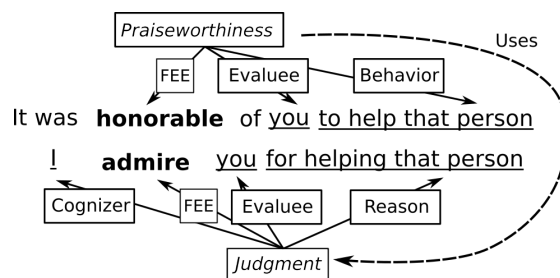


Figure 6: The PRAISEWORTHINESS frame uses the JUDGMENT frame. The COGNIZER role is added to the second sentence in order to satisfy the syntactic constraints of the JUDGMENT frame.

Previous work by Padó (2007) investigated similar cases in German–English translations where the English transitive verb “increase” was translated as the German adjective “höher” (higher) and the semantic roles of the English CAUSE_CHANGE_OF_SCALAR_POSITION frame could not all be realized with the German CHANGE_OF_POSITION_ON_A_SCALE frame. The resulting structures, which they called *frame groups*, structurally correspond to our non-minimal paraphrases. In their observations, they found a large number of frame groups were difficult to classify, and we find that within non-minimal frame pairs, the types of frame groups discovered are not generalizable to other pairs in the **Using** relation. Therefore, we leave a detailed analysis of non-minimal paraphrases to future work and concentrate on minimal paraphrases.

Using subclasses				
Class 1	Class 2	Class 3a	Class 3b	Class 4
Eventuality Uses Object	Object Uses Object	Eventuality Uses Eventuality: Minimal Paraphrase	Eventuality Uses Eventuality: Non-Minimal Paraphrase	Other
95	42	95	197	61

Table 1: Frequency of Relation Types in the **Using** Relation (Total: 490).

3.4 Class 4: Other (Non-EVENTUALITY, Non-OBJECT)

In the final **Using** class, we observe several cases that do not fall into the previous three classes, where frames are not lexicalized and only exist as highly abstract categories in the upper level of the FrameNet hierarchy, consistent with the Ruppenhofer et al. (2010) definition of **Using** discussed in Section 2.1. These instances appear when at least one of the frames in the related frame pair is neither an eventuality nor an object. An example of this type is the TRANSLATING uses MENTAL_ACTIVITY frame pair, where TRANSLATING has predicates such as “translate” and “translation,” but MENTAL_ACTIVITY has no predicates and functions as a purely conceptual frame about cognitive processes or mental states. In this case, paraphrasing across **Using** frames is not possible as one frame is not lexicalized and thus incapable of overt linguistic expression.

3.5 Frequencies of Classes

Table 1 shows the frequencies of the subclasses. The fact that all five are well represented supports our claim about the heterogeneity of **Using**. We observe that class 3 is the most frequent by some margin, accounting for 292 out of the 490 frame pairs, which further supports our claim that FrameNet can serve as a source for concept-level paraphrases.

4 Formalizing Side Conditions for Minimal Concept-driven Paraphrases

In Section 3, we identified a subclass of the **Using** relation, type 3a (minimal paraphrases between eventualities), where the semantic roles in both frames directly correspond to one another. However, as we observed in Section 3.3.1, these paraphrases often come with additional side conditions of a semantic or syntacto-semantic nature that restrict the circumstances under which a sentence using a frame can be paraphrased by another frame. As we found that these side conditions are critical to producing minimal paraphrases across conceptually related frames, the final step of our study is an analysis of the different types of side conditions. We assessed whether they can be formalized within FrameNet, or whether FrameNet needs to be extended conservatively to account for them.

We performed this analysis building on the results of our paraphrasing tests in Section 3, where we recorded the side conditions of minimal paraphrases. We then went through these side conditions for the minimal paraphrase (Class 3a), grouped them, and formalized them. For the formalization, we encode frame-semantic structures as typed attribute-value matrices (AVMs) with coindexation, following various publications on frame semantics (Gawron, 2011; Reiter et al., 2011) and the integration of frame semantics with construction grammar (Kay and Fillmore, 1999) as well as the example of LFG f-structures (Bresnan, 2001). For this representation, we assume that frames can be encoded as AVMs where the semantic roles, plus PREDICATE, form the attributes, and the attributes are themselves AVMs with a feature SEM for the semantics of the role, plus potentially other features.

4.1 Side Condition Type 1: Presence of sentiment

We begin with an example that we have discussed previously: LABELING **Using** JUDGMENT_COMMUNICATION (cf. Example 2, Figure 1). We conjectured above that this paraphrase is only felicitous when the LABEL frame element

of LABELING (the category being given to the ENTITY) carries a negative or positive sentiment, which is then reflected in the polarity of the predicates of the JUDGMENT_COMMUNICATION frame (e.g., “praise” vs. “denounce”). This is a clear case of a semantic (ontological) side condition.

This type of side condition can actually be modeled with information already present in FrameNet via FrameNet’s *semantic types*, which represent general semantic properties of frames, predicates, or semantic roles and have been applied to multiple levels of the FrameNet hierarchy. One semantic type in FrameNet is the “Positive_judgment” and “Negative_judgment” sentiment polarity assigned to lexical units in frames like JUDGMENT_COMMUNICATION or PROMISING.

Assuming that this semantic type information can be represented by a POLARITY feature in the AVM, we can write this side condition as shown in Figure 7. Other frame pairs from **Using** whose paraphrasing relations can be suitably constrained with this type of side condition include DESIRING uses EXPERIENCER_FOCUS, RELIANCE_ON_EXPECTATION uses AWARENESS, and EVENTIVE_COGNIZER_AFFECTING uses SUBJECTIVE_INFLUENCE.

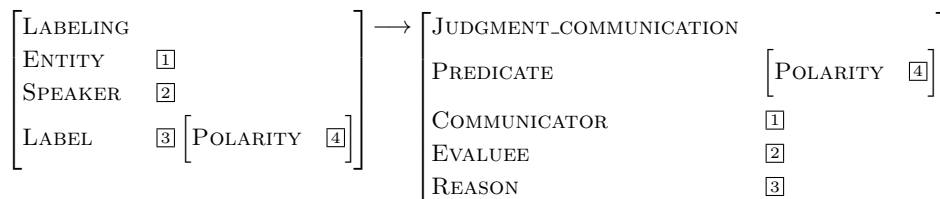


Figure 7: Side Condition 1: Presence of Sentiment.

4.2 Side Condition Type 2: Granularity of Semantic Roles

A second type of side condition applies to frame pairs of **Using** where, although a mapping can be made across the roles in Frame 1 and Frame 2 paraphrases, the frames nonetheless exhibit substantial differences in the granularities that their sets of semantic roles assume. For instance, in the frame ADOPT_SELECTION, an AGENT decides that an ATTRIBUTE should have a specific VALUE:

- (11) It is true that [baroque_{VALUE}] had long been **adopted** [as the style_{ATTRIBUTE}] [for state capitols in the United States_{PURPOSE}].

ADOPT_SELECTION uses the CHOOSING frame, in which a COGNIZER picks an option, expressed by the CHOSEN role:

- (12) Microsoft indicated [it_{COGNIZER}] will probably **pick** [another RISC processor_{CHOSEN}] [before the end of the year_{TIME}].

Thus, the question is how the CHOSEN role corresponds to the two roles ATTRIBUTE and VALUE in the paraphrase from Frame 1 to Frame 2. There are two potential options to construct paraphrases in these cases. First, CHOSEN could just express the VALUE, with the ATTRIBUTE remaining implicit. Paraphrasing the ADOPT_SELECTION sentence from Example (11) would then yield:

- (13) It is true that [baroque_{CHOSEN}] had long been **picked** [for state capitols in the United States_{INHERENT_PURPOSE}].

The second option is to realize the CHOSEN role as a complex phrase incorporating both roles, with the head designating the ATTRIBUTE and the modifier designating the VALUE:

- (14) It is true that [the baroque style_{CHOSEN}] had long been **picked** [for the state capitols in the United States_{INHERENT_PURPOSE}].

Since the first alternative (13) does not offer a general way of paraphrasing fully instantiated ADOPT_SELECTION frames, we concentrate on the second option (14), and find that a side condition characterizing felicitous paraphrases of this type needs to consider two syntactic variants of the original sentence.

The first one is Example 11, where ATTRIBUTE is an NP and VALUE a PP attached to the frame-evoking element. In this case, we can adopt the side condition shown in Figure 8, which combines the two roles into a single phrase. It introduces features (semantic) HEAD and MOD to refer to the constituent parts of a role-bearing phrase.

This side condition, however, does not apply to cases where ATTRIBUTE and VALUE are realized within one complex NP, with the ATTRIBUTE as the head

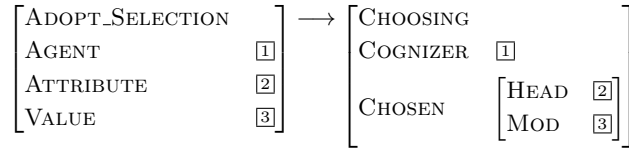


Figure 8: Side Condition 2: Difference in Granularity of Semantic Roles (Split ATTRIBUTE/VALUE)

and the complete phrase as its VALUE, as in many FrameNet annotations such as the following:

- (15) [The session _{AGENT}] **adopted** [a predominantly hostile [tone _{ATTRIBUTE}] towards the forces of the Bosnian Serbs _{VALUE}].

Since in this case, the complete NP already expresses the concept that we need, we adopt the side condition in Figure 9 which simply adopts the VALUE phrase for the CHOSEN role.

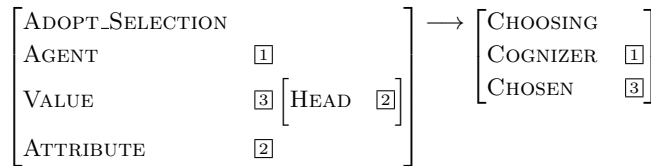


Figure 9: Side Condition 2: Difference in Granularity of Semantic Roles (Combined ATTRIBUTE/VALUE)

This side condition is syntacto-semantic in nature, not constraining the semantics of role fillers (as was the case in Side Condition Type 1), but their realization possibilities. **Using** frame pairs that require the specification of similar side conditions relating to the granularities of their respective role sets include ADDUCING uses STATEMENT and BEYOND_COMPARE uses SURPASSING.

4.3 Side Condition Type 3: Presence of Semantic Roles

As a third category of side condition, we consider cases where the overt realization of a semantic role is important for paraphrasing. We focus on the



Figure 10: Paraphrasing with the BEAT_OPPONENT and the WIN_PRIZE frames.

BEAT_OPPONENT and WIN_PRIZE frames, where the former uses the latter. Both are descriptions of an event in which there could be a WINNER/COMPETITOR and a LOSER/OPPONENT, often in the context of a competition. However, the BEAT_OPPONENT frame describes general competitions and thus can be used to describe informal competitions that do not end in an award or a prize, such as in the following example:

- (16) No matter how hard you try, [they WINNER] always **beat** [you LOSER] [to the huts COMPETITION].

Paraphrasing the above BEAT_OPPONENT sentence with the WIN_PRIZE frame is not possible as the entire purpose of the WIN_PRIZE frame is to emphasize the title or prize that is achieved at the end of the activity, and forms a core component in the WIN_PRIZE conceptual structure.

Paraphrases for BEAT_OPPONENT can thus only be produced with WIN_PRIZE if BEAT_OPPONENT expresses the non-core PRIZE role. Figure 10 shows an example where this leads to a successful paraphrase.

This side condition can not be directly expressed as part of an AVM with coindexation, since unification cannot directly test for the absence of attributes in the input. However, we follow the example of LFG f-structure (Bresnan, 2001), we assume that there is a type of side condition that can check for the presence of attributes and use the notation $A \neq \emptyset$ to denote this side condition. Then we can require the presence of the non-core PRIZE role in addition to the relevant

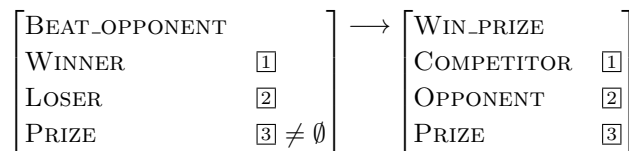


Figure 11: Semantic Side Condition 3: Presence of (non-core) Semantic Role.

coindexations as shown in Figure 11.

Additional frame pairs that have this kind of side conditions requiring the presence of a particular role include `ARRANGING` uses `PLACING`, `BEYOND_COMPARE` uses `SURPASSING`, and `PARTIALITY` uses `TAKING_SIDES`.

5 Conclusion

In this article, we have presented a study of FrameNet’s **Using** relation in terms of its suitability for paraphrasing. In contrast to earlier studies, we have focused not on paraphrases that are backed by relatively straightforward linguistic or ontological relations between predicates (called language-based paraphrases in previous work), but by the disjoint set of paraphrases that are backed by world knowledge (concept-based paraphrases). We have argued that **Using**, a lesser-studied frame-to-frame relation, has the potential of accounting for a large number of such paraphrases, but that the analysis of **Using** must be revised and refined before it can become useful. To that end, we made two contributions.

The first contribution is to inventorize and categorize the frame pairs connected by **Using** into four classes that differ in the types of paraphrases that they support. We identify one of the classes (class 3a) to be most interesting in the sense that it yields concept-based *minimal paraphrases*, that is, paraphrases that involve just one frame each in the original and in the paraphrased sentence. A pertinent question is to what extent our four-class classification captures the major classes not just of **Using**-based paraphrases, but of concept-based paraphrases in general. This hypothesis seems plausible in the sense that classes

1–3 cover the fundamental combinatorial possibilities of objects and events ⁶. An assessment of this question is a matter of future research, for example by analysis of paraphrase corpora (Dolan and Brockett, 2005).

Our second contribution is to investigate the *side conditions* that are necessary to obtain felicitous paraphrases for frames that belong to the third subclass of **Using**. We identified three major types of such side conditions. All of these can be specified either within the existing FrameNet database structure (Baker et al., 2003) or with relatively conservative extensions. Required extensions motivated by the work in this article include the ability to specify side conditions for **Using** relations (which can build on the existing formalization of frame-to-frame relations) and the ability to specify these side conditions, including realization statuses of semantic roles, reference to semantic heads and modifiers of semantic roles, and the extension of the notion of polarity, which currently exists for predicates, to frame elements.

In terms of automatic generation of concept-based paraphrases, this would require an exhaustive identification and implementation of semantic side conditions for the **Using** relation. We believe that a FrameNet resource, when extended in this manner, should support the declarative generation of concept-based paraphrases with relatively simple approaches along the lines of Ellsworth and Janin (2007). Such a setup would provide an attractive scenario to validate, and if necessary, refine semantic side conditions for conceptual paraphrases.

Finally, a limitation of the present study is that we temporarily set aside the class of non-minimal, concept-based paraphrases (class 3b). This was a pragmatic decision; we found it reasonable to first understand the case with more generalizability (minimal paraphrases) before proceeding to the more nuanced case which involves multiple frames on at least one side. We will proceed to analyze non-minimal, concept-based paraphrases in future work.

⁶Valid under the condition that the conceptual paraphrases frame semantics cannot account for are either folded into class 4 or put into a fifth class

References

- Abelson, R. P. (1981). Psychological status of the script concept. *American psychologist*, 36(7):715.
- Aharon, R. B., Szpektor, I., and Dagan, I. (2010). Generating entailment rules from FrameNet. In *Proceedings of ACL*, pages 241–246, Uppsala, Sweden.
- Bach, E. (1981). On time, tense, and aspect: An essay in English metaphysics. In Cole, P., editor, *Radical Pragmatics*, pages 63–81. Academic Press, New York.
- Baker, C. F., Fillmore, C. J., and Cronin, B. (2003). The structure of the Framenet database. *International Journal of Lexicography*, 16(3):281–296.
- Baker, M. (1992). *In Other Words*. Routledge.
- Bergen, B. and Chang, N. (2005). Embodied construction grammar in simulation-based language understanding. In *Construction grammars: Cognitive grounding and theoretical extensions*, pages 147–190.
- Botschen, T., Mousselly Sergieh, H., and Gurevych, I. (2017). Prediction of frame-to-frame relations in the FrameNet hierarchy with frame embeddings. In *Proceedings of the 2nd workshop on representation learning*, pages 146–156.
- Bresnan, J. (2001). *Lexical-Functional Syntax*. Blackwell Publishers, Oxford.
- Castillo, J. J. (2011). A WordNet-based semantic approach to textual entailment and cross-lingual textual entailment. *International Journal of Machine Learning and Cybernetics*, 2(3):177–189.
- Chang, N., Narayanan, S., and Petruck, M. R. (2002). From frames to inference. In *Proceedings of the First International Workshop on Scalable Natural Language Understanding*.
- Chow, I. C. and Webster, J. J. (2007). Integration of linguistic resources for verb classification: FrameNet frame, WordNet verb and Suggested Upper Merged

- Ontology. In *Computational Linguistics and Intelligent Text Processing*, pages 1–11. Springer, Berlin/Heidelberg.
- Coyne, R. and Rambow, O. (2009). Lexpar: A freely available English paraphrase lexicon automatically extracted from FrameNet. In *Proceedings of the Third IEEE International Conference on Semantic Computing*.
- Dolan, W. B. and Brockett, C. (2005). Automatically constructing a corpus of sentential paraphrases. In *Proceedings of the Third International Workshop on Paraphrasing*, pages 9–16.
- Ellsworth, M. and Janin, A. (2007). Mutaphrase: Paraphrasing with FrameNet. In *Proceedings of the ACL-PASCAL Workshop on Textual Entailment and Paraphrasing*, pages 143–150, Prague.
- Fellbaum, C., editor (1998). *WordNet An Electronic Lexical Database*. The MIT Press, Cambridge, MA ; London.
- Fillmore, C. J. (1985). Frames and the Semantics of Understanding. *Quaderni di Semantica*, IV(2):222–254.
- Fillmore, C. J., Baker, C. F., and Sato, H. (2004). FrameNet as a “net”. In *Proceedings of the Fourth International Conference on Language Resources and Evaluation*.
- Fillmore, C. J., Johnson, C. R., and Petruck, M. R. L. (2003). Background to FrameNet. *International Journal of Lexicography*, 16(3):235–250.
- Fodor, J. A. (1961). Projection and paraphrase in semantics. *Analysis*, 21:73–77.
- Gawron, J.-M. (2011). Frame semantics. In *Handbook of Semantics*, volume 1 of 33, pages 664–687. de Gruyter.
- Hasegawa, Y., Lee-Goldman, R., Kong, A., and Kimi, A. (2011). FrameNet as a resource for paraphrase research. *Constructions and Frames*, 3(1):104–127.
- Jackendoff, R. S. (1990). *Semantic Structures*. The MIT Press, Cambridge, MA.

- Katz, J. J. and Fodor, J. A. (1963). The structure of a semantic theory. *Language*, 39:170–210.
- Kay, P. and Fillmore, C. J. (1999). Grammatical constructions and linguistic generalizations: the What’s X doing Y? construction. *Language*, 75:1–33.
- Leech, G. (1974). *Semantics*. Penguin.
- Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. University of Chicago press.
- Narayanan, S. (2014). Bridging text and knowledge with frames. In *Proceedings of Frame Semantics in NLP: A Workshop in Honor of Chuck Fillmore*, pages 22–25.
- Norvig, P. (1987). Inference in text understanding. In *Proceedings of the National Conference on Artificial Intelligence*, pages 561–565, Seattle, WA.
- Padó, S. (2007). Cross-lingual parallelism and translational equivalence: The case of FrameNet frames. In *Proceedings of the NODALIDA Workshop on Building Frame Semantics Resources for Scandinavian and Baltic Languages.*, Tartu, Estonia.
- Petruck, M. R. and de Melo, G. (2012). Precedes: A semantic relation in FrameNet. In *Proceedings of the Workshop on Language Resources for Public Security Applications*, pages 45–49.
- Petruck, M. R., Fillmore, C. J., Baker, C. F., Ellsworth, M., and Ruppenhofer, J. (2004). Reframing FrameNet data. In *Proceedings of The 11th EURALEX International Congress*, pages 405–416.
- Reiter, N., Hellwig, O., Frank, A., Gossmann, I., Larios, B. M., Rodrigues, J., and Zeller, B. (2011). Adapting NLP tools and frame-semantic resources for the semantic analysis of ritual descriptions. In Sporleder, C., van den Bosch, A., and Zervanou, K., editors, *Language Technology for Cultural Heritage: Selected Papers from the LaTeCH Workshop Series*, pages 171–193. Springer.

- Ruppenhofer, J., Ellsworth, M., Petruck, M. R. L., Johnson, C. R., and Scheffczyk, J. (2010). FrameNet II: Extended theory and practice. <https://framenet2.icsi.berkeley.edu/docs/r1.5/book.pdf>. Revision of Nov 1, 2016.
- Schank, R. C. and Abelson, R. (1977). *Scripts, goals, plans, and understanding*. Hillsdale, NJ: Erlbaum.
- Scheffczyk, J., Pease, A., and Ellsworth, M. (2006). Linking FrameNet to the SUMO ontology. In *Proceedings of the International Conference on Formal Ontology in Information Systems*, Baltimore, MD.
- Schreyer, R. (1978). On paraphrase relations. In *Wortstellung in Bedeutung: Akten des 12. Linguistischen Kolloquiums*, pages 225–235, Pavia, Italy. Max Niemeyer, Tübingen.