Frame Semantic Parsing Needs Constructions

Miriam R.L. Petruck
International Computer Science Institute
miriamp@icis.berkeley.edu

Lori Levin
Carnegie Mellon University
Levin@andrew.cmu.edu

Keywords: frames, constructions, frame semantic parsing, natural language processing

Gildea and Jurafsky's (2002) pioneering work, based on early FrameNet data (Johnson et al., 2002), paved the way for the development of semantic role labeling (SRL) systems, some of which have reached a relatively high level of performance (e.g., Swayamdipta et al., 2018; Roth, 2016; Roth & Lapata, 2015, Das et al., 2014). Such systems have proven useful in Natural Language Processing (NLP) tasks, all of which first require frame semantic parsing (FSP, e.g., Liu et al., 2016**)**. Despite exploiting ever more sophisticated computational techniques, greater success in FSP is limited because, for the most part, SRL itself is limited to semantic role detection and labeling with respect to individual lexical items, or what FrameNet has construed as such (e.g., *throw on*, *throw up*, *throw together*) as opposed to grammatical constructions.

While cognitive linguists working in Frame Semantics (e.g., Fillmore 1985, 1986) and Construction Grammar (e.g., Fillmore 1988, Kay and Fillmore 1999) are fully aware of the need to characterize all types of constructions, not just lexical ones, as FrameNet primarily has done (Ruppenhofer et al. 2016), so far the NLP research community has not developed enough computational infrastructure that would support the automatic recognition of constructions, but for notable exceptions (e.g., Bakshandeh & Allen 2015; Dunietz 2018).

This paper presents an analysis of a grammatically complex sentence (#1) to illustrate the kinds of linguistic material that a well articulated FrameNet-type lexicon and constructicon must be able to handle (Fillmore 2006).

1) Car sales were so much lower than expected that stock prices plummeted.

Example (#2) mimics FrameNet's so-called full text annotation, i.e., lexicographic annotation for each lexical item in (#1), with frame names appearing in subscript.

2) [Car $_{\text{Vehicle}}$] [sales $_{\text{Commerce\_sell}}$] were [so $_{\text{Sufficiency}}$] [much $_{\text{Degree}}$] [lower $_{\text{Measurable\_attributes}}$] than [expected $_{\text{Expectation}}$] that [stock $_{\text{Capital\_stock}}$] [prices $_{\text{Commerce\_scenario}}$] [plummeted $_{\text{Change\_position\_on\_a\_scale}}$]

Examples (#3) – (#5) list the constructions that (#1) instantiates: Comparative_inequality, Causation, and Sufficiency, respectively; the examples also indicate the construction evoking elements (CEEs), if identified, and the construction elements (CEs) for each.

(3) **Comparative_ equality**: [Car sales $_{\text{Item}}$] were [so much $_{\text{Degree}}$] [lower $_{\text{CD\_base CD\_marker CEE}}$] [than expected $_{\text{Standard}}$] that stock prices plummeted.
(4) **Causation**: [Car sales were so much lower than expected $_{\text{Cause}}$] [that stock prices plummeted $_{\text{Effect}}$]
(5) **Sufficiency:** Car sales were [so much lower than expected $_{\text{Sufficient\_proposition}}$] [that stock prices plummeted $_{\text{Consequence}}$]

Note that the CEE in (#5) is discontinuous, given that *so* and the finite verb *plummeted* in the *that*-clause evoke the Sufficiency construction. Dunietz et al. (2017) captured numerous syntactic configurations that express causation, despite their linguistic realizations not always including a specific lexical item that would evoke a causation frame. For example, *Car sales were so much lower than expected* is the cause of the effect *that stock prices plummeted* in (#4), where the construction itself provides the causative interpretation; note the Cause and Effect CEs labeled on (#4). While humans can infer such causation (#4), machines cannot do so without the constructional information that the present analysis provides. In other words, frame semantic parsing needs constructions.

This work explicitly brings NLP into Cognitive Linguistics (Dunn 2018), and in so doing strives to invoke cognitive aspects of language processing. Also, it begins to suggest that articulating lexicon and constructicon might best be accomplished by attending to the concerns of NLP, specifically modeling frames and constructions in ways that NLP research and development will find useful.