Précis of

TAKE A LOOK AT THIS! FORM, FUNCTION, AND PRODUCTIVITY OF ENGLISH LIGHT VERB CONSTRUCTIONS

CLAIRE BONIAL

U.S. Army DEVCOM Army Research Laboratory

English light verb constructions (LVCs), such as *make an offer* and *take a bath*, are semi-productive constructions: while some novel combinations of a light verb and eventive or stative noun are acceptable, others are not. LVCs tend to occur in families defined by a shared light verb and semantically similar nominal complements, but it remains mysterious precisely how such families are circumscribed; this presents both theoretical and natural language processing (NLP) challenges. This research first aims to address the void in linguistic resources identifying LVCs in a consistent fashion with the development of annotation guidelines for LVCs within the PropBank project. Using the resulting annotated corpus of LVCs, another theoretically important question of why LVCs exist alongside semantically similar lexical verbs is addressed: corpus evidence demonstrates that the ease and variety with which LVCs can be modified is the primary motivating factor for their use over a lexical verb. Finally, large-scale acceptability studies are used to examine the constraints on LVC productivity; the results reveal the importance of statistical preemption in modeling productivity.

Keywords: light verbs, constructions, semantic roles, productivity, natural language processing

1. INTRODUCTION

A key question in linguistics and cognitive science is how people learn and apply the seemingly idiosyncratic constraints of a language. For example, in English, one can *tell me the facts* but cannot **explain me the facts* (Goldberg, 2011).[1] Similarly, one can *have a drink,* but cannot **have an eat* (Wierzbicka, 1982). How do speakers know this? From a usage-based, Construction Grammar perspective (e.g., Goldberg, 1995, 2006), the crux of this issue is why certain CONSTRUCTIONS (pairings of form and meaning) are compatible, or incompatible, with certain lexical items. Specifically of interest for this research, why is *drink* compatible within the *have* LIGHT VERB CONSTRUCTION (LVC) (Jespersen, 1942) while *eat* is not?

The answer to this question carries ramifications for larger issues of what is stored in the mental lexicon, and how human grammar develops. Some argue that Construction Grammar makes certain claims on the storage and processing of lexical items: phrasal constructions, as pairings of form and meaning, are stored in the mental lexicon in the same way that individual lexical items

are stored in the lexicon. This suggests that constructions are not decomposed or analyzed **compositionally** – according to the semantics of each individual lexical item (Piñango, Mack & Jackendoff, 2006; Wittenberg & Piñango, 2011). A related perspective from another usage-based approach, Emergent Grammar (e.g., Hopper, 1998), predicts that constructions are extended by semantic analogy to an existing, high-frequency exemplar construction. The validity of these views is explored here with respect to LVCs.

The questions surrounding LVC acceptability also present challenges for Natural Language Processing (NLP) systems. In general, NLP systems rely on a combination of training data, which captures usage patterns of lexical items, as well as computer-readable lexicons, which capture some level of the meaning of a word. Complete coverage of LVCs in a lexicon is made impossible by the fact that LVCs are SEMI-PRODUCTIVE: speakers can extend a construction's template creatively with novel combinations of lexical items. However, that productivity is constrained such that not all combinations are acceptable. As a result, we simply cannot list every LVC in a lexicon. Furthermore, novel, creative usages will be quite rare in training data, and will be greatly outnumbered by more conventional usages of the same verbs, precluding fully unsupervised approaches to this problem. Nevertheless, we can supplement NLP systems with knowledge from linguistics and psycholinguistics in order to model patterns of productivity and better estimate likelihoods that a given combination will be acceptable.

2. SUMMARY

A thorough treatment of these issues has been fettered by a lack of resources identifying LVCs in a consistent fashion. Thus, this research first aims to address the void in linguistic resources with the development of annotation guidelines for LVCs within the PropBank project (Palmer et al., 2005) (§4). Using the resulting corpus of LVCs, another theoretically important question of why LVCs exist alongside semantically similar lexical verbs is addressed: corpus evidence demonstrates that the ease and variety with which LVCs can be modified is the primary motivating factor for their use over a lexical verb (§5.1). Finally, large-scale acceptability studies are used to examine the constraints on LVC productivity; the results reveal the importance of statistical preemption in modeling productivity (§5.2). The impacts of this research on (psycho)linguistics and NLP are discussed in closing (§6).

3. BACKGROUND: LIGHT VERB CONSTRUCTIONS

English LVCs are an ideal case for studying integral issues of grammar as LVCs are SEMI-COMPOSITIONAL, SEMI-PRODUCTIVE constructions that tend to have semantically similar lexical verb counterparts (e.g., *make an offer* vs. *offer*) – a fact which runs contrary to assumptions in linguistic theories that two competing forms are rarely maintained in a language, unless they serve distinct purposes (Grice, 1975). Each of these characteristics is explained in more detail below.

The theories of Construction Grammar posit that there is a continuum from wholly fixed, non-compositional idioms like *kick the bucket*, in which no component part expresses the meaning of *die*; to SEMI-COMPOSITIONAL constructions like LVCs, in which the component parts do contribute lexical meaning, and there is some flexibility as to what elements can form an LVC; to purely compositional language, in which words combine freely and productively according to syntactic rules. This continuum is illustrated in Figure 1.

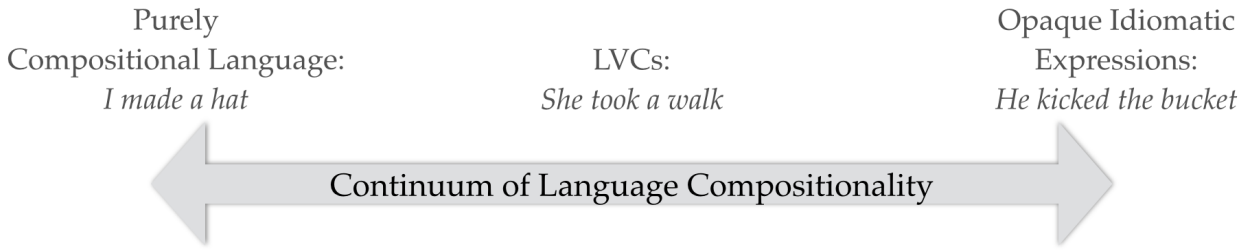

Figure 1: Continuum of language compositionality.

As semi-compositional constructions, LVCs cannot be interpreted in a completely literal fashion (*take a seat* and *take a bath* are not interpreted as events where chairs and bathtubs are taken somewhere, but instead as *sitting* and *bathing* events). Speakers must recognize the unique nature of these constructions and interpret them according to the event semantics denoted by the noun. Furthermore, our computational systems must do this: both natural language understanding and machine translation (among other computational tasks) require that LVCs are delineated from other usages of the same verbs and interpreted uniquely.

Parallel to the continuum of compositionality, there is a continuum of productivity. On one extreme, purely compositional language is thought to be fully productive. On the other, some idiomatic constructions are wholly fixed, allowing for no substitution of lexical items (e.g., **He punted the bucket*). LVCs are SEMI-PRODUCTIVE (Nickel, 1978): while some novel combinations

of a LV and noun are acceptable, others are not. LVCs tend to occur in families defined by a shared LV and semantically similar nominal complements, but it remains mysterious precisely how such families are circumscribed. For example, a variety of nouns denoting communication events can combine with *make*:

(1) make a speech/declaration/proclamation/announcement

Yet other semantically similar nouns may not form acceptable combinations:

(2) ?make a yell

While some research has tried to pinpoint the exact semantic constraints that make a given combination acceptable (e.g., Wierzbicka, 1982), other research has focused on the importance of frequency and analogy in extending particular constructions. Emergent Grammarians (e.g., Hopper, 1998; Bybee, 2006, 2010) hypothesize that semi-productive constructions are extended by semantic analogy to an existing, high-frequency exemplar construction; therefore, novel constructions that are semantically very similar to high-frequency exemplars are predicted to be acceptable.

LVCs in English and Romance languages are somewhat unique cross-linguistically because they tend to have semantically similar synthetic verb counterparts (Zarco, 1999). For example (from *COCA* (Davies, 2008)):

(3) She **appeared** with me on VH1 "Celebrity Rehab."
(4) 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 (e.g., Grice, 1975). What distinct purposes might these forms serve? Why do English LVCs exist alongside counterpart synthetic verbs, especially given that synthetic verbs are arguably the more efficient variant form (Zipf, 1949)?

Unfortunately, these questions and the issues surrounding LVC usage and productivity have been difficult to address without a large-scale resource providing a markup of both LVCs and counterpart verbs. Thus, this research first undertakes the challenge of developing a resource that provides comprehensive annotations of LVCs.

4. APPROACH: DEVELOPMENT OF THE PROPBANK CORPUS OF LVC ANNOTATIONS

The goal of PropBank is to supply consistent, general-purpose labeling of semantic roles across different syntactic realizations. With over two million words from diverse genres, the annotated corpus supports the training of automatic semantic role labelers, which in turn support other NLP areas, such as summarization and question-answering. PropBank annotation consists of two tasks: sense and role annotation. The PropBank lexicon provides a listing of the coarse-grained senses of a verb, noun, or adjective relation, and the roles associated with each sense (known as a ROLESET). The roles are listed as argument numbers corresponding to relation-specific roles. For example:

Offer-01[2]
Arg0: *entity offering*
Arg1: *commodity*
Arg2: *price*

Previous versions of PropBank maintained separate rolesets for the verb *offer*, and the related nouns *offer/offering*, but, as part of this research, these have been combined to provide parallel annotations of related usages (Bonial et al., 2014):

(5) [NE Electric]$_{ARG0}$ **offered** [$2 billion]$_{ARG2}$ [to acquire PS]$_{ARG1}$
(6) [NE Electric]$_{ARG0}$ **made an offer** of [$2 billion]$_{ARG2}$ [to acquire PS]$_{ARG1}$

Prior to this research (see also Hwang et al., 2010), PropBank 1.0 had no special guidelines for LVCs. Therefore, example 6 would have been annotated according to the semantic roles found in the verb roleset, make-01, which is, at best, metaphorically related to the LVC usage (see 9 for a discussion of annotation with make-01). Furthermore, this practice conflated LVC usages like

make an offer with HEAVY, literal usages of the same verb, such as *make a cake*, in which *make* is used in its full, *creation* sense.

LVCs require a unique annotation procedure that represents the event semantics of the construction stemming from the noun, as opposed to the verb. First, however, we must recognize LVCs consistently. This has been problematic, given that definitions of LVCs cross-linguistically and within a language remain nebulous. Nonetheless, LVCs in English are generally defined as consisting of a semantically general, highly polysemous verb and a noun denoting an event or state (e.g., Butt, 2003). More detailed definitional aspects of LVCs remain debatable.

This basic definition does not allow for consistent manual, much less automatic, detection of LVCs since surface-identical forms can be LVCs or heavy usages of the same verbs:

(7) He **took a drink** of the soda. (LVC) = He drank the soda.
(8) He **took** a drink off the bar. (non-LVC) ≠ He drank the bar.

Such usages are indistinguishable with respect to their syntactic constituents (Butt & Geuder, 2001); thus, neither annotators nor automatic systems can rely on syntactic criteria to identify LVCs and must use semantic criteria instead.

The PropBank guidelines for LVC annotation developed as part of this research therefore focus on the semantic nature of arguments (Bonial & Palmer, 2016).[3] This helped to ground the guidelines in theoretical research on English LVCs, which generally assumes that the semantic content of the construction stems from the noun, while the verb provides the syntactic scaffolding for the noun to act as the main predicate (Butt, 2003; Grimshaw & Mester, 1988). Accordingly, the arguments, including the syntactic subject of the verb, will carry the semantic roles of the noun. Thus, for LVC annotation, the annotators are instructed to compare the fit of the semantic roles listed for the verb's roleset to that of the noun's roleset. Figure 2 gives an overview LVC designation heuristics.

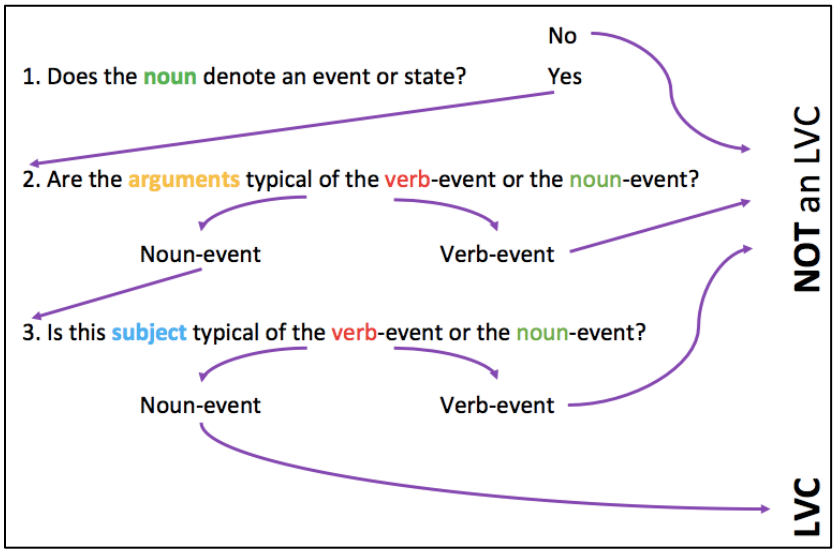

Figure 2: Flow chart for determining LVC status.

Consider the earlier sentence from example (6), involving *make_offer*. Annotators would examine the rolesets to decide which is more fitting:

Make-01

Arg0: *creator*

Arg1: *creation*

Arg2: *created-from*

Selecting the above make-01 roleset requires designation of the following roles:

(9) [NE Electric]$_{ARG0}$ **made** [an offer of $2 billion]$_{ARG1}$ [to acquire PS]$_{ARG-???}$

Offer-01

Arg0: *entity offering*

Arg1: *commodity*

Arg2: *price*

While selecting the above offer-01 roleset requires designation of the roles shown below:

(10) [NE Electric]$_{ARG0}$ made an **offer** of [$2 billion]$_{ARG2}$ [to acquire PS]$_{ARG1}$

The verb-centric annotation in 9 does not capture the semantics of the *offer* event, instead focusing on *make* as if it were a heavy verb. Thus, an *argument* lacks an appropriate tag in the roleset (indicated by '???'). The noun roleset offer-01 provides the appropriate roles to label each argument and allows for the Arg2 *price* to be captured, as seen in annotation 10. These considerations are evidence that the semantic roles stem from the noun relation and the usage should be annotated as an LVC.

The success of these guidelines has been demonstrated in the high agreement rates between annotators.[4] On a task composed solely of the most likely LVs (*give, have, take, make, do*), the agreement rate between two seasoned annotators was 93.8%.[5] As proof of the quality of PropBank's annotations, an automatic LVC detection system has been trained on the OntoNotes 4.99 corpus (Weischedel et al., 2011), of which PropBank is one layer. This system achieves an F-Score of 89% (Chen, Bonial & Palmer, 2015) when tested on the same data used by Tu and Roth (2011) to establish the previous state-of-the-art system with an F-Score of 86.3%.[6] When tested on the more realistic and challenging OntoNotes corpus containing 1,768 LVC instances, the system achieves an F-Score of 80.7% (see system implementation and evaluation details in (Chen, Bonial & Palmer, 2015)).

5. CORPUS & ACCEPTABILITY STUDIES

5.1 CORPUS STUDY: WHY DO LVCS EXIST ALONGSIDE SEMANTICALLY SIMILAR LEXICAL VERBS?

In addition to providing valuable training data, the PropBank corpus provides crucial information for investigating why English LVCs exist in the language alongside semantically similar counterpart lexical verbs. This question is important for an understanding of language use more generally, since the maintenance of both forms runs contrary to assumptions that two competing forms are rarely maintained unless they serve distinct purposes. The PropBank corpus was used to analyze approximately 2,000 LVC annotations and 10,000 counterpart synthetic verb annotations. The aim was to discover evidence of what contexts call for the use of an LVC over a lexical verb.

CORPUS STUDY SUMMARY

PropBank annotations for those LVCs with counterpart lexical verbs (e.g., *give a presentation*), including all semantic roles and modifier arguments (e.g., TEMPORAL, LOCATIVE, MANNER), were tabulated and compared to tabulations of roles and modifier arguments for counterpart lexical verbs (e.g., *present*).

CORPUS STUDY RESULTS

The corpus study shows that LVCs are associated with significantly more modifiers ($M = 1.17$, $SD = 0.45$) than verb counterparts (($M = 0.60$, $SD = 0.23$), $t(36) = 4.93$, $p < .001$). Data were checked for normality and homogeneity of variance, and met the required statistical assumptions. Due to the diverse and often idiomatic usage of these words, they were deemed to be independent, rather than paired, samples. However, using a paired-samples comparison yielded the same results, with LVCs having significantly more modifiers than their specific verb counterparts ($t(18) = 5.60$, $p < .001$). Chart 1 shows the average number of modifiers across the two predicate types.

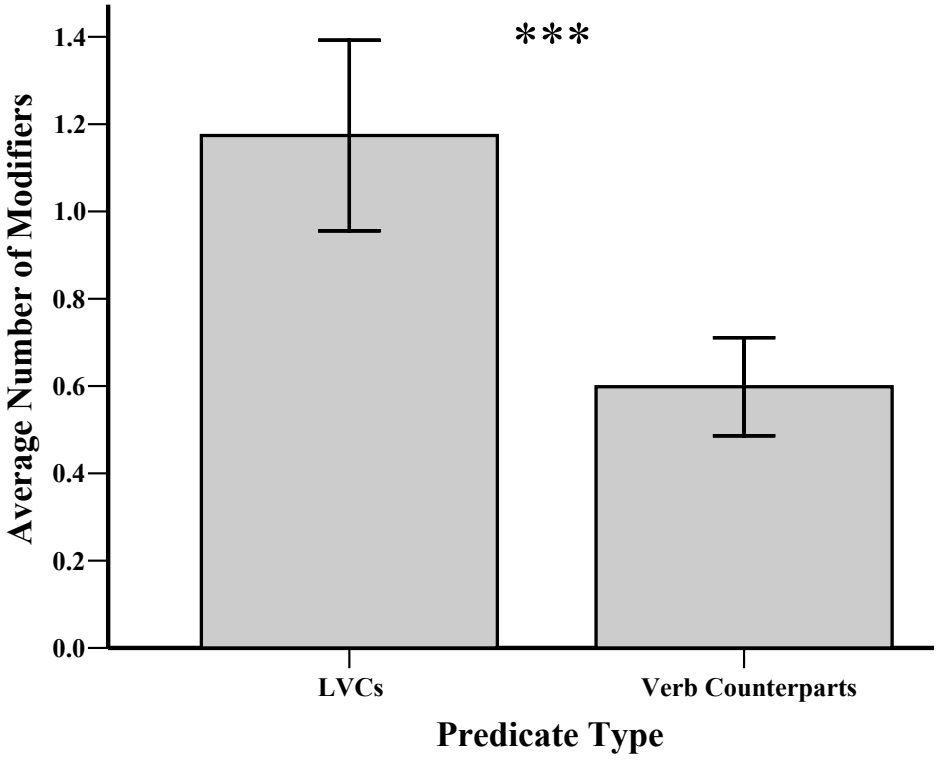

Chart 1: Number of modifiers (mean ± 95% CI) associated with LVCs and verb counterparts.
*** = p < 0.001.

There is also greater variety to the types of modifiers seen with LVCs than counterpart synthetic verbs: nouns are compatible with descriptive elements that can only be expressed periphrastically with synthetic verbs. This can be observed in the following usage example of an LVC with evaluative modification, followed by an invented example that attempts a reformulation using the corresponding verb:

(11) We had a **really good** laugh…
(12) ?We laughed **really well**…

Example 12 shows that the transformation of an evaluative nominal modifier into an adverbial modifier changes the meaning into what will likely be interpreted as a somewhat awkward and ambiguous manner modifier. In general, speakers can exploit LVCs to express detailed descriptions of an event using nominal modification, which is more flexible than verbal modification.

Thus, this corpus study provided 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 their use. Since the lexical verb is not necessarily a viable alternative in the denotation of certain detailed events, the two forms are not competing in these cases, and therefore both forms are maintained. This study also reveals a characteristic of LVCs important for their automatic detection: they consistently include some modification of the noun. This characteristic is overlooked in NLP, as many LVC detection approaches operate under the assumption that only a determiner will occur between the verb and noun (e.g., Stevenson et al., 2004).

5.2 ACCEPTABILITY STUDY: WHAT ARE THE CONSTRAINTS ON LVC PRODUCTIVITY?

Although the PropBank corpus was instrumental in providing training data for the detection of most LVCs, infrequent and novel LVCs unseen in the training data remain a challenge. One could assume that the same LV can combine with any member of a family of semantically related nouns, using information on semantic similarity from WordNet (Fellbaum, 1998) or FrameNet (Fillmore et al., 2002) to establish these families. However, this approach would over-generate possibilities: LVCs are not fully productive, given that some combinations are simply unacceptable. Thus, this research aims to establish why certain combinations are acceptable while others are not, and to

frame this knowledge in a way that can be usefully injected into NLP systems reliant on statistical patterns.

Emergent Grammar predicts that novel constructions are extended by analogy to high-frequency, existing constructions, and this has been demonstrated by Bybee and Eddington (2006) in their research on Spanish *becoming* constructions. The authors conducted acceptability studies of somewhat synonymous verbs, roughly meaning *become,* combined with different adjectives. They examined the acceptability of both attested and unattested combinations. It was found that speakers judged low-frequency or unattested combinations to be acceptable only if they were semantically very similar to an existing, high-frequency combination.

Intuitively, it seems plausible that other semi-productive constructions, like LVCs, would also be extended in this fashion. Thus, the hypothesis was adapted to examine its validity with respect to English LVCs:

FREQUENCY HYPOTHESIS: Speakers will find novel or very low-frequency LVCs acceptable if they are semantically similar to an attested, highly frequent LVC.

In accordance with the findings of Bybee & Eddington (2006), the expectation is that speakers will find very low-frequency LVCs that are similar to a strongly entrenched, high-frequency exemplar to be more acceptable than those that are similar to a weaker, low-frequency exemplar.

To test this hypothesis, it was necessary to establish what families of semantically similar LVCs exist, and what the frequency is of each unique LVC within a family. Although PropBank LVC annotations were used to determine common LVs of focus, the corpus was not large enough to gain a full picture of LVC usage in English. Thus, occurrences of the frequent LVs *give, have, make* and *take* were collected from the Gigaword corpus of over 1.7 billion words (Graff & Cieri, 2003). FrameNet membership was used to determine what LVCs were semantically similar. After computational analysis of frequencies and some manual filtering of false positives, a fairly comprehensive picture of families of LVCs emerged, including very low-frequency LVCs to be tested for acceptability that had either a low-frequency exemplar within their family, or a high-frequency exemplar. "Very low-frequency" LVCs occur 20-50 times in Gigaword, "low-frequency" LVCs occur 100-200 times, "high-frequency LVCs occur 2000-3000 times. These frequency bands were decided upon after careful analysis of histograms displaying how many

unique collocations occur at which frequencies across Gigaword. Examples of the two types of LVCs compared for acceptability here—very low-frequency LVCs with only a low-frequency exemplar in the semantically similar family, and very low-frequency LVCs with only a high-frequency exemplar in the semantically similar family—are shown in Tables 1 and 2.

LVC	**Frequency**
make realization	5
make inference - very low-frequency test LVC	20
make deduction	24
make guess - low-frequency exemplar LVC	102

Table 1: A low-frequency family of LVCs detected in Gigaword: most tokens fall into the low-frequency band of 100-200 instances. The nouns of the LVCs in these families share a FrameNet frame.

LVC	**Frequency**
have dread - very low-frequency test LVC	24
have terror	33
have apprehension	107
have fear - high-frequency exemplar LVC	2342

Table 2: A high-frequency family of LVCs detected in Gigaword: most tokens fall into the high-frequency band of 2,000-3,000 instances.

The very low-frequency test LVCs were presented to 125 participants on Amazon's Mechanical Turk in the context of a sentence exemplifying that LVC taken from Gigaword. Participants provided a rating for the rare LVC by clicking in a box with just two endpoints marked as "Odd," on one end, or "Perfectly fine." The survey interface is shown in Figure 3. Note that participants are never presented with low or high-frequency exemplar LVCs, they are solely judging the

acceptability of the very low-frequency test LVCs, so they are in no way primed to consider similarity to any existing, higher-frequency LVC.

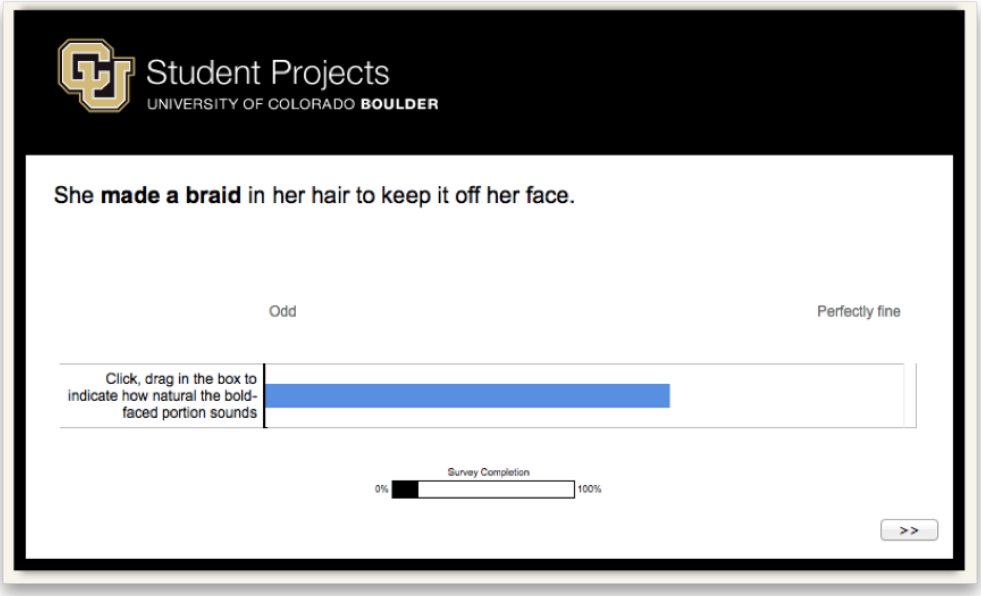

Figure 3: Question interface viewed by participants.

ACCEPTABILITY STUDY SUMMARY

Native English speakers judge the acceptability of very low-frequency LVCs that are either semantically similar to a low-frequency or high-frequency exemplar, and the levels of acceptability will be compared across the two types.

ACCEPTABILITY STUDY RESULTS

It was found that the frequency of the exemplar LVC does significantly correlate with the acceptability of the very low-frequency test LVC, but in a manner opposite of what was predicted: very low-frequency LVCs that are semantically similar to low-frequency LVC exemplars are significantly more acceptable than very low-frequency LVCs that are semantically similar to high-frequency exemplars ($X^2(1)=5.64$, $p=0.01751$).[7] The most plausible explanations for this surprising result were explored: statistical preemption and semantic bleaching.

Statistical preemption is a cognitive process in which speakers implicitly infer from consistently hearing a formulation, B, in a context where one might have heard a semantically

related alternative formulation, A, that B is the appropriate formulation and A is not appropriate (e.g., Suttle & Goldberg, 2011). Here, the highly frequent exemplar LVC blocks the use of the semantically similar, very low-frequency LVC. Thus, when speakers judge the acceptability of a relatively unfamiliar, very low-frequency LVC (e.g., *give a stare*) that is clearly quite similar to a very familiar, high-frequency construction (e.g., *give a look*), the unfamiliar LVC seems odd, given that the familiar construction could have been used. On the other hand, when speakers judge the acceptability of a very low-frequency LVC (e.g., *take a departure*) that is similar to a low-frequency construction (e.g., *take an exit*), the low-frequency exemplar is likely familiar enough that it can serve as the basis of analogical extension of the family to the new, unfamiliar member, but the exemplar isn't so frequent that it blocks the unfamiliar alternative.

Semantic bleaching is a process in which a high-frequency phrase or construction has become entrenched to the point where speakers no longer analyze the individual components of the expression, instead treating it as an unanalyzed whole (Hopper & Traugott, 1993). When speakers judge the acceptability of an unfamiliar LVC, they are comparing it to previously experienced constructions for similarity, in order to decide if the unfamiliar construction should be added to an existing family of constructions. However, the relationship of semantic similarity between the unfamiliar construction (e.g, *have dread*) and the familiar construction (e.g., *have fear*) is not clear after semantic bleaching has taken place because speakers no longer analyze the individual components of the high-frequency expression, precluding recognition of the semantic similarity between those components (e.g., *dread* and *fear*). Therefore, the unfamiliar construction is not accepted as a member of the existing family.

The first interpretation involving statistical preemption seems more likely, given trends in this data and a survey of related work on both statistical preemption and semantic bleaching. Lending further evidence for this interpretation, it was found that very low-frequency LVCs that were from semantically similar families with more than one high-frequency member were significantly less acceptable than those that came from a family with only one higher-frequency member ($X^2(1)=4.84$, $p=0.02775$). Essentially, in families with more than one high-frequency member, there are two entrenched formulations that can block the less frequent alternative formulation. Chart 2 exemplifies such a family.

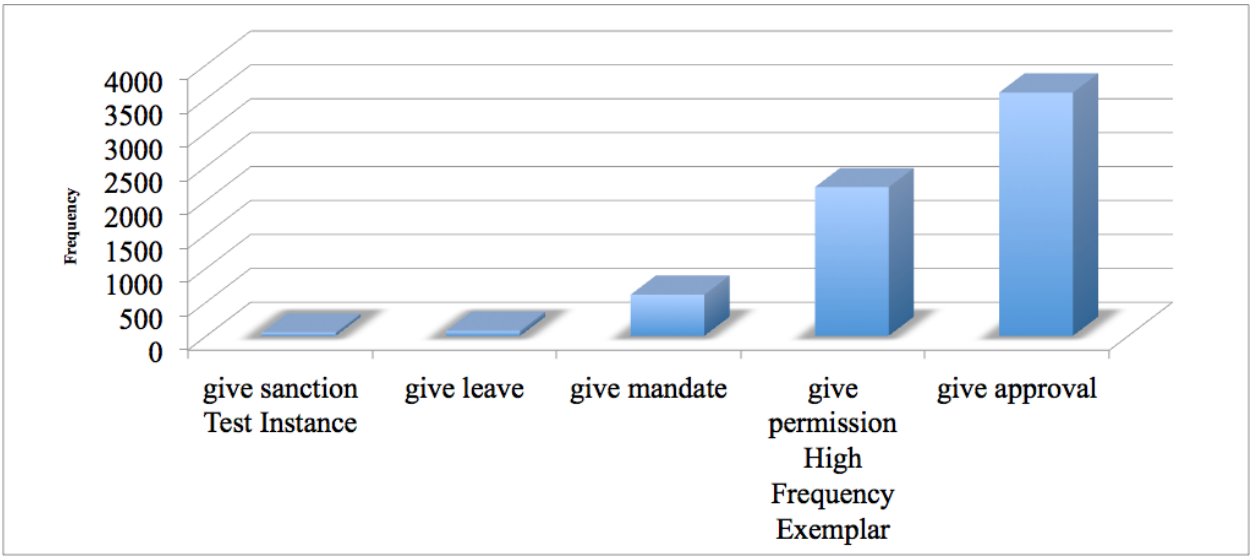

Chart 2: Family with two high-frequency LVCs, in which both *give permission* and *give approval* block the use of the rare variant *give sanction.*

Furthermore, while the process of statistical preemption has been demonstrated quite convincingly in experimental settings (e.g., Boyd, Ackerman & Kutas, 2012), the process of semantic bleaching has not. In fact, although the process of semantic bleaching seems very plausible from a diachronic perspective, a growing body of research in psycholinguistics and cognitive science has demonstrated that speakers do recognize and analyze the individual lexical items within LVCs and idiomatic expressions (e.g., Cutting & Bock, 1997; Wittenberg & Piñango, 2011), contrary to semantic bleaching predictions.

Thus, this research demonstrates that even relatively low-frequency constructions can serve as the basis of analogical extension, which evidences the Emergent Grammar assumption that speakers retain details of each token of linguistic experience. Furthermore, there is a certain point where the higher frequency of a semantically similar form may block another variant, rather than encouraging analogical extension, thereby providing evidence for the somewhat overlooked role of statistical preemption in the process of extending constructions. To take advantage of these findings in NLP, we can use the frequency signatures of LVC families to predict which families will block or allow new members, effectively modeling patterns of LVC productivity and acceptability.

6. CONCLUSIONS, INTERDISCIPLINARY CONTRIBUTIONS

This work draws upon research strands from linguistics, cognitive science, and computer science to make several theoretical and practical contributions. First, the development of a valuable resource of annotated English LVCs allows for continued linguistic research on these constructions and provides training data for the automatic detection of LVCs.[8] Next, this research provides quantitative evidence demonstrating why two seemingly competing forms, LVCs and counterpart lexical verbs, are maintained in the language. Finally, this work provides evidence for the role of statistical preemption in extending semi-productive constructions. This informs our picture of how grammar is constructed as a whole, and provides a computational framework for estimating the likelihoods of productivity and acceptability. This information will be instrumental in the automatic detection of novel LVCs, unseen in training data. As a whole, this research has demonstrated the promising ways in which the various disciplines of (psycho)linguistics, cognitive science and computer science can be brought together to identify a linguistic phenomenon, understand its function in the language, and begin to understand how humans can, and perhaps computers should, process and extend the phenomenon.

REFERENCES

Bates D, M. Maechler, B. Bolker and S. Walker. 2014. *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-7, http://CRAN.R-project.org/package=lme4.

Bonial, Claire & Martha Palmer. 2016. Comprehensive and Consistent PropBank Light Verb Annotation. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pp. 3980-3985.

Bonial, Claire, Julia Bonn, Kathryn Conger, Jena D. Hwang & Martha Palmer. 2014. PropBank: Semantics of New Predicate Types. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14),* pp. 3013-3019.

Boyd, Jeremy K., Farrell Ackerman and Marta Kutas. 2012. Adult learners use both entrenchment and preemption to infer grammatical constraints. In *Proceedings of the 2012 IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL).*

Butt, Miriam. 2003. The Light Verb Jungle. In G. Aygen, C. Bowern & C. Quinn (Eds.) *Papers from the GSAS/Dudley House Workshop on Light Verbs*. Cambridge, Harvard Working Papers in Linguistics, p. 1-50.

Butt, Miriam, and Wilhelm Geuder. 2001. On the (Semi)Lexical Status of Light Verbs. In *Semi-Lexical Categories: On the content of function words and the function of content words*, ed. Norbert Corver and Henk van Riemsdijk. 323-370. Berlin: Mouton De Gruyter.

Bybee, Joan. 2006. From usage to grammar: the mind's response to repetition. *Language* 82(4): 711-733.

Bybee, Joan. 2010. Language, Usage and Cognition. Cambridge: Cambridge University Press.

Bybee, Joan, and D. Eddington. 2006. A usage-based approach to Spanish verbs of 'becoming'. *Languag*e, 82(2), 323–355.

Cattel, Ray. 1984. *Syntax and Semantics: Composite Predicates in English.* Orlando: Academic Press, Inc.

Chen, Wei-Te, Claire Bonial & Martha Palmer. 2015. English Light Verb Construction Identification Using Lexical Knowledge. In Proceedings of *Twenty-Ninth Conference on Artificial Intelligence (AAAI-15)*.

Cutting, J. Cooper and Kathryn Bock. 1997. That's the way the cookie bounces: Syntactic and semantic components of experimentally elicited idiom blends. *Memory & Cognition,* 25(1): 57-71.

Davies, M. 2008-. The Corpus of Contemporary American English (COCA): 425 million words, 1990-present. Available online at http://www.americancorpus.org.

Fellbaum, C. (Ed.) 1998. *WordNet: An Electronic Lexical Database.* MIT Press.

Fillmore, Charles J., Christopher R. Johnson, and Miriam R.L. Petruck. 2002. Background to FrameNet. *International Journal of Lexicography*, 16(3):235-250.

Goldberg, Adele. 1995. *Constructions: A Construction Grammar Approach to Argument Structure.* Chicago: University of Chicago Press.

Goldberg, Adele. 2006. *Constructions at Work: The Nature of Generalization in Language.* New York: Oxford University Press.

Goldberg, Adele. 2011. Corpus evidence of the viability of statistical preemption. *Cognitive Linguistics* 22 1: 131-154.

Graff, David and Christopher Cieri. 2003. *English Gigaword.* Linguistic Data Consortium, Philadelphia.

Grice, H.P. 1975. Logic and Conversation. In P. Cole and J. Morgan (Eds.) *Syntax and Semantics*, vol.3. Academic Press. Reprinted as ch.2 of Grice 1989, 22–40.

Grimshaw, J., and A. Mester. 1988. Light verbs and θ-marking. *Linguistic Inquiry,* 19(2):205–232.

Hopper, Paul J. 1998. Emergent Grammar. In *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, Michael Tomasello (ed.), 155-75. Mahwah, NJ/London: Lawrence Erlbaum Associates.

Hopper, Paul J. and Elizabeth C. Traugott. 1993. *Grammaticalization (=Cambridge Textbooks in Linguistics.)* Cambridge: Cambridge University Press.

Hwang, Jena D, Archna Bhatia, Claire Bonial, Aous Mansouri, Ashwini Vaidya, Nianwen Xue, and Martha Palmer. 2010. PropBank Annotation of Multilingual Light Verb Constructions. In Proceedings of the Linguistic Annotation Workshop held in conjunction with ACL-2010. Uppsala, Sweden, July 15-16, 2010.

Jespersen, Otto. 1942. *A Modern English Grammar on Historical Principles.* Part VI: *Morphology*. With assistance of P. Christopherson, N. Haislund, & K. Schibsbye. London: Goerge Allen & Unwin. Copenhagen: Ejnar Munksgaard.

Nickel, Gerhard. 1978. Complex Verbal Structures in English. *Studies in Descriptive English Grammar,* 63-83.

Palmer, Martha, Daniel Gildea, and Paul Kingsbury. 2005. The Proposition Bank: An annotated corpus of semantic roles. *Computational Linguistics*, 31(1):71– 106.

Passonneau, Rebecca. 2004. Computing reliability for coreference annotation.

Piñango, Maria M, Jennifer Mack and Ray Jackendoff. 2006. Semantic combinatorial Processes in Argument Structure: Evidence from Light-Verbs. In *Proceedings of the Thirty-second Annual Meeting of the Berkeley Linguistics Society.*

Postal, Paul.M. 1974. *On raising: One rule of English grammar and its theoretical implications* (Vol. 5). The MIT Press.

R Development Core Team. 2011. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Suttle, Laura and Adele E. Goldberg. 2011. Partial productivity of constructions as induction *Linguistics* 49 6: 1237-1269.

Tu, Yuancheng, and Dan Roth. 2011. Learning English Light Verb Constructions: Contextual or Statistical. In *Proceedings of the Worshop on Multiword Expressions: from Parsing and Generation to the Real World (MWE 2011),* pp. 31-39.

Weischedel, R., E. Hovy, M. Palmer, M. Marcus, R. Belvin, S. Pradhan, L. Ramshaw and N. Xue. 2011. Ontonotes: A large training corpus for enhanced processing. In Olive, J.; Christianson, C.; and McCary, J., eds., *Handbook of Natural Language Processing and Machine Translation*. Springer.

Wierzbicka, Anna. 1982. Why can you Have a Drink when you can't *Have an Eat? *Language*, 58(4):753–799.

Wittenberg, Eva & Maria M. Piñango. 2011. Processing Light Verb Constructions. *The Mental Lexicon* 6:3, 393–413.

Zarco, M. A. 1999. Interlingual Representation of Complex Predicates in a Multilingual Approach: the Problem of Lexical Selection. Saint-Dizier, P. (Ed.) *Predicative Forms in Natural Language and In Lexical Knowledge Bases*. Dordecht: Kluwer Academic Publishers.

Zipf, George K. 1949. *Human Behavior and the Principle of Least Effort*. Addison-Wesley.

ENDNOTES

[1] Unacceptable usages are preceded by '*'; questionable usages by '?'.

[2] http://verbs.colorado.edu/propbank/framesets-english-aliases

[3] http://verbs.colorado.edu/propbank/EPB-Annotation-Guidelines.pdf

[4] A common measure of the quality and reliability of an annotation schema (e.g., Passoneau, 2004).

[5] http://verbs.colorado.edu/propbank/ita/webtext-p25-SEL-LightVerb.html

[6] F-score is the harmonic mean of a classification system's precision and recall.

[7] R (R Core Team, 2012) and *lme4* (Bates, Maechler & Bolker, 2012) were used to perform a linear mixed effects analysis of the relationship between the subject's naturalness rating of the very low-frequency LVCs and the frequency band of the exemplar LVC.

[8] The LVC annotations have been released as part of the "BOLT English PropBank and Sense -- Discussion Forum, SMS/Chat, and Conversational Telephone Speech" annotated corpus, LDC2020T21 (https://catalog.ldc.upenn.edu/LDC2020T21).