J. Child Lang. **32** (2005), 407–426. © 2005 Cambridge University Press
doi:10.1017/S0305000904006798 Printed in the United Kingdom

NOTE

The role of prediction in construction-learning*

ADELE E. GOLDBERG
Princeton University

DEVIN M. CASENHISER
Princeton University

AND

NITYA SETHURAMAN
Indiana University

(*Received* 21 *August* 2003. *Revised* 20 *April* 2004)

ABSTRACT

It is well-established that (non-linguistic) categorization is driven by a functional demand of prediction. We suggest that prediction likewise may well play a role in motivating the learning of semantic generalizations about argument structure constructions. We report corpora statistics that indicate that the argument frame or construction has roughly equivalent cue validity as a predictor of overall sentence meaning as the morphological form of the verb, and has greater category validity. That is, the construction is at least as reliable and more available than the verb. Moreover, given the fact that many verbs have quite low cue validity in isolation, attention to the contribution of the construction is essential.

INTRODUCTION

As many psychologists have emphasized, human categorization is generally driven by some functional pressure, typically the need to predict or infer

[*] The authors are grateful for helpful discussions with Giulia Bencini, Cindy Fisher, Lynn Gale, Gregory Murphy, Mike Tomasello, Yasuhiro Shirai, and to two anonymous reviewers for helpful comments on an earlier draft. We are also grateful to Liz Bates, for being a wonderful teacher, mentor and friend. This research was funded in part by the first author's NSF Grant SBR-9873450 and fellowship at the Center for Advanced Study in the Behavioral Sciences. Address for correspondence: Adele E. Goldberg, Linguistics, East Pyne Hall, Princeton University, Princeton, NJ 08544, USA. e-mail: adele@princeton.edu

https://doi.org/10.1017/S0305000904006798 Published online by Cambridge University Press

certain properties on the basis of perceived characteristics (Holland, Holyoak & Thagard, 1989; Anderson, 1991; Leake & Ram, 1995; Wisniewski, 1995; Kersten & Billman, 1997; Ross & Makin, 1999; Murphy, 2002). That is, cognitive systems do not generalize randomly or completely. Holland *et al.* (1989), in their monograph on induction, emphasize that generalizations are constrained in that 'the inferences drawn by a cognitive system will tend to be … relevant to the system's goals' (p. 5). In the case of language, the language learner's goal is to understand and to be understood: to comprehend and produce language. There is ample functional pressure to predict meaning on the basis of given lexical items and grammatical characteristics (comprehension); conversely, there is pressure to predict the choice of lexical items and grammatical characteristics given the message to be conveyed (production). Since the sentences the child is learning to understand and produce form an open-ended set, it is not sufficient to simply memorize the sentences that have been heard. The child must necessarily generalize those patterns at least to some extent in order to understand and produce new utterances.

The predictive value of verbs in argument structure patterns

There is a long history in the field of linguistics of considering the main verb to be the key word in a clause (Chomsky, 1965; Lakoff, 1970; Pinker, 1989; Grimshaw, 1990; Levin & Rappaport, 1995). This has also been true in the field of language acquisition (e.g. Tomasello, 1992, 2000). A critical factor in the primacy of verbs in argument structure patterns stems from their relevant predictive value. If we compare verbs with other words (e.g. nouns), verbs are much better predictors of overall sentence meaning, where by 'overall sentence meaning' we basically intend 'who did what to whom,' a level of generalization that is uncontroversially required for adequate sentence comprehension.

Experimental evidence for the idea that verbs play a key role in semantic interpretation is provided by Healy & Miller (1970). Healy & Miller compared the relative contribution of verbs and subject arguments to overall sentence meaning. These two candidates, verb and subject, were presumably chosen because they appear to be the best candidates for representing overall sentence meaning. The subject argument is often referred to as the 'topic' argument in a sentence or what the sentence is 'about' (Kuno, 1972; Reinhart, 1982; Lambrecht, 1994). At the same time, the verb provides a great deal of information about who did what to whom. Healy & Miller constructed 25 sentences by crossing 5 subject arguments (*the salesman*, *the writer*, *the critic*, *the student*, *the publisher*), 5 verbs (*sold*, *wrote*, *criticized*, *studied*, *published*) and one patient (*the book*). Participants were asked to sort the sentences into five piles according

https://doi.org/10.1017/S0305000904006798 Published online by Cambridge University Press

to similarity in meaning. Results showed that participants reliably sorted sentences together that had the same verb much more often than sentences that had the same subject argument. That is, for example, all five sentences with the verb *criticized* were categorized together much more often than five sentences with the subject *the critic*. Given these results, Healy & Miller concluded that the verb is the main determinant of sentence meaning.

Another source of evidence for the idea that the verb is a good predictor of sentence meaning comes from work on analogy. It has been richly documented that relational aspects of meaning are fruitful sources of analogy and similarity judgments (Gentner, 1982). Markman & Gentner (1993), for example, found that in making non-linguistic similarity judgments, similarity is judged to be greater when two representations share the same relations between the entities in each representation. That is, the entities are aligned based on the structure that relates them, rather than on the basis of independent characteristics of the entities. The relevance to language is straightforward. In comparing two sentences, the main relational predicates, the verbs, are more likely to be used than the independent characteristics of the arguments (Tomasello, 2000). Why should this be so? The purpose of analogies is generally one of drawing inferences and making predictions: what can be predicted on the basis of one situation about another situation (Gentner & Medina, 1998). Thus it is the value of verbs as good cues to sentence meaning that results in the child's early learning of verb-centred argument structure patterns ('verb islands').

This paper focuses on the question of why learners ever generalize beyond the verb to the more abstract level of argument structure constructions. The fact there exist generalizations at the relevant level of abstraction is uncontroversial, whether they be termed linking rules, event templates or argument structure constructions. We aim to demonstrate here that there is predictive value in learning such generalizations, thereby offering one type of motivation for speakers to generalize beyond the input (for other types of motivation, see discussion in Goldberg to appear chapter 7).

The value of constructions as predictors of sentence meaning

In the rest of the paper we demonstrate that generalizing beyond a particular verb to a more general pattern is useful in predicting overall sentence meaning, more useful in fact than knowledge of individual verbs. Bates & MacWhinney (1987) have stressed the importance of weighting different cues, dependent on how reliable and available each cue is. We hypothesize that it is the predictive value that encourages speakers to generalize beyond knowledge of specific verbs to ultimately learn the semantic side of linking generalizations, or constructional meaning.

https://doi.org/10.1017/S0305000904006798 Published online by Cambridge University Press

TABLE 1. *Construction types, defined formally*

Label	Form
VOL:	(Subj) V Obj $\text{Obl}_{\text{path/loc}}$
VOO:	(Subj) V Obj Obj2

Precedent for this idea comes from work in the non-linguistic categorization literature. Kruschke (1996) and Dennis & Kruschke (1998) discuss how learners shift attention away from less reliable (i.e. less distinctive) cues toward more reliable cues, when learning overlapping instances that belong to distinct categories. For example if two diseases share one symptom but have their own distinctive symptom, subjects will attend more to the distinctive symptoms than the shared one.

Two construction types are focused on in this paper:

It is clear that constructions are sometimes better predictors of overall meaning than many verbs. For example, when *get* appears in the VOL construction, it conveys caused motion, but when it appears in the VOO construction, it conveys transfer:

(1) a. Pat got the ball over the fence.
get + VOL pattern → 'caused motion'
b. Pat got Bob a cake.
get + VOO pattern → 'transfer'

As quantified below, *get* in isolation has a low CUE VALIDITY as a predictor of sentence meaning. Since most verbs appear in more than one construction with corresponding differences in interpretation, speakers would do well to learn to attend to the constructions. As an indication of the fact that the construction is at least as good a predictor of overall sentence meaning as the verb, we consider the actual predictive value of verbs vs. formal patterns in a corpus of speech to young children.

Clearly if we compare the contribution of verb and construction to subtle aspects of meaning involving manner or means, the verb would be more predictive than the construction. This is necessarily true since constructions rarely encode specific meanings: compare 'X causes Y to receive Z,' the meaning of the ditransitive construction with the meaning of the verbs 'hand' or 'mail.' At the same time, both verbs and constructions have the potential to convey the overall event-level interpretation, roughly 'who did what to whom.' Since the event level interpretation (who did what to whom) is clearly a necessary component of interpretation, we chose to compare the relative contribution of constructions and verbs at this level. Clearly, in order to arrive at a full interpretation of a sentence, the specifics contributed by only the verb (and its arguments) are required as well.

https://doi.org/10.1017/S0305000904006798 Published online by Cambridge University Press

CORPUS EVIDENCE OF THE CONSTRUCTION AS A RELIABLE PREDICTOR OF OVERALL SENTENCE MEANING

Database

We examined the Bates corpus (Bates, Bretherton & Snyder, 1988) on the Child Language Data Exchange System (CHILDES) database (MacWhinney, 1995). This corpus contains transcripts from the Bates/Bretherton Colorado longitudinal sample of 27 middle-class children, 13 boys and 14 girls aged 1;8 and 2;4. There are transcripts for 15 minutes, equally divided into three types of mother–infant interaction: free play, reading of the book *Miffy in the Snow*, and snack time.

Data collection and coding

The speech of 15 mothers to their children aged 2;4 were extracted from the Bates *et al*. transcripts. This data was analysed independently by two coders. A second set of 7 mothers was coded separately once in order to determine whether there was any bias in our sample of 15 mothers. Identical trends were noted in the second sample so we used the sample of 15 mothers that had been coded independently by the two coders. Word and sentence segmentation decisions were respected. All complete child and adult utterances containing a verb were included in the coding. Every utterance was coded for construction type, verb, and speaker + situation + age. After an initial sampling of sentences with overt subjects and sentences without overt subjects, we collapsed the results because the only major difference between the two sets of data was that sentences without subjects were predominantly commands. There did not seem to be any distinctions relevant to the current discussion.

Classifications were based primarily on form: categorizing an utterance as an instance of the VOO pattern required a verb and two NPs. The VOL pattern required a verb with an object NP and some type of locative phrase: a preposition phrase indicating location, a particle indicating location (e.g. *down*, *in*), a locative (*there*, *here*), or some combination. The emphasis on form in our coding scheme was intended to avoid presupposing that utterances had a particular meaning associated with them. The decision to code for locative phrases ('L's) instead of a purely syntactic category such as PPs was based on the fact that ultimate linking generalizations are known to be sensitive to general semantic properties of the oblique argument such as what general type of preposition is involved. Since at least certain prepositions and locative expressions (including *up*, *down*, *off*, *in*, *on* and *here*) are often among the first words uttered (Clark, 1996), we assume that children are able to identify phrases as locative or non-locative.

We ignored variable word orders correlated with questions, topicalizations, etc. Questions such as *What did she put in his eyes?*, for example,

https://doi.org/10.1017/S0305000904006798 Published online by Cambridge University Press

TABLE 2. *Sample from the Bates* et al. (1988) corpus

Label	Verb	Child	Context utterance
VOL:	put:	*KEI-FR28:	put another ball in here.
VOL:	want:	*WAN-ST28:	want you in the house.
VOO:	give:	*GLO-SN28:	Cory gave me my turtle.

were considered instances of the VOL pattern.[1] Utterances considered unacceptable or ungrammatical to adults were included in the coding. We did not attempt to distinguish arguments from adjuncts because we did not want to assume the children had mastered the distinction at 2;4 years. A sample from the coded transcript is given below:

We reviewed all instances of the VOL and VOO constructions in the database of mothers' speech to children in the corpus.

RESULTS FOR THE VOL PATTERN

We first examined whether the formal pattern predicted the semantic caused-motion meaning. 'Cue validity' is the conditional probability that an object is in a particular category, given that it has a particular feature or cue (Murphy, 1982). Two coders classified utterances as either entailing caused motion or not; those that we judged not to entail caused motion were separated and further analysed as discussed below. Agreement between the two independent coders was 99% for classifying mothers' utterances as instances of the VOL pattern. Agreement for classifying VOL utterances as entailing literal caused motion, metaphorical caused motion, caused location or not was 97%. Disagreements were resolved through discussion.

P(A|B) is the probability of A, given B. As detailed below, the cue validity of VOL as a predictor of 'caused-motion' meaning, or P('caused-motion'|VOL), is between 0·62 and 0·83, depending on how inclusive we take the notion of caused-motion to be, and how inclusively we define the VOL formal pattern. We found that 62% (159/256) of the mother's instances of the construction clearly entail literal caused motion. The following examples are representative:

(2a) get some more in it
(2b) bring 'em back over here
(2c) stuff that all in your mouth

[1] It is possible that children (or even adults) may not generalize over these distinct word orders but since examples with non-canonical word order accounted for fewer than 10% of the token utterances, and did not bias the sample in favor of either verbs or constructions, including them should not affect the overall analysis.

https://doi.org/10.1017/S0305000904006798 Published online by Cambridge University Press

(2d) put 'em in the box
(2e) park the car over there
(2f) pick it up here
(2g) take 'em in the house

Another 20 instances involve the verbs *keep*, *have*, *get* or *leave* as in the following type of examples:

(3a) keeping these people in the garage
(3b) does he have something on his head?
(3c) what's she got on?
(3d) leave it right there

The utterances in (3) entail that the subject argument acts to keep or allow the theme argument to stay in a particular location. The subject argument is agentive and the locative phrase is predicated of the direct object argument just as in instances that entail prototypical caused motion. Many researchers have related these instances to cases of caused motion independently (e.g. Talmy, 1976; Matsumoto, 1992; Goldberg, 1995). If we include these cases in the final tally, 70% of VOL utterances imply caused motion or caused location. Another 2% (5) of instances involve the verb *want* as in:

(4) Oh, you want them in a cup?

These instances convey possible future caused motion. If we include these in the tally, the percentage of instances that are related to caused motion increases to 72%. Another 5% (13) of instances involve the verbs *read* or *say* which could be argued to encode metaphorical caused motion (Reddy, 1979; Goldberg, 1995; Ackerman & Webelhuth, 1998). Including these cases would raise the total number of VOL utterances whose meanings are related to caused motion to 77%.

Of the remaining VOL utterances, 25, or 10% involved locative adjuncts. We included these as instances of VOLs because we did not want to assume that the children had mastered the distinction between arguments and adjuncts. If we had excluded them, the total number of utterances included as VOL utterances would have been reduced from 256 to 231. The total percentage of VOL utterances that involved caused motion would be 85%.

The remaining 34 tokens include examples such as the following, which do not convey caused motion:

(5a) What is your foot doing on the table? (The WXDY construction:[2] 16 instances)
(5b) What did Ivy do to her arm? (1 instance)
(5c) find the bird in the snow (utterances with *find*: 2 instances)

[2] (Kay & Fillmore, 1999).

https://doi.org/10.1017/S0305000904006798 Published online by Cambridge University Press

TABLE 3. *Cue validity of VOL construction as a predictor of caused motion*

Strict encoding of caused-motion meaning	Inclusive encoding
0·62	0·85

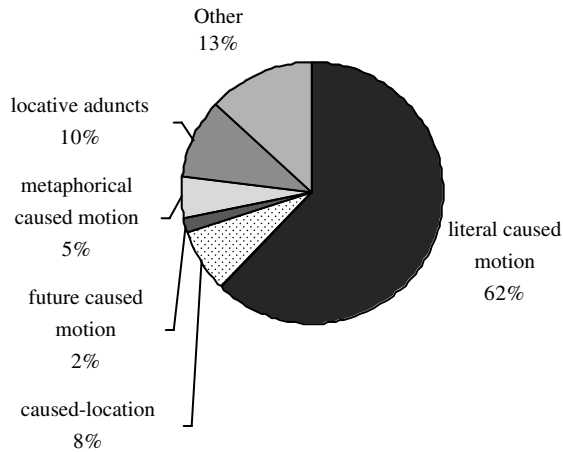


Fig. 1. Proportion of utterances that were related to caused-motion (only 13% in 'other' category is unrelated to caused motion).

(5d) get Papa at the airport (move-from interpretation: 1 instance)
(5e) stand it up (verb particle interpretations: 14 instances)

To summarize, the cue validity of the VOL pattern as a predictor of caused motion meaning is provided in Table 3 below, with a more specific breakdown given in Figure 1.

We then investigated the extent to which individual verbs predicted caused motion meaning. Table 4 shows our calculations of the cue validities of individual verbs that appear at least 5 times in mothers' speech in the Bates *et al.* (1988) corpus.[3] For reliability, a second coder independently classified a sample of 34% (404/1195) of the total number of utterances, including all of the tokens in the corpus, for the verbs *put*, *get*, *take*, *read*, *see*, *stand* and *turn*. The two coders agreed on whether the verb determined the overall sentence meaning reliably 89% (360/404) of the time. The other utterances were also classified independently by two coders, with disagreements resolved through discussion.

[3] Individual verbs that appeared less than 5 times in the VOL pattern accounted for a total of 56 instances. These were: *eat*, *pick*, *set*, *throw*, *want*, *clean*, *close*, *cook*, *dump*, *find*, *give*, *hold*, *keep*, *leave*, *let*, *make*, *move*, *park*, *pour*, *push*, *send*, *stack*, *stick*, *stuff*, *try*, *wake*, *walk*, *wear*, *wipe*.

https://doi.org/10.1017/S0305000904006798 Published online by Cambridge University Press

TABLE 4. *Cue validity of verbs as predictors of their predominant meaning in the VOO pattern*

verb	A. # of instances of stable verb meaning	B. total # of times verb appears in corpus	C. Cue validity (A/B)	D. # of times in VOL pattern	E. Weighted cue validity (C × D/200)
put	113	114	1	99	0·50
do	147	601	0·24	26	0·03
have	15	115	0·13	16	0·01
get	18	108	0·17	14	0·01
take	9	44	0·20	14	0·01
read	14	24	0·58	7	0·02
see	30	86	0·35	7	0·01
stand	8	8	1	7	0·04
say	6	65	0·09	6	0·00
bring	10	10	1	6	0·03
turn	10	20	0·5	6	0·02
Total		1195		200	**0·68**

Column A lists the number of utterances in which the verb's meaning as it is determined to appear most frequently in the VOL pattern, appears overall in the corpus (including appearances in the VOL pattern). Column B lists the total number of times the verb appears in the corpus. Simple cue validity was calculated by dividing column A by column B; this is recorded in Column C. It is clear from Table 4 that there is a wide variability of cue validities across verbs. While a few verbs have perfect or near perfect cue validities (*put*, *bring*, *stand*) in our corpus, other verbs' cue validities were low (*do*, *get*, *have*, *let* and *take*). For the latter verbs, relying on the construction in conjunction with the verb is essential to determining sentence meaning. This fact in itself is sufficient to conclude that attention to the semantic contribution of constructions is required for determining overall sentence meaning.

In order to determine the cue validity of verbs as they appear in the VOL pattern, it is necessary to weight the cue validities of each verb according to how often the verb occurs in the VOL pattern. We therefore multiply the cue validity obtained in Column C by the number of times the verb occurs in the VOL pattern (recorded in Column D) and divide by the number of VOL tokens in the corpus (200). Summing over the weighted cue validities in Column D provides us with the overall cue validity of 0·68 for verbs in the VOL pattern in our corpus. If we compare the 0·62–0·85 cue validity for the VOL pattern as a predictor of caused motion meaning, we can see the construction is roughly as valid a cue.

To see that the weighted average of cue validities is more revealing than the simple average, it is illustrative to consider the following hypothetical situation. Imagine that there were one verb that accounted for 90% of the tokens

https://doi.org/10.1017/S0305000904006798 Published online by Cambridge University Press

of a particular construction and had a cue validity of 1. In this particular case, the predictive value of verbs in the construction would clearly be quite high. Yet if there happened to be ten other verbs appearing in the construction, each accounting for 1% of the tokens, and each having low cue validity (in the limiting case close to 0), the average cue validity would only be approximately 0·09. This is not the number we're after, since it does not reveal the fact that the most likely verb to appear is highly predictive. The weighted average in this circumstance would be roughly 0·90, accurately reflecting the predictive nature of verbs in this hypothetical construction. Thus it is the weighted average of cue validities that more accurately reflects the predictive value of verbs.[4]

Let us examine the cue validities of a few individual verbs in a bit more detail. 114 instances involved the verb *put* which is a perfect predictor of the overall sentence's meaning, since its interpretation was virtually invariant in our corpus.[5] However, the next most frequent verbs in the VOL construction were *do*, accounting for 26 tokens, *have* (16 tokens), *take* (14 tokens), and *get* (14 tokens). These verbs are not reliably associated with the meaning they have in the VOL pattern insofar as each of these verbs occurs more frequently in a different construction, with a different meaning. We consider each of these in turn.

The verb *do* appears 26 times in the VOL construction. Twenty-three of these cases were main verb uses, with a general meaning of 'to act.' Another 124 appearances of *do* in the corpus were also main verb uses with the same meaning, for a total of 147. However, this is out of a total number of 601 instances of *do* in the corpus; in the balance of these uses, *do* serves as an auxiliary marker indicating only aspect and/or polarity. The relevant cue validity is P(main verb use | *do*) = 147/601 = 0·24. Thus *do* is, in and of itself, not a terribly reliable predictor of overall sentence meaning when it appears in the VOL frame.

Get appears 10/14 times in the VOL construction with a caused-motion interpretation (e.g. *or do you want to play with it and get it all over the floor?*), and 8 other times in the corpus with a similar caused-motion interpretation. The same morphological form occurs in a variety of different constructions with different meanings, including the simple transitive (primarily expressing possession as in *get a washcloth*); the transitive resultative (expressing a caused change of state as in *get them undone*), the intransitive motion construction (as in *get out of the car*), the ditransitive (expressing transfer as in *get me some of those*) and with a VP complement (expressing modality as in

[4] The simple average of cue validities is 0·47, while the weighted average of cue validities is 0·68. Thus calculating cue validities without weighting them would only strengthen our claim that the cue validity of constructions is at least as high as that of verbs.

[5] *Put* appeared once in the VOL frame as *put them to bed*, which we determined did not clearly express caused motion.

https://doi.org/10.1017/S0305000904006798 Published online by Cambridge University Press

TABLE 5. *Cue validity of the VOO pattern as a predictor of the meaning of transfer*

Strict encoding of transfer	Inclusive encoding of transfer
0·61	0·94

you got to chew your cookie up). The cue validity of *get* as an indicator of caused motion is thus P('caused motion' | *get*) = 18/108 = 0·17.

Take appears 9 times in the VOL construction with a meaning of 'carry' (e.g. *take their chairs in the house*); the other 5 times it appears in the VOL construction, it has the meaning 'remove.' It occurs another 30 times in the simple transitive construction, none of which conveyed the idea of 'carrying'. For example, twenty-six of these transitive uses involved simple action scenes such as *take a bite/a nap/some pictures*. Thus the cue validity of *take* as a predictor of caused-motion meaning is P(caused-motion | *take*) = 9/44 = 0·20.

Have appears 16 times in the VOL construction, including 15 instances that entail location as in *have something on his head*. One instance of *have* in the VOL pattern was *have a good time at the store*. The vast majority of instances of the 115 instances of *have* appeared in the simple transitive construction, conveying possession (*can I have one of your apples*) or action (*you just had a nap*). In other uses, it appeared with a VP complement conveying deontic modality as in *now you have to wipe your mouth*, or as a perfect auxiliary. The cue validity of *have* in the VOL pattern (as an indicator of caused location) is 15/115, or 0·13.

RESULTS OF AN ANALYSIS OF THE VOO PATTERN

Comparable results exist for the VOO pattern. We first examined whether the formal pattern predicted the meaning of 'transfer.' There were a total of 54 VOO utterances in our database. After initial discussion of criteria to be used, agreement was 100% for classifying mothers' VOO utterances as entailing transfer: literally, metaphorically or not (n = 54). If we include instances that involve metaphorical transfer (specifically those involving the verbs *read* and *tell*), 51 of those 54 (94%) convey transfer. If we exclude instances of *read* and *tell*, 34/54 (61%) code transfer. These figures are given in Table 5, with the breakdown given in Figure 2.[6]

[6] In response to a reviewer's question, we wish to note that the meanings of 'transfer' and 'caused motion' are distinct: one can transfer ownership without causing anything to move (as in giving someone a house), and conversely, one can cause something to move without any transfer of ownership (as it putting one's jacket on a table). There is a

https://doi.org/10.1017/S0305000904006798 Published online by Cambridge University Press

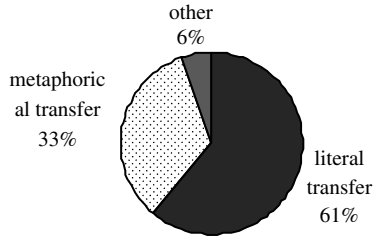


Fig. 2. Proportion of utterances that involved literal or metaphorical transfer.

We then investigated the extent to which the verbs predicted overall sentence meaning. Agreement among the two coders was 97% for classifying mother's uses of verbs as predictive of overall sentence meaning (who did what to whom) (n = 307).

We determined for each utterance whether the predominant meaning ('who did what to whom') conveyed by the verb's use in the VOO pattern held in each of the remaining utterances involving the same verb. For example, we decided that *tell* involved a speaker, a listener, and some kind of content conveyed in all uses of the VOO pattern and also in every other use in our corpus: therefore *tell* has a perfect cue validity of 1. On the other hand *read* involved a reader, a listener and some kind of content in its 3 appearances in the VOO pattern and in 11 other utterances (phrased using 'someone read something to someone'); in the remaining 10 utterances in our corpus, *read* did not involve a listener. Therefore the cue validity of *read* was determined to be: $(3+11)/24 = 0{\cdot}58$.

Table 6 shows our calculations of the cue validities of each of the 13 verbs that appeared in the VOO pattern in our corpus. As in Table 4, Column A lists the number of utterances in which the verb's meaning as it is determined to appear most frequently in the VOO pattern, appears overall in the corpus (including appearances in the VOO pattern). Column B lists the number of total number of times the verb appears in the corpus. Simple cue validity is recorded in Column C and was calculated by dividing column A by column B. To determine cue validity across verbs, it is necessary to weight this number according to how often the verb occurs in the VOO pattern. We therefore multiply the cue validity obtained in Column C by the number of

metaphorical relationship between the two (and in our analyses we are careful to distinguish literal vs. metaphorical interpretations). The two words *put* and *give*, which correspond roughly to the meanings associated with the VOL and VOO patterns respectively, have an LSA correlational rating of only 0·48. Evidence that speakers are able to readily distinguish the two meanings comes from Bencini & Goldberg (2000)'s sorting task, in which subjects readily sorted expressions of caused motion separately from those encoding transfer.

https://doi.org/10.1017/S0305000904006798 Published online by Cambridge University Press

TABLE 6. *Cue validities of verbs in the VOO pattern as predictors of their predominant meaning in the VOO pattern*

verb	A. # of times verb has stable meaning	B. # of instances in corpus	C. Cue validity (A/B)	D. # of times in VOO pattern	Weighted cue validity (C × D/54)
Give	14	14	1	11	0·20
Tell	37	37	1	11	0·20
Get	7	108	0·06	7	0·01
Build	6	24	0·25	5	0·02
Make	3	56	0·05	4	0·00
Read	14	24	0·58	3	0·03
Show	6	6	1	3	0·06
Bring	2	10	0·2	2	0·01
Feed	2	2	1	2	0·04
Pour	2	4	0·5	2	0·02
Buy	1	9	0·11	1	0·00
Fix	1	7	0·14	1	0·00
Call	4	6	0·67	2	0·02
TOTAL		307		54	**0·61**

times the verb occurs in the VOO pattern (recorded in Column D) and divide by the number of VOO tokens in the corpus (54). Summing over the weighted cue validities in Column D provides us with the overall cue validity of 0·61 for verbs in the VOL pattern in our corpus.

The weighted average cue validity for verbs is 0·61. We again see that the overall cue validity of constructions is at least as high as the cue validity of verbs. If we are generous in deciding what utterances involve transfer, the cue validity of the construction is markedly higher than the cue validity for verbs.

Once again there is a wide variability of cue validities across verbs. While a few verbs had perfect cue validities (*feed*, *give*, *show*, *tell*) in our corpus, other verb's cue validities were quite low (*fix*, *get*, *make*). Again, regardless of the overall cue validity of verbs, this fact in itself indicates that attention to the construction's contribution is key to determining who did what to whom.

What about verbs in other constructions? It may be that verbs are more predictive for some constructions than others. For example, in the simple intransitive construction, the verb used supplies almost all of the lexical content. There is a large difference in sentence meaning between *The vase broke* and *She shouted*. Still, even in these cases, the verbs involved are far from perfect predictors of overall sentence meaning. *Break* can appear both transitively and intransitively – to know whether an agent is known or relevant, one needs to know which construction was used. *Shouted*, too, can be

https://doi.org/10.1017/S0305000904006798 Published online by Cambridge University Press

used as a verb of communication (e.g. *She shouted the directions*) or simply a verb of sound emission (e.g. *She shouted for joy*).

It is intriguing that the numbers are so close, for the weighted average cue validities for verbs 0·68 (VOL) and 0·61 (VOO). It is possible that there are stable generalizations about the overall cue validity of verbs: they may be predictive of sentence meaning roughly two thirds of the time. Calculations on other constructions and other corpora are needed to confirm this figure.

CATEGORY VALIDITY

We have discussed cue validity, the probability that an item belongs to a category, given that it has a particular feature: P(cat|feature), and we have found that when the category is taken to be overall sentence meaning, constructions have roughly equivalent cue validity compared with verbs. There is also a second relevant factor. CATEGORY VALIDITY is the probability that an item has a feature, given that the item belongs in the category: P(feature|cat). Thus category validity measures how common or available a feature is among members of a category. The relevant category is again, sentence meaning. In this case, we want to know how likely it is that a given verb or construction will appear, given a particular general sentence meaning.

The category validity of particular verbs as a feature of the semantic category caused motion was determined by hand-coding each utterance in a randomly selected subset of the data analysed above for whether each of the mother's utterances expressed caused motion or not. In our sample, which involved the entire transcripts of four mother's speech, there were 47 utterances that conveyed caused motion involving 12 different verbs. There is often one verb that accounts for the lion's share of tokens of particular constructions (Goldberg, 1999; Sethuraman, 2002; Goldberg, Casenhiser & Sethuraman, 2004); in particular, *put* accounts for many of the tokens of the caused motion construction, but since the transitive and resultative constructions can also convey caused motion (with verbs such as *send*, *bring*, *carry*), the category validity for even *put* is not particularly high; in our corpus; 29/47 expressions conveying caused motion involved the verb *put*, resulting in a category validity of 0·62. The probability that a sentence with caused motion meaning contained the verb *bring* was only 0·02, since only 2% of the utterances expressing caused motion used *bring* (1/47). Similarly low probability was found for another 7 verbs in the corpus (*drive*, *make*, *open*, *ride*, *stack*, *leave*, *wear*). The category validity for *dump* was 0·04, *stick* was 0·06 and *turn* was 0·11.

The average category validity of all verbs that may convey caused motion is equal to 1/n, where n = the number of verbs that express caused motion, or

https://doi.org/10.1017/S0305000904006798 Published online by Cambridge University Press

in our sample, 1/12 = 0·08. Clearly as the sample size increases, the average category validity for verbs is lowered. The actual average category for verbs approaches 0, since more than a hundred different verbs can be used to convey caused motion (n > 100; average cue validity = $1/100^{+} < 0{\cdot}01$). Another relevant number is the MAXIMUM CATEGORY VALIDITY, since the maximum category provides an estimate of the category validity associated with the 'best guess' of a relevant verb. In our sample, *put* had the highest category validity of 0·62; the other 11 verbs conveying caused motion had markedly lower category validities.

The category validity of a construction as a feature of the semantic category caused motion, is the probability that the particular construction will be involved, given the interpretation of caused-motion. There were only three constructions used to convey caused motion in our corpus (the VOL, the resultative and the transitive construction). If we make the very conservative assumption that these three constructions are independent, we find an average category validity of 0·33 for constructions. The average category validity for constructions may also go down as the sample size increases; but since there are less than a handful of constructions that can be used to convey caused motion, the average category would not dip below 0·20. The VOL pattern had the maximum category validity for constructions at 0·83 (39/47 utterances that expressed caused motion involved the VOL pattern). Since *put* is only grammatical in the VOL pattern (**She put on the table*; **She put the book*), the probability that the category 'caused motion' is associated with the feature *put* is the same as the probability that the category 'caused motion' is associated with the features *put* and VOL:

$$P(put \mid \text{'caused motion'}) = P(put \ \& \ \text{VOL} \mid \text{'caused motion'}).$$

Since hundreds of verbs in addition to *put* can appear in the VOL pattern with caused motion interpretation, P(*put* & VOL | 'caused motion') is necessarily less than P(VOL | 'caused motion'). Therefore, P(*put* | 'caused motion') < P(VOL | 'caused motion'). That is, the VOL pattern must have a higher category validity than *put*; it is more *available* as a cue to caused motion meaning. The comparison between verbs and constructions is given in Table 7.

On both measures, the average category validity and the maximum category validity, the construction has a higher score than the verb. All things being equal, if two cues have roughly equal validity, the higher category validity of one cue will naturally result in a greater reliance on that cue in categorization tasks (Estes, 1986; Hintzman, 1986; Bates & MacWhinney, 1987; Nosofsky, 1988). If we assume that the results for the constructions analysed here are representative, it seems that constructions are better cues to sentence meaning than verbs insofar as they are as reliable (with equivalent cue validity) and more available (having higher category validity).

https://doi.org/10.1017/S0305000904006798 Published online by Cambridge University Press

https://doi.org/10.1017/S0305000904006798 Published online by Cambridge University Press

TABLE 7. *Category validity for the constructions and verbs as features of the category of caused motion meaning in a random sample of 47 utterances (to illustrate) and more generally, based on rational assumptions*

Average category validity for verbs and constructions that may be used to convey caused motion:

	In sample of 47 utterances	Asymptotic category validity
n	(n = 12)	(n > 100)
$\sum_{i=1}$ P(**verb**$_i$ | 'caused motion')/n	= 0·08	< 0·01
where verb$_i$ is a verb that may encode caused motion		
n	(n = 3)	(since n < 5)
$\sum_{i=1}$ P(**construction**$_i$ | 'caused motion')/n	= 0·33	> 0·20
where construction$_i$ is a construction that may encode caused motion		

Maximum category validity for caused motion meaning

	In sample of 47 utterances	In general
Verb: P(*put* | 'caused motion')	= 0·62	P(*put* | 'caused motion') = P(*put* & VOL | 'caused motion')
Construction: P(VOL | 'caused motion')	= 0·83	P(*put* & VOL | 'caused motion') < P(VOL | 'caused motion') Therefore, P(*put* | 'caused motion') < P(VOL | 'caused motion')

GENERAL DISCUSSION

Previous experimental work provides evidence consistent with the idea that constructions are at least as good predictors of overall sentence meaning as are verbs. Bencini & Goldberg (2000) compared the semantic contribution of construction with that of the morphological form of the verb, by having subjects perform a sorting task in which they were asked to classify sentences according to 'overall sentence meaning.' Subjects received 16 sentences, created by crossing four verbs (*slice*, *throw*, *get*, *take*) with four constructions (transitive-VO, ditransitive-VOO, caused-motion-VOL and resultative-VOP). Despite a well-documented 1-dimenstional sorting bias which would seem to favor sorts based entirely on verb forms (Medin, Wattenmaker & Hampson, 1987), subjects displayed an equal propensity to sort by construction as by verb. Bencini & Goldberg hypothesized that constructional sorts were able to overcome the one-dimensional sorting bias to this extent because constructions are in fact better predictors of overall sentence meaning than the morphological form of the verb. The implications for language learning are clear. Learners would do well to learn to identify construction types, since their goal is to understand sentences.

Liang (2002) replicated Bencini & Goldberg (2000) with Chinese learners of English of varying proficiencies, finding that subjects produced relatively more construction-based sorts as their English improved. These results indicate that the ability to use language proficiently is correlated with the recognition of constructional generalizations Bencini & Goldberg (2000) has also been replicated with German learners of English (Gries & Wulff, 2004).

Kaschak & Glenberg (2000) demonstrate that subjects rely on constructional meaning when they encounter nouns used as verbs in novel ways (e.g. *to crutch*). In particular they show that different constructions differentially influence the interpretations of the novel verbs. For example, *She crutched him the ball* (ditransitive) is interpreted to mean that she used the crutch to transfer the ball to him, perhaps using the crutch as one would a hockey stick. On the other hand, *She crutched him* (transitive) might be interpreted to mean that she hit him over the head with the crutch. They suggest that the constructional pattern specifies a general scene and that the 'affordances' of particular objects are used to specify the scene in detail. It cannot be the semantics of the verb that is used in comprehension because the word form is not stored as a verb but as a noun. Ahrens (1995) conducted an experiment with a novel verb form. She asked 100 native English speakers to decide what *moop* meant in the sentence *She mooped him something*. 60% of subjects responded by saying that *moop* meant 'give,' despite the fact that several verbs exist that have higher overall frequency than *give* and could be used in that frame, including *take* and *tell*. Similarly, Kako (in press) finds that subjects' semantic interpretations of constructions and their semantic interpretations of verbs that fit those constructions are highly correlated,

https://doi.org/10.1017/S0305000904006798 Published online by Cambridge University Press

concluding as well that syntactic frames are 'semantically potent linguistic entities.' The present paper is the first attempt that we are aware of to quantify and compare the reliability and availability of constructions and verbs as cues to sentence meaning.

A question arises as to why constructions should perform at least as well as verbs in predicting overall sentence meaning. The answer we believe stems from the fact that in context, knowing the number and type of arguments conveys a great deal about the scene at hand. To the extent that verbs encode rich semantic frames that can be related to a number of different basic scenes (Goldberg, 1995), the complement configuration or construction will be as good a predictor of sentence meaning as the semantically richer, but more flexible verb.

The present study suggests an account of why children form the argument structure generalizations they do. We observe that children initially generalize at the level of specific verbs plus argument slots (Tomasello's 'verb islands') because the verb in an argument frame is the best single word predictor of overall sentence meaning. We argue further that children eventually generalize beyond specific verbs to form more abstract argument structure constructions because the argument frame or construction has roughly equivalent cue validity as a predictor of overall sentence meaning as the morphological form of the verb, and has much greater category validity. That is, the construction is at least as reliable and is more available. Moreover, given the fact that many verbs have quite low cue validity in isolation, attention to the contribution of the construction is essential.

REFERENCES

Ackerman, F. & Webelhuth, G. (1998). A theory of predicates. In Stanford: CSLI Publications.

Ahrens, K. (1995). The mental representation of verbs. UCSD.

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review* **98**(3), 409–29.

Bates, E., Bretherton, I. & Snyder, L. (1988). *From first words to grammar: individual differences and dissociable mechanisms*. New York: Cambridge University Press.

Bates, E. & MacWhinney, B. (1987). Competition, variation, and language learning. In B. MacWhinney (ed.), *Mechanisms of language acquisition*. Hillsdale, NJ: Erlbaum.

Bencini, G. & Goldberg, A. (2000). The contribution of argument structure constructions to sentence meaning. *Journal of Memory and Language* **43**(4), 640–51.

Bruce, L. (1988). *Serialization: from syntax to lexicon*. Unpublished manuscript.

Chomsky, N. (1965). *Aspects and the theory of syntax*. Cambridge: MIT Press.

Clark, E. V. (1996). *Early verbs, event types and inflections* (Vol. 9). Mahwah, NJ: Erlbaum.

Dennis, S. & Kruschke, J. K. (1998). Shifting attention in cued recall. *Australian Journal of Psychology* **50**(3), 131–8.

Estes, W. K. (1986). Array models for category learning. *Cognitive Psychology* **18**, 500–48.

Gentner, D. (1982). Why nouns are learned before verbs: linguistic relativity versus natural partitioning. In S. A. Kuzcaj II (ed.), *Language development, Vol 2: Language, thought, and culture* (Vol. 2). Hillsdale, NJ: Erlbaum.

https://doi.org/10.1017/S0305000904006798 Published online by Cambridge University Press

Gentner, D. & Medina, J. (1998). Similarity and the development of rules. *Cognition* **65**(2–3), 263–97.
Goldberg, A. E. (1995). *Constructions: a construction grammar approach to argument structure*. Chicago: Chicago University Press.
Goldberg, A. E. (1999). The emergence of argument structure semantics. In B. MacWhinney (ed.), *The emergence of language*: Erlbaum.
Goldberg, A. E. (to appear). Constructions in context. Oxford University Press.
Goldberg, A. E., Casenhiser, D. & Sethuraman, N. (2004). Learning argument structure generalizations, *Cognitive Linguistics* **15**(3), 289–316.
Gries, S. T. & Wulff, S. (2004). Foreign language learners have constructions: evidence from sorting and corpora. *3rd International Conference of Construction Grammar*. Marseilles, France.
Grimshaw, J. (1990). *Argument structure*. Cambridge, MA: MIT Press.
Healy, A. F. & Miller, G. A. (1970). The verb as the main determinant of sentence meaning. *Psychonomic Science* **20**, 372.
Hintzman, D. L. (1986). ''Schema abstraction'' in a multiple-trace memory model. *Psychological Review, Vol 93*(4), Oct 1986, 1411–1428 American Psychological Assn, US.
Holland, J. H., Holyoak, K. J. & Thagard, R. E. R. (1989). *Induction: processes of inference, learning and discovery*. Cambridge, MA: MIT Press.
Kako, E. (in press). The semantics of syntactic frames, *Language and Cognitive Processes*.
Kaschak, M. P. & Glenberg, A. M. (2000). Constructing meaning: the role of affordances and grammatical constructions in sentence comprehension, 508–29.
Kay, P. & Fillmore, C. J. (1999). Grammatical constructions and linguistic generalizations: the what's X doing Y? construction. *Language* **75**, 1–34.
Kersten, A. W. & Billman, D. (1997). Event category learning. *Journal of Experimental Psychology: Learning, Memory and Cognition* **23**(3), 638–58.
Kruschke, J. K. (1996). Base rates in category learning. *Journal of Experimental Psychology: Learning, Memory and Cognition* **22**(1), 3–26.
Kuno, S. (1972). Functional sentence perspective: a case study from Japanese and English. *Linguistic Inquiry* **3**(3), 269–320.
Lakoff, G. (1970). *Irregularity in syntax*. New York: Rinehart and Winston, Inc.
Lambrecht, K. (1994). *Information structure and sentence form*. CUP.
Leake, D. B. & Ram, A. (1995). Learning, goals, and learning goals: a perspective on goal-driven learning. *Artificial Intelligence Review* **9**(6), 387–422.
Levin, B. & Rappaport. (1995). Unaccusativity: at the syntax–lexical semantic interface. Cambridge, MA: MIT Press.
Liang, J. (2002). How do Chinese EFL learners construct sentence meaning: verb-centered or construction-based? Unpublished masters thesis, Guangdong University of Foreign Studies.
MacWhinney, B. (1995). *The CHILDES project: tools for analyzing talk* (2nd ed.). Hillsdale, NJ: Erlbaum.
Markman, A. B. & Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Psychology* **25**(4), 431–67.
Matsumoto, Y. (1992). *Abstract motion and English and Japanese verbs*. Unpublished manuscript, Tokyo Christian University.
Medin, D. L., Wattenmaker, W. D. & Hampson, S. E. (1987). Family resemblance conceptual cohesiveness, and category construction. *Cognitive Psychology* **12**(2), 242–79.
Murphy, G. L. (1982). Cue validity and levels of categorization. *Psychological Bulletin* **91**(1), 174–7.
Murphy, G. L. (2002). *The big book of concepts*. Cambridge: MIT Press.
Nosofsky, R. (1988). Similarity, frequency and category representations. *Journal of Experimental Psychology: Learning, Memory and Cognition* **14**(1), 54–65.
Pinker, S. (1989). *Learnability and cognition: the acquisition of argument structure*. Cambridge, MA: MIT Press/Bradford Books.

https://doi.org/10.1017/S0305000904006798 Published online by Cambridge University Press

Reddy, M. (1979). The conduit metaphor. In A. Ortony (ed.), *Metaphor and thought*. Cambridge: CUP.
Reinhart, T. (1982). *Pragmatics and linguistics: an analysis of sentence topics*. Bloomington, Indiana: Indiana University Linguistics Club.
Ross, B. H. & Makin, V. S. (1999). Prototype versus exemplar models. In *The concept of cognition*. Cambridge, MA: MIT Press.
Sethuraman, N. (2002). The acquisition of verbs and argument structure constructions, *Linguistics*. San Diego: UCSD.
Talmy, L. (ed.). (1976). *Semantic causative types* (Vol. 6). New York: Academic Press.
Tomasello, M. (1992). *First verbs: a case study of early grammatical development*. Cambridge: CUP.
Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition* **74**(3), 209–53.
Wisniewski, E. J. (1995). Prior knowledge and functionally relevant features in concept learning. *Memory and Cognition* **21**(2), 449–68.

https://doi.org/10.1017/S0305000904006798 Published online by Cambridge University Press