Constructs, Frames and Event Structure

William Croft and Meagan Vigus
Department of Linguistics
MSC 03 2130
1 University of New Mexico
Albuquerque NM 87131-0001

Abstract

Charles Fillmore developed both the theory of construction grammar and the theory of frame semantics. Fillmore’s original frame semantic model included broad frames such as Commercial Transaction and Risk. FrameNet II (the current incarnation) has a complex lattice of frame-to-frame relations, in which the original frames are often abstract frames, and/or have been decomposed into distinct concrete frames. The frame-to-frame relations are of different types beyond simple taxonomic relations. The concrete frames are often more restricted in their argument structure construction possibilities than the original, more general frames. We argue that some frame-to-frame relations effectively represent the internal structure of the core events in the frame; the event structures are associated with different argument structure constructions. Introducing event structure into the frame semantic representation offers the potential to simplify the FrameNet lattice. Brief analyses of the Commercial Transaction and Risk frames are presented.

In this paper we explore the relationship between construction grammar, in particular argument structure constructions, frame semantics and event structure. Charles Fillmore developed the theory of frame semantics out of case grammar (Fillmore 1977a; 1982; 1985), and the theory of construction grammar at about the same time (Fillmore, Kay, and O’Connor 1988).

A semantic frame is a body of knowledge made coherent by the human activity that holds it together: ‘a system of categories structured in accordance with some motivating context...The motivating context is some body of understanding, some pattern of practices, or some history of social institutions against which we find intelligible the creation of a particular category in the history of the language community’ (Fillmore 1982, 119). Semantic frames are frequently centered on a type of event (a pattern of practices, a social institution) such as eating, placing something, or engaging in a commercial transaction (buying, selling etc.).

The implementation of semantic frames in FrameNet (Fillmore, Johnson, and Petruck 2003; Ruppenhofer et al. 2016) consists of an informal prose description of the event that forms the core of the semantic frame, and an enumeration of frame elements (FEs) of varying degrees of centrality and uniqueness to the frame. For example the Theft frame has core FEs of a Perpetrator, a Victim, Goods and Source (location of the Goods), as well as non-core FEs such as Instrument, Manner, Means, Purpose, Explanation, Frequency, Place and Time. An example is Leslie stole the watch from Kim.

However, the nature of frames in FrameNet differs from that in Fillmore’s earlier work, and in earlier versions of FrameNet. For instance, in Fillmore (1977b, 59), all of the following sentences are considered to be instances of the commercial transaction ‘scene’ (as it was called in that paper):

1. John bought the sandwich from Henry for three dollars.
2. Henry sold John the sandwich for three dollars.
3. John paid Henry three dollars for the sandwich.
4. The sandwich cost three dollars.

In the current version of FrameNet (FrameNet II), each of these sentences are found in a different frame: Commerce_buy, Commerce_sell, Commerce_pay, and Expensiveness respectively. The unity of the commercial transaction ‘scene’ is captured by an abstract Commercial_transaction frame and a set of frame-to-frame relations. FrameNet II includes a large range of frame-to-frame relations, forming a complex lattice of frames.

Figure 1 gives the relevant part of the frame lattice including the frames for buy, sell and pay (Fillmore 2008, 54). The commercial transaction event is analyzed as a complex frame with two components, transferring the goods and transferring the money, linked by a Subframe relation. Each of the latter two frames can be assigned two different Perspectives: from the point of view of the buyer or the point of view of the seller. In order to include cost, FrameNet uses a still more abstract Commerce_scenario frame, of which...
Commercial_transaction is a Subframe; the Expensiveness frame is in a “Uses” relation to the Commerce_scenario frame.

One effect of the decomposition of the commercial event into multiple frames is that the argument realization, in particular the core argument realization, of the participants (frame elements or FEs) is uniform within the more specific frames. The commercial transaction examples have the following FE argument realizations; each one expresses the FEs in different grammatical roles (subject, object, for phrase, from phrase):

(5) Buyer buy Goods from Seller for Money
(6) Seller sell Buyer Goods for Money
(7) Buyer pay Seller Money for Goods
(8) Goods cost Money

In other words, the FrameNet II frames divide lexical items, particularly verbs, into separate frames effectively depending on the argument structure constructions they occur in. The semantic commonalities for a frame in the more general (and earlier) sense of that term are captured by abstract frames and a variety of frame-to-frame relations.

This represents a long-term tension in frame semantics between a more general notion of frame as a ‘motivating context’ and a notion of frame more specific to particular syntactic realizations of argument structure based on case grammar. Fillmore (1977b) calls the more general semantic structure a ‘scene’ and the argument structure that highlights or perspectivizes a part of the scene a ‘case frame’. Baker et al. (1998), which describes an early version of FrameNet, calls the more general structure a ‘frame’ and the argument-structure-specific entity a ‘scene’. In that version of FrameNet, the only frame-to-frame relation is inheritance. In FrameNet II, both the more general and more specific structures are frames, and the positing of abstract frames and different types of frame-to-frame relations allow the representation of both types of frame as the same data structure.

In this paper, we suggest another way to resolve this tension between ‘frame’ as motivating context and ‘frame’ as a more specific semantic structure related to argument structure/case frame. We propose that the more specific semantic structure can be accounted for by analyzing the event structure of the motivating context and using the analysis of event structure to handle the argument realization variation. In other words, event structure is not equated to frame structure (though in some cases there may be only one event structure that is lexically realized for a particular frame). Event structure semantics is instead (or in addition) associated with argument structure constructions (Goldberg 1995; 2006; Croft 2003; 2012; 2015).

FrameNet has not exploited event structure in its frame representation. For each frame there is an informal prose description that specifies the interactions between frame elements (participants, broadly construed) in the frame as the event(s) central to the frame unfold over time. FrameNet’s prose description is a neutral choice for representation, since there are many models of event structure proposed in the formal semantic, cognitive semantic and computational linguistic literatures.

The consequence is that event structure is analyzed in frame-to-frame relations instead. The Inchoative relation links a stative event frame to the frame for the process that results in the state. The Causative relation links a processual event frame to the frame that includes an external cause (human agent or physical entity). The Subframe relation links a complex event to component subevents, which may be expressed with different verbs, and are sometimes in temporal sequence, described by the Precedes relation. The Perspective relation and sometimes the Uses relation is used for “focusing” on one participant or set of participants in a frame (in the general sense) as opposed to others. For example, Commerce_buy focuses on the Buyer and the Goods, while Commerce_collect, as in "Henry charged John three dollars for the sandwich", focuses on the Seller and the Money.

We propose that analyzing the structure of the events in frames allows one to capture the relationships represented in the Inchoative, Causative and Perspective relations, and the relationships represented in some cases of the Subframe and Uses relations. This is a strong claim whose fulfillment rests on two bases. The first is an analysis of event structure that can account for these relations. We will use the event analysis presented in detail in Croft (2012) and currently being elaborated. The second is to apply the event structure analysis to the FrameNet frames. Obviously this is impossible to do in a short paper, and is a task we have hardly begun. We will restrict ourselves to brief event structure analyses of the commercial event frame (Croft, Taoka, and Wood 2001; Croft 2012) and an event structure analysis of the RISK frame (Fillmore and Atkins 1992; 1994).

### Event Structure Analysis

Although there are a variety of analyses of event structure that have been proposed, there are certain commonalities among the analyses. The primary dimensions of event structure are aspect and the (mostly) causal network of interactions among the participants in the event.

Aspect describes how an event unfolds over time. Models of the aspectual structure of events in different theoretical traditions have converged on a phasal model, in which events unfold over distinct phases of states, transitions and processes. Croft proposes a richer phasal model than most others, including states, transitions and processes. Croft’s representation also allows for different profiled phases in the aspectual contour of an event. For example, *I see Mount Tamalpais* profiles the state of the mountain being in the visual field, while *Suddenly I saw Mount Tamalpais* profiles the transition from not seeing to seeing the mountain.

Aspectual structure does not play a major role in FrameNet. Nevertheless, one of the criteria for splitting frames in the FrameNet II manual is if the events are aspectually distinct (Ruppenhofer et al. 2016). Also, one of the frame-to-frame relations is Inchoative. Instead, we would propose that inchoative and corresponding state predicates, as in the two senses of *see* just described, would belong to the same frame, and the distinction captured by a difference
in the profiled phase of the aspectual structure of the event in the frame.

FrameNet structure is primarily focused on the participants—frame elements—and their interactions. This forms the basis of the frame analysis, namely the set of frame elements, with their interactions described informally in prose. There is less consensus on how to represent the relations among participants in an event, but it would probably be safe to say that there is general agreement on representing the relations among participants as a network of causal and noncausal interactions.

Croft (1991; 2012), building on Talmy (1976; 1988) argues that the event structure realized in a single clause with a single argument structure forms a causal chain, that is, a directed, acyclic, nonbranching structure, and that argument realization is determined in part by the relative position of participants in the causal chain.

Of course, events in the real world involve interactions that do not form a causal chain. Some interactions among participants are noncausal, for example the spatial relation between the figure and ground, or the possession relation between possessor and possessed. Croft argues that these relations are construed as directed, for example with the figure antecedent to ground, as in examples 9-10.

(9) Bill poured the port into the decanter.

Bill \(\Rightarrow\) port \(\Rightarrow\) into decanter

(10) Bill filled the decanter with the port.

Bill \(\Rightarrow\) with port \(\Rightarrow\) decanter

When the figure is the direct object, the ground is expressed with a path preposition. When the ground is the direct object, the figure is expressed with the preposition with. Croft argues that prepositions can be divided into two categories, subsequent oblique prepositions that govern participants subsequent to the direct object participant in the causal chain, and antecedent oblique prepositions such as with that govern participants antecedent to the direct object participant in the causal chain. English subsequent prepositions include spatial path prepositions, metaphorical (nonspatial) to, and for; antecedent prepositions include with, by, of and metaphorical from.

Certain segments in the causal chain are profiled depending on the argument structure construction, just as only certain phases of the aspectual contour are profiled depending on the tense-aspect construction. In examples 9-10, the double arrow represents the profiled part of the causal chain, delimited by the subject and object referents which are boldfaced. In example 9, only the chain from Bill to the port is profiled, whereas in example 10, the chain from Bill to the decanter is profiled. Profiling generally "focuses" the participants that delimit the profiled part of the chain, and are realized as subject and object.

The aspectual structure and causal chains can be integrated into a single event structure representation (Croft 2012; 2015; Croft, Pešková, and Regan 2016). This is achieved by decomposing events into subevents for each participant. Each participant’s subevent can be thought of as what the participant does, or what happens to the participant, over the time course of the event. Each participant’s subevent enters into a causal or other relation with the subevents of the other participants over the time course of the event. This representation will be illustrated in the analysis of the Risk frame below, where subevent structure is relevant for understanding the argument structure constructions in which Risk verbs occur.

**Disentangling Causal Networks: The Commercial Transaction Frame**

The other way in which participants in an event do not obviously form a causal chain is when the event has multiple participants in a complex network of relations. A classic example is the commercial transaction event: ‘two people are active, and each of the two performs two acts, the buyer that of taking the goods and that of surrendering the money, the seller that of taking the goods and that of surrendering the money’ (Fillmore 1977b, 58). This network of relations is represented in Figure 2 (Croft 2012, 248).

![Figure 2: Network of relations among participants in the commercial transaction event](image)

Both human participants are initiators towards each other, but the other relations are all noncausal: possession relations which change over the course of the event, and a relation of equality of value between the money and the goods.

But any specific verb and argument structure can represent only a causal chain, so only part of the network can be expressed in a clause with that verb and argument structure construction. Croft et al. (2001) describe three principles constraining the construal of a causal chain in such a network:

- The causal chain must be a single connected path in the network
- The direction of the causal chain must match the direction of the segments in the selected path
- The expression of participants in the causal chain as obliques must use antecedent and subsequent oblique forms appropriate to their position in the causal chain relative to the participants realized as subject and object

All of the commercial transaction verbs and the argument structure constructions they occur in conform to the network of interactions in Figure 2 and the three constraints; this is also true in Russian and Japanese (Croft, Taoka, and Wood 2001; Croft 2012). Admittedly, that is not too difficult, since the interactions among the participants in the commercial transaction frame constitutes a fully connected graph. Also,
we must allow for the double object construction, in which the noncausal possession relation is realized grammatically through argument phrases in the same grammatical position (object). The double object examples can be replaced with to-construction examples which express more clearly the causal chain.

(11) John bought the sandwich from Henry for three dollars.

\[
\text{from Henry} \rightarrow \text{John} \rightarrow \text{sandwich} \rightarrow \text{for $3}
\]

(12) Henry sold the sandwich to John for three dollars.

\[
\text{Henry} \rightarrow \text{sandwich} \rightarrow \text{to John} \rightarrow \text{for $3}
\]

(13) John paid three dollars to Henry for the sandwich.

\[
\text{John} \rightarrow \text{$3} \rightarrow \text{to Henry} \rightarrow \text{for sandwich}
\]

(14) The sandwich cost three dollars.

\[
\text{sandwich} \rightarrow \text{$3}
\]

The prepositions from and to realize antecedent and subsequent obliques respectively, in their metaphorical use for transfer of possession in these examples. Hence a from-marked participant must be antecedent to the direct object participant, and the to-marked participant must be subsequent to the direct object participant.

The preposition for is a bit more complicated. In example 11, the subject referent gives up the item described by the for phrase (the money). Croft, Taoka and Wood (2001) call the verb in this type of construction a get type verb, since the subject ends up with the direct object referent (the goods). In examples 12 and 13, the subject referent ends up with the item described by the for phrase (the goods). Croft, Taoka and Wood (2001) call the verb in this type of construction a give-type verb, since the subject referent gives up the direct object referent (the money).

The semantic relation of for is different in give type and get type verbs. In give type verbs, the for of purpose, a broadening of the for of benefit of Jackendoff (1990, 183-184), is used: the action is carried out by the subject referent with the aim of getting the referent of the for phrase. Other examples of the for of purpose are given below (Jackendoff 1990, 184):

(15) Bill looked for Harry.

(16) Bill aimed for the target.

In get type verbs, substitution for is used: the action involves a replacement of one entity (the money) with another entity (the goods) in a similar role, in this case the role of a certain value. Other examples of substitution for are given below (Croft, Taoka, and Wood 2001, 589):

(17) The director substituted Mary for Janet in the lead role.

(18) John stood in for Mary.

(19) I feel for you.

In terms of this analysis, the different commercial transaction frames that are children of Commercial_transaction in FrameNet in Figure 2 are capturing events that select different paths (causal chains) in the network of interactions of the participants (frame elements) in the commercial transaction frame. We propose that there need only be one Commercial_transaction frame, and the event structure analysis will allow for the different realization of frame elements in different roles in different argument structure constructions.

Participants and Their Subevents: The Risk Frame

Fillmore and Atkins (1992, 79) defines the Risk frame as making reference to ‘the possibility of an unwelcome outcome’. The main participants in this frame are shown below: the Fillmore & Atkins (1992) terminology is first, followed by the equivalent frame element in the Run_risk frame, and then the equivalent frame element from the Daring frame. The Daring frame does not have any equivalent frame elements for the Valued Object and Harm, for reasons explained below. The terminology from Fillmore and Atkins (1992) will be used in this section.

Actor/Protagonist/Agent: entity that performs the Deed

Deed/Action/Action: action that bring about the potential of Harm to the Valued Object

Valued Object(VO)/Asset: entity that may be hurt, lost, or otherwise damaged if the Harm occurs

Harm/Bad outcome: potential negative outcome of the Deed

Purpose/Purpose/Purpose: potential positive outcome of the Deed

This frame includes all of the meanings of the verb risk itself, along with danger, peril, hazard, and venture (Fillmore and Atkins 1992, 80). Although Fillmore and Atkins (1992) and FrameNet include many different parts of speech, the analysis here will focus only on verbs.

In the current FrameNet analysis, the verb risk is split between the Run_risk frame and the Daring frame. The Daring frame includes the verb risk when it occurs with the Deed expressed as a direct object. The Daring frame includes the verbs risk, chance, dare, hazard, and venture. The Run_risk frame includes all other argument structures of the verb risk. The verbs in the Run_risk frame are risk and wager. The other verbs included in the original Risk frame of Fillmore and Atkins (1992), endanger and imperil, occur in the Endangering frame.

FrameNet also includes a general Risk_scenario frame. However this frame only links together frames involving the different parts of speech of risk. The relevant frames and frame-to-frame relations for the original Risk frame are shown in Figure 3.

The Risky situation, Run_risk, and Being_at_risk frames are Perspectives on the Risk_scenario. The Wagering frame uses the Run_risk frame; the Endangering frame is a Causative of the Being_at_risk frame. The Risk_scenario frame itself does not contain any lexical units. The Risky_situation and Being_at_risk frames do not include any verbs. Interestingly, the Daring frame is not related to any of
Figure 3: Risk frames in FrameNet

the frames in Figure 3. Here, not only has FrameNet split a single frame into multiple frames, but one of those frames is not related to the others.

Like the Commercial Transaction frame, the different argument structures that risk and related verbs occur in can be captured in a single frame. Unlike the Commercial Transaction frame, the differences in argument realization do not come from a complex network of causal interactions. The participants in the risk event can be represented in a single causal chain. The differences in argument structure come from how the participants and their subevents are expressed.

Using the representations of participants and subevents from Croft (2012), the relations between elements of the risk frame are shown below in Figure 4 and Figure 5.

In the event structure decomposition in Croft (2012), each participant is associated with its own specific subevent, as we noted above. The subevents represent the aspectual structure, that is, the phases undergone by the participant over the time course of the event.

Figure 4 shows the relations between participants and subevents when the Actor is a separate entity from the Valued Object. The participants are shown on the left with their grammatical realizations; the events are on the right. The horizontal dimension represents time and the vertical dimension qualitative change. Each subevent consists of three phases: a rest phase before the event, the profiled transition phase asserted by the verb, and a result phase after the profiled phase. The nonprofiled phases are represented by dotted lines.

In the case of the Risk frame, not all of the participants identified by Fillmore & Atkins (1992) are truly participants, that is, persons, places or things. The Actor and Valued Object are true participants, i.e. they are entities and not events. The Deed and Harm, however, are subevents associated with the Actor and the Valued Object respectively. The Purpose is also usually expressed as a subevent; the participant in the Purpose subevent may be a Beneficiary.

While it is clear that the Deed actually occurs, the Harm and the Purpose are just possible subsequent subevents. The Deed may not lead to Harm to the Valued Object, and the Deed (and whatever effect it has on the Valued Object) may not achieve its Purpose.

In this event, the Deed represents the Actor’s subevent. The Actor’s performance of the Deed is profiled and definitely occurs; therefore it is marked with a solid line. The vertical arrows in the figure represent causal interactions. The Actor’s performance of the Deed causes the possibility of Harm to the Valued Object. Since this is only a possibility and not a certainty, the qualitative change to the Valued Object is represented with a dashed line. The possibility of Harm to the Valued Object causes the possibility of a Purpose subevent. As with the Harm, since this is only a possibility, it is represented with a dashed line.

Either the participants or their subevents may be realized in argument phrases. In example 20 below, the participants are realized as arguments and the exact nature of the subevents is left unspecified.

(20) Why did he risk his life for a man he did not know? (Fillmore and Atkins 1992, 88)

When the participants are realized as arguments, they are expressed by noun phrases. The Actor is realized as subject, the Valued Object as (grammatical) object, and the Purpose (here, a Beneficiary) with for, a subsequent oblique.

When the subevents are realized as arguments in the clause, they have a wider range of grammatical encoding. In example 21, the Deed is expressed as a noun and realized as subject; the Harm is expressed as the object.

(21) Aye, such heedlessness – such invitation to injury, as if to a friendly playmate – can imperil our battle planning, risking loss of personnel and materiel (FrameNet)

A single argument structure can realize both participants and subevents as arguments, as in example 22 below.

(22) He had risked two of his submarines by sending them to the edge of the American beaches. (Fillmore and Atkins 1992, 90)

The Actor is expressed as subject and the Valued Object as direct object. The Deed that the Actor performs is realized as a gerundial phrase introduced by the antecedent oblique by. An antecedent oblique expresses a participant antecedent to the direct object in the causal chain. Since the Deed along
with the Actor as antecedent to the Valued Object, its expression with an antecedent oblique preposition is expected.

Figure 5 shows the event type in which the Actor and the Valued Object are the same entity. Here, the Deed and the Harm are both subevents of the same entity (the Actor/Valued Object). Argument structure constructions expressing this event can realize either the Deed or the Harm as direct object.

In example 23, the Deed is the direct object; in example 24, the Harm is realized as the direct object. In both examples 23 and 24, the Actor is the subject and the Purpose is expressed with the subsequent oblique phrase (in order) to.

(23) Lance risked staying at the ridge alone to watch.
   (Fillmore and Atkins 1992, 96)

(24) We are prepared to risk a substantial increase in unemployment in order to bring inflation to an end.
   (Fillmore and Atkins 1992, 94)

When the Deed and the Harm are both expressed, as in example 25, the Deed is realized with the antecedent oblique by and the Harm as the direct object.

(25) Mrs. Gore even risked the wrath of the record industry by campaigning to have warning labels put on particularly offensive records. (FrameNet)

This is consistent with the representation in Figure 5 because the Deed subevent occurs before the (potential) Harm subevent and is accordingly expressed with an antecedent oblique. In fact, the Deed is a somewhat more complex event structure since it involves an additional participant to the ones in Figure 5, namely the record companies (the source of the wrath possibly incurred by Mrs Gore).

In the FrameNet analysis, argument structures like those in examples 20, 21, 22, 24, and 25 are represented in the Run risk frame. Argument structures like that in example 23, where the Actor is subject and the Deed is the direct object are separated, are found in the Daring frame. In the Daring frame, the Actor and the Valued Object are always the same participant with just its subevent; this is why the Daring frame lacks the Valued Object and Harm subevents.

As this analysis has shown, however, argument structures where the Deed and the Harm are direct objects are actually closer to each other semantically, in that the Actor is necessarily construed as the Valued Object as well. Since all of these argument structures involve the same participants and the same subevents, we propose that they should be represented together in a single frame in FrameNet.

**Conclusion**

Our analyses of the Commercial Transaction frame(s) and the Risk frame(s) suggest that exploiting a rich decomposition of event structure will allow us to put together frames in FrameNet that have been split essentially because of differences in event structure, and hence argument realization, in the verbs in the more general frame. In the Commercial Transaction frame, different verbs trace different paths through the network of interactions among the participants or frame elements, leading to different verbs appearing in different argument structure constructions. In the Risk frame, there appears to be a single causal chain, but alternative grammatical realization of either a participant (nominal) or the participant’s subevent (action nominal), and the possibility of one participant playing two roles (Actor and Valued Object), again lead to a variety of argument structure constructions.

A decompositional analysis of event structure allows us to eliminate some frame-to-frame relations and simplify the lattice of frames in FrameNet. Of course, it would require a major effort to introduce event structure analysis into FrameNet and determine just how much of the FrameNet lattice can be simplified. Ideally, the majority of frame-to-frame relations remaining in FrameNet would be inheritance relations.

Even so, we suspect that what a frame would inherit would be primarily the basic semantic structure of participants/frame elements, their subevents, and the network of their interactions. Different predicates will profile different parts of the network of interactions; this cannot be predicted merely from frame membership. It is unlikely that the event profiles, as manifested in argument structure constructions, will all be inherited from the parent frame by the child frames. Nevertheless, there are certainly generalizations about which argument structure constructions and hence which event structures occur in particular frames. And the richer representation of event structure should allow for additional types of inferences about the events in the frames.

**Acknowledgements**

This research was partly funded by grant number HDTRA1-15-0063 from the Defense Threat Reduction Agency to the first author.

**References**


