Why Verb Meaning Matters to Syntax

Eunkyung Yi & Jean-Pierre Koenig[1]

1 Introduction

Since Gruber (1965) and particularly Fillmore (1968), it is known that we can predict a lot of the syntax of verbs from their meaning. To quote Green (1974) or thirty years later Koenig and Davis (2006):

'Syntactic properties of lexical items depend crucially and in regular ways on the meaning of those lexical items' (Green 1974)
'It is widely accepted that the semantic content of a lexical entry determines to a large extent its syntactic subcategorization' (Koenig & Davis 2006)

There are at least two ways in which this relation between verb meaning and the syntactic contexts in which verbs can occur manifests itself. First, the syntactic pattern can carry meaning. For example, the English ditransitive construction with which this paper is concerned is often claimed to be associated with a transfer of possession meaning (see Pinker 1989, Goldberg 1993 among others, although see Van Valin & La Polla 1997 for a dissenting view). Some researchers associate the meaning directly with the syntactic pattern itself (see Goldberg 1995, and Ramchand 2008 for a very different tack on the association of a transfer of possession meaning with the pattern). Others associate the meaning with the output of a rule mapping a set of lexical entries onto another set of lexical entries (Pinker 1989 is a representative example of such a view). Either way, a question that arises is what kinds of verb meaning components are cross-linguistically relevant for syntactic purposes (see Pinker 1989, Grimshaw 1993, Koenig et al.

2015). But, another aspect of this association between verb meaning and syntactic frames, the one we concentrate on in this paper, is that verbs that tend to have similar meanings tend to occur in similar syntactic frames. Consider the three following syntactic frames and some of the verbs that can occur in these frames.

(1) a. John gave his son a toy.
    b. *give, offer, hand, lend, promise, tell, grant…*

(2) a. John loaded the truck with hay.
    b. *load, spray, smear, cram, engrave, mark…*

(3) a. John kept his son from entering the room.
    b. *keep, prevent, bar, prohibit, discourage, deter…*

Intuitively, the verbs in each of these syntactic frames have something in common semantically, at least more so than verbs across the syntactic frames exemplified in (1)–(3). Thus, *give* seems more similar to *promise* than to *load*; *spray* seems more similar to *smear* than to *keep*, and so forth. Not all verbs that can occur in a particular syntactic frame are equally similar to each other, though. For example, *load* is certainly not as similar, intuitively, to *smear* than *spray* is. But, overall, most verbs that occur in most syntactic frames seem to denote kinds of situations that are similar to each other, possibly through some kind of metaphorical similarity, as Goldberg (1995) claims. As Levin (1993: 11) puts it:

> 'Studies of diathesis alternations show that verbs in English and other languages fall into classes on the basis of shared components of meaning.'

The question that occupies us in this paper is why this is the case. What accounts for the fact that verbs that occur in the ditransitive construction are semantically similar to each other? In other words, why is it the case that verbs that occur in the ditransitive construction share some meaning features rather than constitute a grab bag of verbs as dissimilar from each other as any randomly chosen set of verbs with three arguments? We will call *semantic clustering* the fact that verbs that occur in the same syntactic frames are typically semantically similar to each other. Semantic clustering "makes sense" at an intuitive level. There are in the order of fifty or so distinct syntactic frames in English and, according to surveys our lab conducted, college-educated speakers know around 4,000 verbs and about 12,000 verb meanings. Since there are many more verbs (and even more verb meanings) than syntactic frames, there must be a way of grouping verbs that

occur in each syntactic frame. Grouping verbs on the basis of their shared meaning features seems eminently "reasonable." Our goal in this paper is to go beyond the intuition that "it makes sense" and identify some mechanisms that may *cause* semantic clustering and provide initial empirical support for these mechanisms.

Note that, although grouping based on similarity of the kinds of situation that verbs denote (henceforth, situation-type) is "reasonable," it is not impossible for the group of verbs that occur in a syntactic frame to have little to do semantically with each other. Consider the set of verbs that allow indefinite proto-patient arguments to be unexpressed in (4).

(4) a. John baked all day.
  b. *bake, drink, dust, fish, read, sing, sow, wash, write...*

Verbs that allow indefinite proto-patient omission do not seem to form a natural class of situation-types. Rather, what seems to determine the omission of the proto-patient is, to grossly oversimplify, its predictability given the verb (Resnik 1996). Although predictability of the proto-patient is still a semantic property, the grouping of verbs that allow object omission is not based on similarity in situation-types. The existence of patterns like proto-patient omission suggests that semantic clustering is not an a priori necessity and requires an explanation.

The answer we provide is that semantic clustering is due to *priming*, i. e., the fact that exposure to certain stimuli influences our response to subsequent (similar) stimuli. In the case at hand, the relevant kind of stimuli is exposure to the association between a verb meaning and the syntactic frame the verb occurs in. We suggest that two different priming mechanisms can lead to semantic clustering. Both mechanisms operate concurrently and we will not be able to determine how they interact in this paper. For our purposes, it will suffice to show that each mechanism seems to be at play in language production and may provide a partial account of semantic clustering.

We draw our inspiration for the first priming mechanism from an observation made by Goldberg and her colleagues (Goldberg et al. 2004). They observe that for many syntactic frames, one verb occurs in them with disproportionally high frequency. Focusing on the ditransitive frame exemplified in (1a) above, they note that *give* occurs first and the most in child directed speech. They further suggest and provide experimental evidence that the skewed distribution of verbs in the

ditransitive frame helps children learn the construction. That *give* accounts for most of the tokens of the ditransitive frame is confirmed by Bresnan et al.'s (2007) corpus study. We found that, in their corpus, more than 80 % of the ditransitive frame tokens have *give* as their main verb. In our own study of the entire British National Corpus, which we discuss below, *give* accounts for 59 % of the ditransitive frame tokens. We call verbs such as *give* that account for the lion's share of the tokens of a syntactic frame a (global) *semantic anchor*.

2   The global semantic anchor hypothesis

Our first hypothesis, which we call *the global semantic anchor hypothesis*, is that when a verb accounts for such a large proportion of tokens of a syntactic frame, it can serve as an anchor and lead speakers to use the same syntactic frame when describing similar situation-types. The logic of our hypothesis is as follows. When thinking of a message that consists of the description of a situation, a particular situation-type concept is activated. Other concepts that share features with that situation-type concept are also activated. In some cases, one of the concepts sharing features with the situation-type concept about to be expressed is named by the *semantic anchor*. Since the lemma for the semantic anchor is strongly associated with a particular syntactic frame, the semantic anchor's syntactic frame will be strongly activated. As a result, the chances of that frame being chosen to verbalize the situation description will increase (see Reitter et al. 2011). Over time, this increase in activation of that frame for verbs that share features with the semantic anchor will result in those verbs being more likely to occur in the same syntactic frame as the semantic anchor.

If the global semantic anchor hypothesis is correct, verbs that are more similar to the ditransitive semantic anchor, *give*, ought to occur more in the ditransitive frame. Put differently, the more semantically similar a verb is to a global semantic anchor, the more similar it ought to be syntactically, i. e., we predict a correlation between the degree of semantic similarity between a verb and the global semantic anchor and the degree to which a verb occurs in the anchor's preferred syntactic frame. We tested the global semantic anchor hypothesis in two corpus studies.

The first corpus study looked at the ditransitive frame and its alternate prepositional object frame illustrated in (5a) and (5b), respectively; the second corpus

study looked at the material object frame and its alternate locatum object frame illustrated in (6a) and (6b), respectively.

(5) a. John gave his boy a toy.
    b. John gave a toy to his boy.

(6) a. John loaded grocery bags into the car.
    b. John loaded the car with grocery bags.

Both studies used the same methodology. We first computed degree of semantic similarity and frequency of occurrence in a syntactic frame (ditransitive, prepositional object, material object, and locatum object frames). We then examined if there was, as predicted, a correlation between the two measures. We measured semantic similarity between verbs and semantic anchors, using Latent Semantic Analysis (hereafter LSA, Landauer et al. 1998) a computational technique that approximates the semantic similarity between two words (two texts, more generally) by measuring how similar the contexts of occurrence of these two words are. (The contexts of occurrence are the texts in which both words occur or the texts in which words that co-occur with both words being compared occur.) We measured syntactic similarity between verbs and semantic anchors by computing how biased verbs are to occur in the anchor's preferred syntactic frame in the British National Corpus. Since our hypothesis pertains to whether semantic similarity affects the *choice* of syntactic frame in sentence production, we only consider verbs that alternate between the ditransitive and prepositional object frames or material object and locatum object frames.

2.1   The ditransitive alternation

In the case of the ditransitive frame, we computed the LSA semantic similarity cosine between *give* and the 108 alternating verbs mentioned in Levin (1993) that occur in the British National Corpus (hereafter BNC). To measure semantic similarity between *give* and each of 108 other alternating verbs, we used the one-to-many application utility available on-line at http://lsa.colorado.edu.[2] We

---

[2] LSA applications take word forms as input. Thus, word forms that are least ambiguous grammatically constitute a better choice for similarity estimation. For the verb lemma *give*, we chose as input word the past tense form *gave* as it is the only form that is invariably used as a main verb. Choosing only the past form of a verb, however, may not always be the best option. For example, the present and past tense forms of the verb *read* are the same and are identical to the noun form, e. g., *a good*

measured syntactic similarity to *give* by computing the proportion of times each verb occurs in the ditransitive frame (that is, we divided the number of times a verb occurred in the ditransitive frame in the BNC by the number of times it occurred in either the ditransitive or prepositional object frames). Our prediction is that, as semantic similarity to *give* increases, so will the proportion of times a verb occurs in the ditransitive frame. We performed a partial correlation analysis between LSA similarity to *give* and proportion of occurrence in the ditransitive in the BNC, to partial out the effect of frequency, since more frequent pairs of words occur in more texts and therefore will tend to be measured as more semantically similar by the LSA. We did find a correlation, as predicted, between semantic similarity to *give* and proportion of occurrence in the ditransitive frame ($r = .427$; $p < .001$). The global semantic anchor hypothesis is thus supported by this first corpus study.

One potential concern with the result of our study is that the correlation we found may be epiphenomenal as it may reduce to the other factors Bresnan et al. (2007) have found to influence the choice of the ditransitive frame (e. g., whether or not the theme and recipient arguments are pronouns). To determine whether semantic similarity to *give* influences the choice of the ditransitive frame *independently* of the factors Bresnan and her colleagues identified, we conducted a logistic regression on Bresnan et al.'s corpus.[3] Their collection of sentences was from the Switchboard and the Wall Street Journal corpus and only included 38 verbs (rather than the 108 verbs our study included) and 2,360 sentences (rather than the 63,495 sentences our study included). When we included 'Verb Similarity to *give*' as an additional factor, we found it to be a significant predictor of the use of the ditransitive frame even when Bresnan et al.'s predictors were included in the model. We also conducted a separate logistic regression analysis on our full set of 108 alternating verbs and 63,495 sentences. Because it was impractical to hand-code that many sentences for all the factors mentioned in Bresnan et al.'s study, we restricted our analysis to pronominality of the theme and recipient ar-

---

read. We therefore took multiple verb forms as input for verbs that were compared to *gave*: the base form, the third person singular present form, and the past tense form, e. g., *offer, offers* and *offered.*

[3] Their collection is publicly available for download at the publisher's website of *Quantitative methods in linguistics* by Keith Johnson (2008), Blackwell. Some variables mentioned in their 2007 paper (*person, number, structural parallelism* and *concreteness of theme*) are omitted in the publicly available data set.

guments, which can be easily automatically coded for and are the best predictors of the choice of the ditransitive or prepositional object frames. Semantic similarity to *give* was still a predictor of ditransitive use when these additional factors were included. In brief, our original finding that semantic similarity to *give* correlates with proportion of the ditransitive frame does not reduce to other factors known to affect the choice of syntactic frame.

2.2 The locative alternation

Our second corpus study used the same methodology as our first study but concentrated on the locative alternation illustrated in (4). We investigated 45 verbs that participate in that alternation. Frequency of occurrence in the material object frame and semantic similarity to anchor verbs were measured as in the first study. We chose as semantic anchors *rub* for the material object frame as it is the most frequent verb in the frame. But finding a good anchor for the frame is more difficult in this case than it was for the ditransitive frame, as no verb accounts for a large portion of the material object frame. So, whereas *give* accounts for 59 % of *all* ditransitive tokens, *rub* only accounts for 13 % of the material object tokens. It is perhaps not surprising, then, that we did not find a correlation between semantic similarity to *rub* and proportion of occurrence in the material object frame.

The results of our two corpus studies are at odds. Our ditransitive study shows that, as predicted by the global semantic anchor hypothesis, similarity to a verb strongly associated with a syntactic frame increases the likelihood of occurrence in the anchor preferred syntactic frame. These results suggest that shared semantic features affect choice of syntactic frame and contribute to semantic clustering. The locative alternation, though, serves as a cautionary tale. Global semantic anchors like *give* might be rare. The material object frame does not seem to have a good anchor (as we pointed out, *rub* does not account for the lion's share of tokens of their preferred frame). So not all constructions where verbs seem to semantically cluster might have a distribution as skewed as the ditransitive frame.

Semantic anchors like *give* have a global clustering effect: Most of the verbs that participate in the ditransitive frame cluster around it, at least the verbs that entail caused change of possession (see Rappaport & Levin 2008 and Yi, Koenig & Roland 2014 for details). But verbs can play a more local role in semantic

clustering, by which we mean that verbs that can occur in a syntactic frame might clump into small semantic clusters without any single verb anchoring the entire syntactic frame, as *give* does for the ditransitive construction (a point already made in Pinker 1989). That clustering might be more local for some frames is suggested by the material and locatum object frames. Consider the following three groups of verbs that participate in the locative alternation:

- **Group 1:** *mark, engrave, etch* ... (A mark or a symbol is created on the surface of some object.)
- **Group 2:** *brush, powder, rub, spray* ... (Some material is added on the surface of some object.)
- **Group 3:** *load, stock, pile* ... (Something is moved and put in another place.)

Intuitively, semantic similarity within each of these three groups is high, but semantic similarity across groups is low. LSA similarity measures confirm this intuition. Whereas the average LSA semantic similarity between verbs within the three groups was .23, .28, and .28, respectively, semantic similarity across groups was only between .13 and .15. So, even though there might not be any good global anchor for the material or locatum object frames, there might be several good anchors that could lead to semantic clustering, although each anchor would lead to clustering into narrower classes.

We would like to suggest that priming could be at play here, as it was in the clustering effect of global semantic anchors like *give*, but that the effect of priming takes a different form. In the case of global semantic anchors, it is the strong association between *give* and the ditransitive frame that helps verbs that share semantic features with *give* activate more strongly the ditransitive frame even in the absence of any contextual occurrence of *give*.

In the case of *local* clustering, it is the more frequent co-occurrence of verbs with similar meanings than verbs with dissimilar meanings that lead to the clustering. More specifically, we suggest that it is the fact that verbs with similar meanings are more likely to co-occur in the same texts (e. g., paragraphs) that helps semantic clustering. That verbs with more similar meanings tend to co-occur in the same texts more than verbs with less similar meanings is at the root of the relative success of computational techniques for measuring semantic similarity, be it LSA or the Hyperspace Analogue to Language (HAL, Lund et al. 1995).

The assumption on which these techniques rest, as we mentioned earlier, is that words with similar meanings tend to co-occur more or to co-occur with more words that tend to co-occur. We also know from much work in sentence processing that the use of a syntactic frame increases the likelihood of the same frame being reused in the not so distant future (see Bock 1986, and Jaeger & Snider 2008 for corpus evidence of syntactic priming). What we are suggesting is that shared semantic features between verbs increases the likelihood of reusing a syntactic frame and that this increased likelihood can also lead to semantic clustering. We call *local semantic anchor hypothesis* the hypothesis that semantic clustering may also be the result of (1) the tendency of verbs with similar meanings to co-occur more and (2) the increased likelihood of the re-use of a syntactic frame when verbs are semantically similar.

3   The local semantic anchor hypothesis

To test the local semantic anchor hypothesis, we conducted two pairs of syntactic priming experiments. The goal of these experiments was to determine whether the use of a verb in a syntactic frame primes the reuse of that frame more when a subsequent verb is more similar in meaning than when the subsequent verb is less similar in meaning. The first pair of experiments examined the priming of the ditransitive vs. prepositional object frames, whereas the second pair of experiments examined the priming of the locatum vs. material object frames. All experiments used a sentence reading and recall paradigm (Griffin & Weinstein-Tull 2003, Potter & Lombardi 1998). In this paradigm, participants read a pair of sentences and complete aloud partial sentence prompts. Prompts appear in reverse order of presentation of the sentences they read. Thus, they might read in succession the sentences in (7) and then, successively complete aloud sentences (8a) and (8b).

(7) a. The producer promised a large part to the actress
    b. The CEO guaranteed all employees a Christmas bonus

(8) a. The CEO guaranteed ⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯
    b. The producer promised ⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯

The logic of this read-and-recall paradigm is as follows. The syntactic frame of sentences that have just been produced can prime participants to produce subsequent sentences that exhibit the same syntactic frame. So, the frame in the just produced sentence (8a) might prime participants to use the same frame when completing aloud (8b), i. e., it might lead participants to shift the frame of the sentence they read from the prepositional object frame to the ditransitive frame. Priming is thus measured by a shift from the syntactic frame used in the sentence a participant has read to the syntactic frame of the sentence the participant produces in recall. (There could also be a shift from the sentence frame they read last, e. g., (7b), to the frame they produce next, e. g., (8a). But, this shift is less likely, as the memory trace of the sentence that was just read is stronger and in all four experiments such shifts were inconsistently significant. For this reason and because of space considerations, we do not discuss these shifts further in this paper.)

(9) presents an example stimulus set for Experiment 1.

(9) *Reading Phase*:

    a. The producer **promised** a large part to the actress

    b. [One of the three prime conditions]

       High semantic similarity condition:

          The CEO **guaranteed** all employees a Christmas bonus

       Low semantic similarity condition:

          The ball boy **bounced** the player a new ball

       Control condition:

          Organic food **is increasing** in popularity recently

    Recall Phase:

    b. [One of the three prime conditions presented in reading]

       The CEO **guaranteed**/The ball boy **bounced**/Organic food **is increasing**

       _____

    a. The producer **promised** _____

For each stimulus set, participants first read two sentences presented word by word (200ms per word) on the middle of the screen (so-called Rapid Serial Visual Presentation). The first sentence was a prepositional object sentence (9a). The second (9b) was either a ditransitive sentence whose verb was highly seman-

tically similar or not very semantically similar to the verb in the first sentence or a control sentence (most of the times, an intransitive sentence). Participants then attempted to accurately recall, when prompted, the sentences they read in reverse order, i. e., the second sentence they read first, and the first sentence they read second. The sentence they recalled first served as a prime to the second sentence. First recalled sentences fell into three groups, (i) sentences whose verbs were highly similar semantically to the verb in the sentence they would have to recall next (*guaranteed* and *promised* in (9)); (ii) sentences whose verbs were not very similar semantically to the verb in the sentence they would have to recall next (*bounced* and *promised* in (9)), but whose verbs were alternating verbs; (iii) sentences whose verbs were not alternating verbs, mostly intransitive verbs. If the local semantic anchor hypothesis is correct, priming should increase when the two verbs are highly similar semantically as compared to when they are not very semantically similar. Priming was measured as increase in shifts from the prepositional object frame of the sentence that was read to the ditransitive frame of the first recalled sentence when compared to the baseline percentage of shifts provided by the control (mostly intransitive) sentences.

Table 1: Example stimulus sets for Experiments 1 and 2

	Experiment 1	**Expriment 2**
Reading Phase	The producer **promised** a large part to the actress (Prepositional Object frame (PO) frame)	The producer **promised** the actress a large part (Ditransitive or Double Object (DO) frame)
	The CEO **guaranteed** all employees a Christmas bonus (DO; high semantic similarity) The ball boy **bounced** the player a new ball (DO; Low semantic similarity) Organic food **is increasing** in popularity recently (Intransitive control)	The CEO **guaranteed** a Christmas bonus to all employees (PO; high semantic similarity) The ball boy **bounced** a new ball to the player (PO; Low semantic similarity) Organic food **is increasing** in popularity recently (Intransitive control)
Recall Phase	The CEO **guaranteed** _____/ The ball boy **bounced** _____/ Organic food **is increasing** _____ The producer **promised**_____	The CEO **guaranteed** _____/ The ball boy **bounced** _____/ Organic food **is increasing** _____ The producer **promised**_____

The logic of Experiments 2–4 is the same as that of Experiment 1. Tables 1–2 include example stimulus sets for all four experiments.

Materials for all experiments were normed for naturalness and, most importantly, for semantic similarity. For each experiment, we computed the LSA sim-

Table 2: Example stimulus sets for Experiments 3 and 4

	Experiment 3	**Expriment 4**
Reading Phase	The kid **smeared** mom's lipstick on her face (material object (MO) frame)	The kid **smeared** her face with mom's lipstick (locatum object (LO) frame)
	The New Yorker **spread** a toasted bagel with cream cheese (LO; high semantic similarity)	The New Yorker **spread** cream cheese on a toasted bagel (MO; high semantic similarity)
	The freight driver **loaded** the huge truck with lots of boxes (LO; low semantic similarity)	The freight driver **loaded** lots of boxes on the huge truck (MO; low semantic similarity)
	The congressman **decided** to run for the next election (Intransitive control)	The congressman **decided** to run for the next election (Intransitive control)
Recall Phase	The New Yorker **spread** _____/ The freight driver **loaded** _____/ Organic food **is increasing** _____	The New Yorker **spread** _____/ The freight driver **loaded** _____/ Organic food **is increasing** _____
	The kid **smeared** _____	The kid **smeared** _____

ilarity of various pairs of verbs and we also had participants in a norming study evaluate the semantic similarity of the verbs on a Likert scale from 1 (least similar) to 7 (most similar). We used verb pairs where LSA measures and human judgments matched. We chose verbs that were either highly semantically similar or not very semantically similar to our target verbs (the second verb used in the recall phase in Tables 1–2). Table 3 provides mean similarity scores for all four experiments.

Table 3: Mean similarity scores for high and low semantic similarity pairs for the ditransitive (Experiments 1–2) and locative (Experiments 3–4) alternations

	Dative	Locative
HIGH similarity pairs	5.55	4.96
LOW similarity pairs	1.81	1.67

Our overall prediction was that verbs in the first recalled sentence would lead to stronger priming when the prime and target verbs are more semantically similar than when they are less semantically similar. To test this prediction, we compared the effect of both high and low semantic similarity primes to control primes. More specifically, priming was assessed by comparing shifts from the reading phase frame to the first recalled sentence frame in the experimental conditions (i. e., ditransitive frame (Experiment 1), prepositional object frame (Experiment 2), locatum object frame (Experiment 3), material object frame (Experiment 4)) vs. an intransitive control condition. Sixty University at Buffalo undergraduate

students participated in each experiment. In Experiments 1, 3, and 4, we found, as predicted, significant priming when the prime verb was highly similar semantically to the target verb, compared to control primes, but not when the prime verb was not semantically similar to the target verb as Tables 4 and 5 show. In Experiment 1, participants were more likely to shift from the prepositional object frame to the ditransitive frame after producing a ditransitive sentence that contained a highly semantically similar verb than after producing a (mostly intransitive) control prime. But, they showed no such tendency after producing a ditransitive sentence that contained a semantically dissimilar verb. In Experiment 3, participants were only more likely to shift from the material object frame to the locatum object frame when recalling the first sentence they read after producing a locatum object sentence that contained a highly semantically similar verb. Finally, in experiment 4, participants were only more likely to shift from the locatum object frame to the material object frame when recalling the first sentence they read after producing a material sentence that contained a highly semantically similar verb.

Table 4: Percentages of increase in shifts from baseline (control condition) in Experiments 1–4

Experiment	Target shift	Low similarity condition	High similarity condition
Exp 1	PO to DO	2.3	4.5
Exp 2	DO to PO	20.5	19.5
Exp 3	MO to LO	2.4	5.2
Exp 4	LO to MO	4.4	11.4

Table 5: Results of the mixed effect logistic regression analyses for Experiments 1–4

Experiment	Condition	Coefficient Estimate	Std error	*p* value
Exp 1	Low vs. Ctrl	0.89	0.69	0.2
	High vs. Ctrl	1.24	0.61	.04*
Exp 2	Low vs. Ctrl	0.46	0.2	.02*
	High vs. Ctrl	0.55	0.2	.007*
Exp 3	Low vs. Ctrl	0.48	0.45	0.29
	High vs. Ctrl	0.89	0.44	.04*
Exp 4	Low vs. Ctrl	0.4	0.29	0.17
	High vs. Ctrl	0.79	0.29	.007*

Experiment 2 did not conform to our predictions in that both highly semantically similar and semantically dissimilar prime verbs lead to significantly more

shifts from the ditransitive frame of the reading phase to the prepositional object frame in the recall phase as compared to control primes. We surmise that the different pattern of results in Experiment 2 is due to the fact that the shift in this experiment was to an overwhelmingly preferred prepositional object frame. Recall that in the case of Experiment 2, shifts were from the dis-preferred ditransitive frame in the reading phase to the preferred prepositional object frame in the recall phase. As is well known, the prepositional object frame is strongly preferred among alternating verbs, including the verbs we used in our materials. Furthermore, ditransitive frame sentences most often include a pronominal recipient, which our materials did not include. The combined effect of these two preferences is that for the verbs used in Experiment 2, over 95 % of the time, when the recipient is not encoded as a pronoun, the prepositional object frame is used. This explains that *in the control condition* shifts to the ditransitive frame in Experiment 1 occurred about 3 % of the time, but shifts to the prepositional object frame in the control condition in Experiment 2 occurred over 40 % of the time. It seems that when a syntactic frame is as overwhelmingly preferred as the prepositional object frame for full lexical NP recipients, syntactic priming occurs regardless of semantic similarity. More precisely, the increase in syntactic frame shifts due to semantic similarity is small when compared to the effect of syntactic preferences (about 2 % vs. about 20 %, see Table 4). A likely consequence is that any potential effect of semantic similarity is dwarfed by the effect of such a strong syntactic preference.

Overall, the results of our four experiments suggest that semantic similarity increases priming of syntactic frames: Speakers are more likely to incorrectly recall the frame they read when they just produced a sentence with an interfering alternative frame whose verb was highly semantically similar. The mechanism that we suggested may underlie local semantic clustering thus receives experimental support. Interestingly, we found effects of semantic similarity on priming, and support for the local semantic anchor hypothesis, for both the locative alternation as well as for the dative alternation. This is important, since, as we saw earlier, our corpus studies did not support the global semantic anchor hypothesis for the locative alternation. We suggested that global semantic anchors as strong as *give* for the ditransitive construction might be quite rare and that, since semantic clustering is rather general and obtains across many syntactic frames, some other mechanism must be at play. We suggested that two other possibly

overlapping mechanisms might account for semantic clustering for alternations without a global semantic anchor. The first is a more circumscribed version of the semantic anchor hypothesis: Within narrower classes, there may be semantic anchors (maybe within classes akin to Pinker's narrow classes). Unfortunately, the small numbers of verbs in the locative alternation's narrow classes prevent us from testing this hypothesis. But a second possibility is that clustering might be the result of pairwise similarities, in that pairs of verbs that are semantically similar might share their relative propensity to occur in the same syntactic frame(s). Such clustering might lead to the intuition researchers have had that verbs that occur in a syntactic frame share semantic features. It is this second possibility our experiments tested and found support for.

4   General discussion

Most of our research, as linguists, focuses on determining what grammars are like. In the case of the effect of the meaning of verbs on the syntactic frames in which they can occur, that has meant figuring out how much of the syntax of verbs can be predicted from their meaning and what aspect of the meaning of verbs tends to be relevant for predicting their syntax. In this paper, we asked a question linguists ultimately strive for, an explanation for why grammars are the way they are, here why verbs with similar meanings tend to occur in similar syntactic frames. We suggested two mechanisms might provide a partial answer to this question.

The first, which we call *global semantic anchoring*, is operative when a verb accounts for much of the tokens of a frame, such as *give* for the ditransitive frame. Goldberg et al. (2004) suggest that syntactic frames like the ditransitive construction, are associated with "primary scenes" and that verbs like *give* are simply names of these "primary scenes," thus explaining their frequent occurrence in the frame. We hypothesized that these "primary scenes" verbs semantically anchor the syntactic frame, so that other verbs' occurrence in the syntactic frame will correlate with their semantic similarity to the anchor. The mechanism underlying this global anchoring, we suggest, works as follows. When thinking of a message and an event description for a planned utterance, speakers activate verbs that share semantic features with that event description. The more semantically similar two verbs are, the more the activation of one verb will activate the other

verb. When a global semantic anchor is among the semantically similar verbs that are also activated, the activation of the global anchor's preferred syntactic frame is relatively high, because the anchor verb is strongly associated with that frame. As a result of this increase in activation of the anchor's preferred frame, speakers are more likely to choose that frame when producing their intended event description.

We provided evidence for the effect on syntactic frame selection of global semantic anchors like *give* through a corpus study. We showed that there was a correlation between a verb's semantic similarity to *give* (as measured by Latent Semantic Analysis) and the choice of the ditransitive frame rather than the prepositional object frame in the British National Corpus. We further showed that the influence of a verb's semantic similarity to *give* does not reduce to other factors known to affect the choice of the ditransitive frame (Bresnan et al. 2007). A comparable study of the locative alternation, however, suggests that what we call *global semantic anchors* like *give* for the ditransitive frame might not be that common and thus cannot constitute the only mechanism underlying semantic clustering. Indeed, we did not find a correlation between putative semantic anchors and the material object frame. There are several possible reasons that we suggested might explain the absence of a global semantic anchor for this frame. First, no verb accounts for so many of the tokens of the frame in the BNC (e. g., *rub*, the verb that occurs the most in the material object frame accounts for only 13 % of the tokens of this frame in the BNC). So, it is not clear that there is a strong anchor for this frame. Second, the material object frame includes verbs that are quite distant semantically. As less semantic features are shared between verbs, the activation of the concept associated with any verb selected by the speaker will activate less the concept associated with the anchor.

As we mentioned, Pinker (1989) noted that not all verbs that participate in a syntactic frame form a coherent semantic class and he distinguished between narrow and broad classes of alternating verbs. Only narrow classes constitute natural semantic classes. What this suggests is that while semantic priming and global semantic anchors can play a role in explaining semantic clustering, this cannot be the entire story. We suggested that a similar, but more local mechanism might play a role in explaining semantic clustering. We hypothesized that a verb that is just heard in a particular syntactic frame will prime semantically similar verbs to occur in the same syntactic frame. That the occurrence of a syntactic frame

will prime the subsequent reuse of that frame is of course something that has been well established since Bock's seminal work. But, for that well-established observation to help explain semantic clustering, we must assume that verbs that share semantic features tend to be more likely to occur in the same stretch of discourse than verbs that are semantically further apart. And of course, this is what computational techniques that measure semantic similarity assume. So, the relative success of LSA or HAL suggests that there is something to the idea that semantically similar verbs will occur in the vicinity of each other.

The second part of our paper showed that *local semantic anchors* can indeed play a role in online sentence production. We hypothesized that semantic similarity would increase syntactic priming so that verbs would prime other verbs to occur in the same syntactic frame more than verbs that are not semantically similar. We conducted four syntactic priming experiments that tested the local semantic anchoring hypothesis. Two of these experiments involved the ditransitive/prepositional object frames and two involved the locatum object/material object frames. Each experiment tested the hypothesis that semantic similarity increases priming in the form of increasing shifts from alternating structure A in reading to alternating structure B in recall after speakers have just produced a sentence using structure B that contained a semantically similar verb. Three of the four experiments we conducted conformed to our predictions. Semantically similar prime verbs lead to more shifts to the alternating structure in recall than verbs that were semantically dissimilar. But in Experiment 2, the use of the prepositional object frame in recall did not vary as a function of the semantic similarity of the verb in the prime and target sentences. We surmised that the reason for the irrelevance of semantic similarity in this case lies with the fact that the prepositional object frame is overwhelmingly preferred when recipients are expressed as non-pronominal NPs (over 95 % in the BNC for the verbs included in our experiments). Priming of such a preferred structure is so strong that the relatively small effect of semantic similarity becomes invisible.

Overall, then, we found empirical support for both of the mechanisms that together or separately might begin to provide an explanation for semantic clustering. Global and local semantic anchors might be an important part of what gives rise to semantic clustering. Several issues remain to be explored, however. First, it is yet unclear what role global anchors play. The role of *give* for the ditransitive frame might be the exception, rather than the norm. Much ink

has been shed on that frame in theories of linking lexical semantics to syntactic frames from Fillmore and Green on. The ditransitive frame, alongside passives, has also received most of the attention in the psycholinguistic study of syntactic priming. But a cursory look at other syntactic frames, including the locatum object/material object frames, suggests generalizing from the ditransitive to the bulk of syntactic frames might be unwarranted.

Second, criteria for selecting potential global semantic anchors must be refined. In this paper, we selected anchors on the basis of the proportions of tokens of the syntactic frame in the BNC they accounted for: Semantic anchors were simply the verbs that accounted for the largest number of tokens of the syntactic frame in the BNC. But other criteria or additional criteria might be needed. For example, we would expect anchors for a particular syntactic frame to occur significantly more often in that frame than in the alternative frame(s). Otherwise, how could they anchor that frame? In such circumstances what is more important? Is there a combination of the two criteria that is optimal? (See Yi et al. 2014 for a proposal.)

Third, what counts as a distinct syntactic frame is not always clear. Here again, the ditransitive construction might be quite unique. Aside from rather infrequent causatives (e. g., *Her victory made her the most sought after athlete*), the sequence of two post-verbal NPs in English is criterial of the ditransitive construction. But the post-verbal sequence NP+PP can be ambiguous between what *could be* considered several distinct syntactic frames (e. g., *John threw the ball into the basket* and *John loaded the suitcases into the car*).

Fourth, if local semantic anchors play an important role in semantic clustering, it would suggest that syntactic frames are a collection of related frames, each targeting a small, semantically coherent cluster of verbs, not unlike Pinker's (1989) narrow classes. If distinct semantic anchors might account for each small semantic clustering, what accounts for the collection of small semantic clusters? Take the small clusters of verbs like *load* and *engrave* that participate in both the material object and locatum object frames. Both semantic clusters are quite coherent semantically and members of each cluster alternate between the locatum object and the material object frames. If semantic anchoring accounts for membership in each small semantic cluster, what accounts for the fact that both clusters alternate between the same two syntactic frames (assuming, for now, that this is the case)? Obviously, we cannot provide an answer to these questions in this short, rather speculative paper. What we hope to have shown is that semantic anchors and

"priming" in a broad sense might explain this recurring fact of the syntax of basic clauses: Verbs that participate in the same or a similar set of syntactic frames look alike semantically.

References

Bock, J. K. 1986. Syntactic persistence in language production. *Cognitive Psychology* 18: 355–387.

Bresnan, J., A. Cueni, T. Nikitina & R. H. Baayen. 2007. Predicting the dative alternation. In G. Boume, I. Kraemer & J. Zwarts (eds.), *Cognitive foundations of interpretation*, 69–94. Amsterdam: Royal Netherlands Academy of Science.

Fillmore, C. 1968. The case for case. In E. Bach & R. Harms (eds.), *Universals in linguistic theory*, 1–87. New York: Holt, Rinehart and Winston.

Goldberg, A. E. 1993. Another look at some learnability paradoxes. In E. V. Clark (ed.), *Proceedings of the twenty-fifth annual child language research forum*, 60–75. Stanford, CA: CSLI Publications.

Goldberg, A. E. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: Universtiy of Chicago Press.

Goldberg, A. E., D. M. Casenhiser & N. Sethuraman. 2004. Learning argument structure generalizations. *Cognitive Linguistics* 15: 289–316.

Green, G. 1974. *Semantics and syntactic regularity*. Bloomington: Indiana University Press.

Griffin, Z. M. & J. Weinstein-Tull. 2003. Conceptual structure modulates structural priming in the production of complex sentences. *Journal of Memory and Language* 49: 537–555.

Grimshaw, J. 1993. *Semantic structure and semantic content in lexical representation.* Department of Linguistics, Rutgers University.

Gruber, G. 1965. *Studies in lexical relations*. MIT dissertation.

Jaeger, F. & N. Snider. 2008. Implicit learning and syntactic persistence: surprisal and cumulativity. *Proceedings of the 30th Meeting of the Cognitive Science Society*, 1061–1066.

Koenig, J.-P. & A. R. Davis. 2006. The KEY to lexical semantics representations. *Journal of Linguistics* 42: 71–108.

Koenig, J.-P., D. Roland, H. Yun & G. Mauner. 2015. Which event properties matter for which cognitive task? In R. de Almeida & L. Manouilidou (eds.), *Cognitive*

science perspectives on verb representation and processing, 213–234. Dordrecht: Springer.

Landauer, T. K., P. W. Foltz & D. Laham. 1998. An introduction to latent semantic analysis. *Discourse Processes* 25: 259–284.

Levin, B. 1993. *English verb classes and alternations.* Chicago: Universtiy of Chicago Press.

Lund, K., C. Burgess & R. A. Atchley. 1995. Semantic and associative priming in a high-dimensional semantic space. *Proceedings of the 17th Annual Conference of the Cognitive Science Society*, 660–665. Mahwah, NJ: Lawrence Erlbaum Associates.

Potter, M. & L. Lombardi. 1998. Syntactic priming in immediate recall of sentences. *Journal of Memory and Language* 38: 265–282.

Pinker, S. 1989. *Learnability and cognition: the acquisition of argument structure.* Cambridge, MA: MIT Press.

Rappaport, M. & B. Levin. 2008. The English dative alternation: the case for verb sensitivity. *Journal of Linguistics* 44: 129–167.

Ramchand, G. 2008. *Verb meaning and the lexicon: a first phase syntax.* Cambridge: Cambridge University Press.

Reitter, D., F. Keller & J. D. Moore. 2011. A computational cognitive model of syntactic priming. *Cognitive Science* 35: 587–637.

Resnik, P. 1996. Selectional constraints: an information-theoretic model and its computational realization. *Cognition* 61: 127–159.

Van Valin, R. D., Jr. & R. J. LaPolla. 1997. *Syntax: form, meaning and function.* Cambridge: Cambridge University Press.

Yi, E., J.-P. Koenig & D. Ronald. 2014. The role of frequency-driven association between verb and syntactic frame on syntactic frame selection. Manuscript submitted for publication.

Authors

Eunkyung Yi & Jean-Pierre Koenig
University at Buffalo
eyi@buffalo.edu, jpkoenig@buffalo.edu