

# Choosing an Event Description: What a PropBank Study Reveals about the Contrast between Light Verb Constructions and Counterpart Synthetic Verbs

Claire Bonial

Army Research Laboratory, Adelphi, MD

Light Verb Constructions (LVCs) in English and Romance languages are somewhat unique cross-linguistically because they tend to have semantically similar synthetic verb counterparts (e.g., Zarco, 1999). For example (from the *Corpus of Contemporary American English* (Davies, 2008)):

1. She **appeared** with me on VH1 "Celebrity Rehab."
2. This afternoon, Bahrain's King Hamad **made** a rare **appearance** on television.

This runs contrary to assumptions in linguistic theories that two competing forms are rarely maintained in a language unless they serve distinct purposes (Grice, 1975; Cattell, 1984; Goldberg, 1995). Why do English LVCs exist alongside counterpart synthetic verbs, especially given that synthetic verbs are arguably the more efficient variant form (Zipf, 1949)? This question has been difficult to study without a large-scale resource providing a markup of both LVCs and counterpart verbs. Such a resource has been difficult to create given that English LVCs can be surface-identical to corresponding full senses (Butt & Geuder, 2001; Tu & Roth, 2011), and even manual identification of LVCs can be challenging, as there are many borderline cases with verbs showing varying degrees of lightness. The present research describes the development and refinement of guidelines for the annotation of LVCs in the English PropBank (Palmer et al., 2005) and the subsequent analysis of approximately 2,000 LVC annotations and 10,000 counterpart synthetic verb annotations. The aim of the analysis was to discover evidence of what contexts call for the use of an LVC over a synthetic verb (annotations of 18,000 eventive and stative nouns found outside LVCs were also examined for comparison). The corpus study shows that LVCs are modified significantly more often than counterpart lexical verbs: there are, on average, 1.2 modifiers per LVC and 0.6 modifiers per counterpart lexical verb. Further annotation of the semantic function of the modifiers demonstrates that there is also greater variety to the types of modifiers seen with LVCs than counterpart synthetic verbs: nouns are compatible with descriptive elements, such as quantifiers and relative clauses, that can only be expressed periphrastically with synthetic verbs. LVCs lacking a clearly related verb (e.g., *give an overview*, *make an effort*) were also considered, and found to have equally high levels of modification when compared to LVCs with verb counterparts. Finally, it is also noted that although LVC modification includes certain types of determiners that allow speakers to modulate event aspect, no consistent relationship

between the aspect of a synthetic verb and that of its LVC counterpart was found. Thus, this corpus study provides distributional evidence that the ease and variety with which LVCs can be modified, in order to provide nuanced and detailed descriptions of events, is the primary motivating factor for the use of an LVC instead of a counterpart synthetic verb.