**The nature of generalization in language**

ADELE E. GOLDBERG

*Abstract*

*This paper provides a concise overview of* Constructions at Work *(Goldberg 2006). The book aims to investigate the relevant levels of generalization in adult language, how and why generalizations are learned by children, and how to account for cross-linguistic generalizations.*

*Keywords: constructionist approach, usage-based, argument structure constructions, learning, categorization, statistical preemption, category induction, information structure, islands, scope, linking rules, argument omission, argument realization.*

How do learners acquire generalizations such that they can produce an open-ended number of novel utterances based on a finite amount of input? Why are languages the way they are? In order to address these long-standing questions, many linguists with varying backgrounds have converged on several key insights that have given rise to a family of *constructionist approaches*. These approaches emphasize that speakers' knowledge of language consists of systematic collections of form-function pairings that are learned on the basis of the language they hear around them. This simple idea is, to my mind, what sets constructionist approaches apart from traditional generative grammar. The learning of language is constrained by attentional biases, principles of cooperative communication, general processing demands, and processes of categorization.

The term *constructionist* is intended to evoke both the notion of "construction" and the notion that our knowledge of language is "constructed" on the basis of the input together with general cognitive, pragmatic and processing constraints. It is intended to be a more inclusive term than Construction Grammar, as the latter is a particular instance of a constructionist approach. Constructions themselves are defined as follows:

> Any linguistic pattern is recognized as a construction as long as some aspect of its form or function is not strictly predictable from its component parts or from other constructions recognized to exist. In addition, patterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency (see Chapter 3 for discussion). (Goldberg, 2006:5).

For clarification of the notion of "sufficient frequency" see Gurevich and Goldberg (forthcoming).

*Constructions at Work: the nature of generalization in Language* (hereafter, *CW*) aims to explore the nature of generalizations in language, both

in adult grammar and in children's developing grammar, drawing parallels between our linguistic knowledge and more general conceptual knowledge. This paper aims to provide a thumbnail sketch of the arguments and analyses presented in the book. Sections below correspond to chapters of the book.

**1. Constructions**

Constructions are stored pairings of form and function, including partially or fully filled words, idioms, and general linguistic patterns. Examples are given below:

| | |
|---|---|
| Word | e.g., *tentacle, gangster, the* |
| Word (partially filled) | e.g., *post-N, V-ing* |
| Complex word | e.g., *textbook, drive-in* |
| Idiom (filled) | e.g., *like a bat out of hell* |
| Idiom (partially filled) | e.g., *believe* <one's> *ears/eyes* |
| Covariational Conditional | The Xer the Yer (e.g., *The more you watch the less you know*) |
| Ditransitive | Subj V Obj1 Obj2 (e.g., *She gave him a kiss; He fixed her some fish tacos.*) |
| Passive | Subj aux VPpp ($PP_{by}$) (e.g., *The cell phone tower was struck by lightening.*) |

Table 1. Examples of constructions, varying in size and complexity

Referring to both words and phrasal patterns as *constructions* is intended to highlight their basic commonalities: both are learned pairings of form and function, and both may be completely or partially filled (and completely or partially compositional).

In previous work I have argued that even basic sentence patterns of a language can be understood to involve phrasal constructions (Goldberg 1995). It is implausible to claim that *whoosh* is stored as a verb and yet it appears as a main verb in the attested example in (1). The patterns in (2)-(3) are likewise not naturally attributed to the main verbs involved:[1]

[1] In what follows, I attempt to use attested examples from the web where possible. At the same time, in order to avoid including irrelevant, lengthy and potentially distracting reading material, I have opted to shave the example sentences down, omitting continuations of various sorts. I have also taken the

1. Something has whooshed into my life (forums.digitalpoint.com/)
2. They could easily co-pay a family to death. (NYT, 1/14/02)
3. [The] new Bond squats and crunches his way into purists' hearts. (defamer.com/Hollywood)

While the examples in 1-3 are particularly noticeable, the fact is that the main verb typically underdetermines the overall argument structure of a sentence. Verbs commonly appear with a wide array of complement configurations. Consider the verb *cook* and the various constructions in which it can appear (labeled in parentheses):

4a. The chicken cooked all night. (intransitive inchoative)
b. Pat cooked the steaks. (transitive)
c. Pat cooked the steak well-done. (resultative)
d. Pat cooks. (deprofiled object)
e. Pat cooked Chris some dinner. (ditransitive)
f. Pat cooked her way into the Illinois State bake-off. (*way* construction)

In all of these expressions *cook*'s meaning evokes the notion of preparing food using heat. It is the argument structure constructions that provide the direct link between surface form and general aspects of the interpretation such as something changing state (4a), someone acting on something (4b), someone causing something to change state (4c), someone acting generically (4d), someone intending to cause someone to receive something (4e), someone (metaphorically) moving somewhere despite difficulties (4f) (Goldberg 1995, 2006).

Kaschak and Glenberg (2000) have demonstrated experimentally 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 it 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. Kaschak and Glenberg 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.

---

liberty of inserting periods and correcting spelling where appropriate. I also do include constructed examples when minimally different examples are important to the exposition, or of course, when the sentence in question is ungrammatical. If no url or reference is given, the reader should assume the example was constructed.

Constructions often *underspecify* word order. That is, as emphasized in *CW* (chapter 1), argument structure constructions do not themselves determine the overt order of arguments. The caused-motion construction, for example, in 5a-d below can be captured by a set of grammatical relations, {subject, verb, direct object, oblique}. As intended by the set notation, word order is not specified by this construction. In the case of simple declarative sentences, argument structure constructions combine with Subject-Predicate and VP constructions to specify the order of complements. Other general constructions including a wh-question construction (5a), topicalization construction (5b), heavy NP construction (5c) and adverbial construction (5d) can combine with the caused motion construction to yield alternative orderings of constituents.

| | |
|---|---|
| 5.a. What did she give to the old folks? | waw.emporia.edu |
| b. Into the pail she tossed her books. | |
| c. I gave to him everything he wanted. | www.yourrights.org.uk |
| d. I handed it quickly to Allen. | Allenbooks.google.com |

Learners need to attribute relevant effects to the appropriate constructions, even though multiple constructions co-occur in most utterances. Accounts of our ability to parcel out responsibility from complex structures to their individual subparts has been discussed extensively under the rubric of Bayesian causal networks (Pearl, 1988, 2000; Tenenbaum & Griffiths, 2002). We understand that the water turns on because I twisted the faucet's handle, not because I have a toothbrush in my hand. We likewise come to know that the caused-motion construction contributes the relational meaning that something causes something else to move, while the question construction determines that a *wh*-word appears sentence initially.

The observation that language has an infinitely creative potential (Chomsky 1957, 1965) is accounted for by the free combination of constructions, where constructions can have open slots and underspecified aspects of their overt realization. Constructions are combined (unified) freely to form actual expressions as long as they don't conflict.

## 2. Surface Generalizations

Broader generalizations are revealed by attending to the surface structure of a language than by relating formal patterns to their near paraphrases. While this point may appear to be obvious to those who eschew derivations, the traditional dominance of transformational and derivational frameworks has had some often overlooked implications. For example, many cognitive linguists continue to focus on "dative" or "locative" constructions without recognizing that both are, at least in the case of English, part of a larger "caused-motion" generalization. That is, there is no solid language-internal reason to treat either of the first two sentences as instances of different constructions than the latter four sentences:

| | |
|---|---|
| 6a.She gave a book to him. | "dative" |
| b. She loaded the hay onto the wagon. | "locative" |

c. She tossed the book to him.

d. She tossed the book toward him.
e. She tossed the book toward the wagon.
f. She tossed the book onto the wagon.

That is, (6a-f) are all instances of the same "caused-motion' construction.

Similarly, many linguists continue to treat (regular) ditransitives and "benefactive" ditransitives as distinct constructions because of their distinct paraphrases. However, both types of ditransitive examples pattern alike both semantically and syntactically (see discussion in CW pgs 26-33). *CW* critiques an analysis by a leading defender of derivations (Baker 1997), demonstrating that what appear to be arguments in favor of derivations are often in fact arguments in favor of attention to surface structure instead (p.28-33).

When faced with pairs of sentences that share words and involve different argument structure patterns, similarities are most naturally attributed to the words and differences to the argument structure patterns. Paraphrase relationships are captured by explicit reference to the lexical items, particularly relational lexical items such as verbs.

We need to attend carefully to the way that a construction interacts with other constructions, as generative linguists generally do. These interactions often follow from an analysis of the functions of the particular constructions involved (cf. *CW's* chapter 7 for discussion of many island and scope effects, for example). If constraints on interactions among constructions do not follow from independently motivated facts, then we owe an account of how these conventionalized constraints are learned.

**3. Item specific knowledge and generalizations**

We clearly retain a great deal of specific information about how individual lexical items can be used. There are many items in every language that have highly unusual distributions. In English words like *ago, mere, blithering, dare, rather* and *notwithstanding* have unique distributions (cf. Culicover 1999). Moreover, in our everyday speech, it is often the case that one particular formulation is much more conventional than another, even though both conform to the general grammatical patterns in a language (Pawley and Syder 1983). This quickly becomes clear in cross-linguistic comparisons. For example, in English we ask *do you believe me?,* whereas in Persian one is more likely to ask "did you believe his words?" (*Harfe u ra bavar mikardi*). An English speaker may believe the grass is always greener, while a Persian speaker is more likely feel that the neighbor's hen is a goose (*Morghe hamsayeh ghaz ast*). In English we ask *How are you?* whereas in Persian people routinely ask, how is your health? (*Hale shoma chetore*?). Such idiomatic expressions pervade our everyday speech. Knowing them is part of knowing a language, and clearly their specifics are not determined by universal principles but must be learned on an item-by-item basis (cf. also Jackendoff 2002).

At the same time, it would be a mistake to conclude that speakers do not form generalizations over words, that such generalizations are merely epiphenomenal. If speakers did not necessarily form generalizations over

individual words, we would expect to find languages in which the distribution of words was much less systematic than it is. We do not find languages in which the transitive construction must be characterized anew for each transitive verb. Far from being an arbitrary collection of stipulated descriptions, our knowledge of linguistic constructions, like our knowledge generally, forms an integrated and motivated network. The advantages of learning generalizations are outlined in section 6.

The constructionist approach to grammar allows both broad generalizations and more limited patterns to be analyzed and accounted for fully. In particular, constructionist approaches are generally USAGE-BASED: facts about the actual use of linguistic expressions such as frequencies and individual patterns that are fully compositional are recorded alongside more traditional linguistic generalizations (Barlow & Kemmer, 2000; Langacker, 1988). The usage based approach to grammar allows us to capture more aspects of our knowledge of language, better, than do approaches that deal only in sweeping generalizations on the one hand, or only in the fine grained minutiae of actual usage events on the other.

The relevant point is emphasized by Langacker's rule *vs.* list fallacy. Available evidence suggests that both generalizations ("rules") and item-specific knowledge ("lists") are recorded (cf. discussion in *CW* 2006, chapter 3). Instances are represented at some level of abstraction due to selective encoding; that is, since not all features of an item are represented, the representation is necessarily partially abstract. Moreover, generalizations across instances are also made. Memory for individual instances decays over time, while generalizations that are reinforced grow in strength. Overall, humans organize their knowledge into patterns while retaining an impressive amount of specific information.

In a striking interdisciplinary convergence, a similar position has been developed within the field of categorization. Very early accounts of categories adopted general abstract summary representations; a subsequent wave of "exemplar based" models of categorization held sway in the field for a period following. Most recently, categorization researchers have argued for an approach that combines exemplar-based knowledge with generalizations over that knowledge (Anderson, 1991; Murphy, 2002; Ross & Makin, 1999).

Inheritance hierarchies have long been found useful for representing all types of generalizations. The construction-based framework captures linguistic generalizations within a given language via the same type of inheritance hierarchies used for represented non-linguistic generalizations (Goldberg, 1995, 2003; Hudson, 1990; Lakoff, 1987; Pollard & Sag, 1994). Broad generalizations are captured by constructions that are inherited by many other constructions; subregularities are captured by positing constructions that are at various midpoints of the hierarchical network. Low level constructions represent exceptional patterns. The idea that phrasal patterns *emerge* as generalizations over uses with particular words is briefly outlined below.

**Learning Generalizations**

## 4: How argument structure constructions are learned

A question arises as to where knowledge of constructions comes from. A growing number of linguists and psycholinguists, armed with a better understanding of the power of statistical learning and general categorization processes, are taking another look at the question of whether the generalizations can be learned (Childers and Tomasello 2002; Elman et al. 1996; Kam and Newport 2005; Kuhl 2000; Gomez 2002; Saffran et al. 1996**;** Saffran et al. 1999; Saffran 2001a, 2001b, Saffran and Wilson 2003; Tomasello 2003; Wonnacott et al. 2007). Work in this area has demonstrated that domain-general statistical processes can lead to the learning of subtle linguistic generalizations. The majority of this work, however, has focused on the learning of sounds, words or simple phrase structure rules.

Previous work on phrasal *construction* learning has consistently involved either a form or a meaning that already exists in the ambient language. For example, a number of studies have investigated the learning of the transitive construction (involving English word order or a novel order) (e.g., Akhtar, 1999; Childers & Tomasello, 2001; Wonnacott et al., to appear), determiners with novel form but familiar meaning (Hudson and Newport 1999) or some other familiar constructional meaning (e.g. Kaschak and Glenberg 2004). None of these studies investigate what is involved in associating a novel meaning with a novel phrasal form.

It is possible to address the issue of whether and how novel constructions can be learned from the input directly. The experiments summarized in *CW* chapter 4 train subjects on pairings of novel phrasal forms and novel meanings, while testing learners' ability to generalize beyond the input (see also Goldberg et al. 2004; Casenhiser and Goldberg 2005).

We created a novel construction, associating SOV word order with scenes of APPEARANCE. In a series of experiments, we exposed subjects to sixteen instances of the novel construction, all involving novel verbs. In a forced-choice comprehension task, children (mean age 6;4) were able to associate new instances of the novel construction with its meaning of appearance significantly more often after training than in the control condition in which they watched the same video without sound; they were also able to distinguish the novel construction from instances of the transitive construction, involving new nonsense verbs. Insofar as the construction was truly novel, this simple fact serves to demonstrate that constructions (or novel "linking rules") can be learned and learned quickly. The novel construction learning experiments indicate that people can learn tentative construction-level generalizations over items quickly on the basis of little input.

The novel construction-learning work has also begun to investigate facilitory factors involved in construction-learning. We have demonstrated that a high frequency exemplar facilitates accurate linguistic generalization both in adults and six year olds; that is, holding overall type and token frequency constant, learners are better able to generalize when one type accounts for a large proportion of the tokens. It seems that the high frequency instance acts as an anchor point for the generalization. The implications of this work are

potentially far reaching as tokens of individual constructions are typically centered around a small number of words (often a single word), or around a semantic prototype, even when they potentially occur with a much broader range of words or meanings (Brenier & Michaelis, to appear; Cameron-Faulkner et al., 2003; Deane, 2003; Diessel, 2001; Goldberg, 1996, 1998; Hunston & Francis, 1999; Kidd et al., 2006; Schmidt & Ungerer, 2002; Sethuraman, 2002; Stefanowitsch & Gries, 2003; Thompson & Hopper, 2001; Zwicky, 2002). These corpus findings suggest that exactly this sort of tailor-made input is available to language learners for a variety of constructions. This fact is not unexpected, given Zipf's observations that highly frequent words account for most linguistic tokens (Zipf, 1935).

The fact that the very frequent use of one exemplar in a pattern facilitates the learning of the semantics of that pattern is consistent with the idea that generalizations emerge on the basis of patterns of usage. After hearing many sentences with *put* in the pattern given in 7), children come to associate the meaning of *put* with the pattern even when the verb is not present as in 8):

7) She put a finger on that.
8) He done boots on. (STE, 28 months, Bates et al., 1988)

The result is that the meaning of caused motion (roughly "X causes Y to move $Z_{loc}$") comes to be associated with the Subject Verb Object $Oblique_{path/loc}$ formal pattern. Similar facilitory effects of skewed input have been found in non-linguistic categorization as well (Elio & Anderson, 1984; Goldberg & Casenhiser, 2006). The learning mechanism has been successfully modeled in a connectionist network (Borovsky & Elman, 2006). These parallels motivate viewing construction learning as a type of category learning.

More recent work has produced a number of additional results. See Goldberg et al. 2007 for additional control conditions; Boyd and Goldberg, forthcoming, for a demonstration that both < theme locative verb> and < locative theme verb> orders can be learned and distinguished from one another; we also demonstrate that undergraduates are willing to use the novel construction in production tasks. New research also indicates that the learning involved lasts over a 7-day delay (Boyd, Gottschalk and Goldberg, forthcoming). Work with younger children is still in progress. It will also be interesting to investigate how learners integrate their newly acquired knowledge of a construction with their prior knowledge of other constructions (Abbot-Smith & Behrens, 2006).

**5. How generalizations are constrained**

Children are not strictly conservative, producing only what they have heard, and yet they are not reliably corrected when they produce overgeneralizations. How can children retreat from or avoid overgeneralizations?

Two factors are argued to play a central role: *statistical preemption* or repeatedly witnessing a word in a competing pattern, and a pattern's *degree of openness*: the variability of the items that occur in a given pattern. That is, one

way that overgeneralizations can be minimized is based on the idea that more specific knowledge always preempts general knowledge in production, as long as either would satisfy the functional demands of the context equally well. In the case of morphological preemption (or blocking) this idea is already familiar. While the agentive nominalizing suffix, *–er,* for example, is partially productive (one who is prone to doodling can be called a *doodler*), it does not apply to words for which there already exists an agentive nominal counterpart. For example, while someone can *ref* games, he is not a *reffer*, because *referee* preempts the creation of the new term *reffer*.

Goldberg (1995) argued that a statistical form of preemption could play an important role in learning to avoid expressions such as (9), once a speaker's expectations are taken into account in the following way. In a situation in which an expression such as (9) might have been expected to be uttered, the learner can infer that (9) is not after all appropriate if, consistently, an expression such as (10) is heard instead.

9. ?? She explained him the problem.
10. She explained the problem to him.

Just this sort of statistically based preemption involving related, but non-synonymous constructions has in fact been demonstrated empirically (Brooks and Tomasello (1999) and has been successfully modeled (Alishahi & Stevenson, 2005; Regier, 1996). Cf. also Marcotte (2005) for discussion.

Several theorists have suggested an additional process related to degrees of *entrenchment*. In particular, hearing a pattern with sufficient frequency has been argued to play a key role in constraining overgeneralizations (Braine & Brooks, 1995; Brooks & Tomasello, 1999; Theakston, 2004). For example, Theakston (2004) demonstrated that children in an experimental setting were more likely to overgeneralize verbs that were used infrequently (e.g., to use *giggle* transitively) than verbs that are highly frequent (e.g., to use *laugh* transitively). The difference was attributed to the difference in frequency between verbs like *giggle* and *laugh*. This sort of explanation, however, does not address the fact that verbs that frequently appear in one argument structure pattern can in fact be used creatively in new argument structure patterns, without ill-formedness. For example, *sneeze* is a frequent verb and almost always occurs in an intransitive construction, and yet it is not ill-formed when used in other argument frames such as those in (11) and (12):

11. She sneezed a huge-super-sonic-mind-blowing-ecstatic sneeze...
    www.plotki.net (cognate object construction)
12. She sneezed the water from out of her nose.
    www.writing.com/main/ (caused motion construction)

The preemptive process, unlike the notion of simple high token frequency, does not predict that expressions like (11) or (12) would be preempted by the overwhelmingly more frequent use of *sneeze* as a simple intransitive, because

the different types of expressions do not mean the same things. The finding that high frequency verbs are less likely to be overgeneralized than low-frequency verbs is consistent with the idea that it is preemption that prevents overgeneralization, not the frequency of the verb per se. That is, the preemptive context in which *laugh* might have been expected to occur transitively but instead is witnessed intransitively (in a periphrastic causative construction) occurs more frequently than the same preemptive context for *giggle.* In fact, a google search reveals almost ten times more hits for *made her laugh* than for *made her giggle*. Thus frequency plays a role in the process of statistical preemption exactly because the preemption is statistical. Upon repeated exposures to one construction in lieu of another construction with closely related meaning, the learner can infer that the second construction is not conventional. As noted above, this requires that a given pattern occur with sufficient frequency. Thus, effects that might be ascribed to entrenchment are better attributed to a statistical process of preemption, critically involving the role of semantic or pragmatic relatedness.

*Type frequency/Degree of Openness of a pattern*

Statistical preemption is not sufficient to account fully for children's retreat from overgeneralizations. Constructions may be either too low frequency or too semantically or pragmatically specialized for another construction to effectively preempt their appearance (cf. discussion in Goldberg, 1995 chapter 5). Moreover, upon first encountering a novel verb, speakers presumably know something about the contexts in which it can appear and the contexts in which it cannot appear, without there being a possibility of a preemptive context (since it is a new word).

Several authors have proposed that *type frequency* correlates with productivity (Bybee, 1985; Bybee, 1995; Clausner & Croft, 1997; Goldberg, 1995). Constructions that have appeared with many different types are more likely to appear with new types than constructions that have only appeared with few types. For example, argument structure constructions that have been witnessed with many different verbs are more likely to be extended to appear with additional verbs. To some extent, this observation has to be correct: a pattern is considered extendable by learners if they have witnessed the pattern being extended.

At the same time, the degree of relatedness of new instances to instances that have been witnessed is likely to play at least as important a role as simple type frequency. Constructions that have been heard used with a wide *variety* of verbs are more likely to be extended broadly than constructions that have been heard used with a semantically or phonologically circumscribed set of verbs. That is, learners are fairly cautious in producing argument structure constructions that involve using a familiar verb in a novel way (for a review, see Tomasello 2000); they can only be expected to confidently use a new verb in a familiar pattern when that new verb is relevantly similar to verbs they have already heard used in the pattern. The greater the degree to which previously attested instances fill the semantic or phonological space that includes the

potential target instance, the more confident speakers will be in using the target instance (Brooks & Tomasello, 1999). This factor ("degree of coverage") has been demonstrated to be relevant to induction more generally (Osherson et al. 1990). A combination of both conservative extension based on similarity to a cluster of attested instances and statistical preemption allows us to explain how overgeneralizations are avoided.

In some ways the task of learning would seem to be made easier if speakers never generalized beyond what they had positive evidence for. For example, it would seem to simplify the task if languages used each particular verb it its own particular frame, without generalizing across verbs or using verbs in novel ways. However, in the following section we investigate the benefits of generalizing beyond the input to the level of argument structure constructions.

## 6. Why constructions are learned.

The question of *why* constructional generalizations are learned is not a question that has commonly even been formulated. But if we do not invoke innate syntactic knowledge that simply kicks in upon hearing fixed-in-advance triggers in the environment, the question clearly arises. Cognitive systems do not generalize randomly or completely. As many have emphasized, human categorization is generally driven by some functional pressure, typically the need to predict or infer certain properties on the basis of perceived characteristics (Anderson, 1991; Holland et al., 1989; Kersten and Billman, 1997; Leake and Ram, 1995; Murphy, 2002; Ross and Makin, 1999; Wisniewski, 1995). 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 motivation to be able to predict meaning on the basis of given lexical items and grammatical characteristics (comprehension); conversely, there is motivation to be able to predict the choice of lexical items and grammatical characteristics given the message to be conveyed (production). Because contexts are ever changing, 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.

Both verbs and constructions have the potential to convey the general, event-level interpretation of a clause, roughly "who did what to whom." Since this event level interpretation is a necessary component of interpretation, it is worth comparing the relative contribution of constructions and verbs at this level. Of course, 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.[2]

It is clear that constructions are sometimes better predictors of who did what to whom than many verbs. For example, when *get* appears with a direct

[2] 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*.

object and prepositional phrase, it conveys caused motion (13); when it appears with two postverbal objects, it conveys transfer (14); and when *get* appears with a locative complement it conveys motion (15) :

(13) Dallas got the ball into field goal range (Boston Herald.com)
get + direct object + locative phrase → "caused motion"

(14) We got him a small 20-piece John Deere farm set.
(sirhollands.blogspot.com)
get + direct object + second object → "transfer"

(15) He got out of college (www.caderbooks.com).
get + locative phrase → "motion"

When faced with promiscuous and chameleon-like verbs like *get,* it behooves learners to attend to the constructions involved. That is, *get* can be seen to have low *cue validity* as a predictor of sentence meaning. Cue validity is the conditional probability that an entity is in a particular category, given that it has a particular feature or cue (Murphy, 1982).

In order to quantify the cue validity of verbs and constructions as predictors of overall event interpretation (who did what to whom), a corpus-based study and a classification task were designed. The corpus study had independent coders determine the frequency and the meanings of particular verbs and constructions in the Bates corpus (Bates et al., 1988) on the Child Language Data Exchange System database (MacWhinney, 1995). Analysis revealed that verbs and constructions had roughly equivalent cue validity as predictors of overall sentence meaning under a strict interpretation of what counts as the same meaning for constructions (roughly .6 for both verb and construction). Under a more inclusive criterion that included metaphorical interpretations of constructional meanings, constructions are far better predictors of overall sentence meaning than verbs (with cue validity of .9) (cf. Goldberg, Casenhiser and Sethuraman 2005 for the original study).

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, overall event-level interpretation. Both measures of average category validity and of maximum category validity favor constructions over verbs. All things being equal, if two cues have roughly equal cue validity, the higher category validity of one cue will naturally result in a greater reliance on that cue in categorization tasks (Bates and MacWhinney, 1987; Estes, 1986; Hintzman, 1986; Nosofsky, 1988). Thus 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).

In a separate study, Bencini & Goldberg (2000) performed a classification task that aimed to compare the semantic contribution of the construction with that of the morphological form of the verb when determining overall sentence meaning. The stimuli were sixteen sentences created by crossing four verbs with

four different constructions. Adult subjects were asked to sort these sixteen sentences, provided in random order, into four piles based on "overall sentence meaning." Subjects could sort equally well by verb: e.g., all four instances of *throw* being put into the same pile, regardless of construction; or subjects could sort by construction: all four instances of, e.g., the ditransitive construction being put into the same pile. The stimuli were designed to minimize contentful overlap contributed by anything other than the lexical verb. No other lexical items in the stimuli were identical or near synonyms. Results showed that subjects were just as likely to sort by construction as they were to sort according to the single dimension of the morphological form of the verb (see also Gries & Wulff, 2004; Liang, 2002).

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

In addition, there is a second factor that may well play a role in encouraging speakers to form argument structure constructions. This factor involves the phenomenon of *constructional priming*. Producing or hearing instances of one grammatical pattern primes speakers to produce other instances of the same. Kathryn Bock and colleagues (Bock 1986; Bock and Loebell, 1990; Bock et al., 1992) have shown in a number of experimental studies that passives prime passives, ditransitives prime ditransitives, and instances of the caused motion construction prime other instances of the construction. Structural priming also occurs in natural unmonitored speech or text (e.g., Levelt and Kelter 1982).

Bock's original claim was that syntactic tree structures, not constructions with associated functions, were primed. However, recent work indicates that shared semantic and morphological properties play a role (Chang et al., 2003; Griffin & Weinstein-Tull, 2003; Hare & Goldberg, 1999). Priming has been argued to represent implicit learning in that its effect is unconscious and long-lasting (Bock and Griffin, 2000; Chang et al., 2000). Thus the existence of structural priming may be an important factor underlying the fact that there are generalizations in languages. The same or similar patterns are easier to learn and produce. Once again, the mechanism is not particular to language—repetition of virtually any task results in routinization.

Thus we can see that a learner who did not generalize over instances would be at a distinct communicative disadvantage: learning the meaning of a new word would not tell him anything about how it could be used. In addition, since repetition of the same abstract patterns shows priming effects, learners who failed to generalize would also be at a disadvantage in terms of their processing of language.

**Explaining Generalizations**

The following three sections, which correspond to chapters 7-9 of *Constructions at Work,* focus on the nature of linguistic generalizations in the adult grammar including: island constraints and scope (section 7), subject-auxiliary inversion (section 8) and cross-linguistic generalizations in argument realization (section 9).

**7: Island constraints and scope**

Information structure, or how information is packaged in a clause so as to convey the relevant information status of various propositions, is a complicated topic. While semantics has come into its own as an explanatory force in linguistics, with linguists of all persuasions paying closer attention to lexical and constructional meaning, information structure has been largely left to specialists. Nonetheless, by building on the previous work in this area that does exist, we will see that attempting to bridge the gap between information structure and syntax allows us to begin to unravel some long-standing puzzles often assumed to only be amenable to formal treatments (cf. also Deane, 1991; Erteschik-Shir, 1979; Erteschik-Shir, ; Ioup, 1975; Kluender, 1998; Kluender & Kutas, 1993; Van Valin & LaPolla, 1997). These puzzles include both constraints on long-distance dependencies ("island constraints") and scope phenomena. By recognizing the information structure properties associated with particular constructions, it is possible to predict, to a great extent, which constructions will behave as "islands" with respect to long-distance relations. The "gap" that is identified with the filler constituent cannot be within a constituent that has particularly low discourse prominence (i.e., that is "backgrounded"). This is because the filler constituent in long-distance dependency constructions plays a prominent role in information structure: it is anomalous to treat an element as at once backgrounded and discourse-prominent.

**BACKGROUNDED** elements are defined to be constituents that do not correspond either to the primary topic nor part of the potential focus domain.

| | Example (relevant constituent underlined) |
|---|---|
| PRIMARY TOPIC | She saw a movie. |
| within the potential FOCUS DOMAIN | She left early. |
| BACKGROUNDED elements | The man who she told him about called. |

The idea that certain constructions serve non-discourse prominent functions explains why the same constructions that are islands for "extraction" are also unavailable for containing direct replies to questions (Morgan 1975).

The information structure account does not explain all known facts, but it does predict the following generalizations in a straightforward way:[3]

1. Subject constituents are islands
2. Relative clauses are generally islands.

[3] By way of comparison, only the first two generalizations on the list are wholly accounted for by the standard "subjacency" explanation, without additional stipulation.

3. Complements of manner-of-speaking verbs and factive verbs are islands.

4. Grammaticality judgments should correlate with the degree of "backgroundedness," when length and complexity are held constant (see Ambridge and Goldberg, forthcoming).

5. Direct replies are sensitive to islands (Morgan 1975)

6. Exclamative ah! is sensitive to islands (James 1972)

7. The recipient argument of an active ditransitive, as a secondary topic, resists being a gap.

8. The recipient argument of a passive ditransitive, as a primary topic, is free to be a gap.

9. Presentational relative clauses are not always islands.

10. Definite relative clauses are stronger islands than indefinite relative clauses.

11. Parentheticals are islands.

*A critical role for processing demands*

As alluded to above, there is more than a single factor at work. We know that there is an extra processing load involved when arguments appear in displaced positions relative to their canonical expression (cf. Gibson, 2000; Gibson et al., 2005; Kitagawa & Fodor, 2003; Kluender, 1998; Kluender & Kutas, 1993). This processing burden combines with the information structure clash to result in unacceptability. Thus the ill-formedness of *wh*-words within backgrounded constructions is mitigated in *wh in-situ* languages. At the same time, there should be *some* evidence of the information structure clash even in in-situ languages that permit wh-words within backgrounded constructions. Some tantalizing evidence suggests this may be true (cf. *CW*:151-155 for discussion).

*Topicality and Quantifier Scope*

The information structure properties of constructions predict their predominant assignment of scope. Evidence brought to light by Ioup (1975), Kuno (1991), Michaelis and Francis (forthcoming) is used to argue that relative topicality of arguments predicts their relative scope possibilities. The idea that topicality is intimately related to quantifier scope interpretation can be seen to be intuitive once one examines what it means to have wide scope. A wide scope interpretation of a variable is one in which the variable is given or fixed, and a second variable is interpreted with respect to it. That is, it is within the context provided by the wide scope operator that variables with more narrow scope are interpreted. This is clearly reminiscent of what topics are: the topic is given or fixed, while the comment is predicated of it**.** Chafe notes that topics typically set up "a spatial, temporal or individual framework within which the main predication holds" (Chafe 1976).

This account offers an explanation for the fact that the recipient argument of the ditransitive strongly prefers to have wide scope over the theme argument. That is, while (16) implies that one poor patient was buried in forms, (17) allows for an interpretation that the forms were distributed across patients.

16)The doctor handed one patient every form. One > every

17)The doctor handed every form to one patient. (ambiguous)

The recipient argument in the ditransitive has wider scope than the theme argument (cf. 16), because the recipient argument of the ditransitive construction is more topical. The caused-motion construction in (17) on the other hand, does not constrain either argument to be topical and so either scope assignment is possible.

While the field of information structure is complex and it requires recognizing statistical regularities, a case can be made that information structure and processing are absolutely central to issues that lie at the heart of linguistic theorizing. Ignoring the information structure properties of constructions is like trying to explain a highway system without understanding what the system is designed to do.

**8: Grammatical categorization: Subject-Auxiliary Inversion**

Mainstream generative grammar has traditionally held a belief in "autonomous syntax"—the idea that formal patterns are, in general, most profitably analyzed without reference to their functions. This early view has recently morphed into the much weaker claim that certain purely syntactic generalizations exist. In particular, it has been argued that unrelated functions are associated with the same form (Jackendoff, 2002; Prince, 1996).

In fact, there may not be much to argue about on this point. The notion that the same construction can have different unrelated functions has precedent in lexical ambiguity. The two meanings of *bat*, the flying mammal and the baseball stick, are unrelated. Thus the same form has two different functions, and these two lexical constructions do not share a related meaning, but only form. At the same time, lexical ambiguity is much less common than lexical polysemy: typically, shared form is an indication of *related* meaning.[4] The baseball meaning of *bat*, for example, itself has many different conventional extended senses including the following:

18a. Brian Bruney is at bat.
b. He's batting 50/50.
c. The old woman batted the intruder.
d. She batted out letters of recommendation all morning.
e. They batted ideas about all afternoon.

[4] There is some evidence that learners find ambiguous meanings more difficult to learn, while finding polysemous meanings easier to learn (Casenhiser 2004, 2005). This makes sense insofar as it is easier to extend an existing category than to create an entirely new category.

Since constructionists regard words and phrasal patterns as the same basic type of entity: learned pairings of form and function, we expect that for phrasal constructions as well, polysemy will be the norm, although constructional ambiguity may also exist. The particular case of subject-auxiliary inversion (SAI) is the focus of chapter 8. SAI has been cited as a paragon example of rampant ambiguity; moreover it has been claimed to involve a *system* of syntactic facts (e.g. Newmeyer 2000).

Is English SAI an instance of constructional ambiguity or constructional polysemy? And is there any system of syntactic generalizations that do not make reference to the functions of the constructions involved? I argue that SAI is a polysemous family of constructions, and while it is a conventional and syntactic fact about English that SAI exists in this family of constructions, there is no system of syntactic generalizations about it beyond the basic fact that it exists in English.

The analysis of Subject-Auxiliary inversion (SAI) in *CW* builds on previous accounts that have offered insightful generalizations about its function (Diessel, 1997; Jackendoff, 1972; Lakoff & Brugman, 1987; Michaelis & Lambrecht, 1996). It is argued that several properties including non-positive polarity, non-predicate focus information structure, non-assertive, non-declarative speech act function, and dependent clause status combine to characterize the family of SAI constructions. As is the case with categorization generally, none of these attributes is necessary or sufficient; non-positive polarity is, at the same time, the most heavily weighted attribute (Lakoff & Brugman, 1987). It is this attribute that serves to motivate the form of SAI.

Seven out of 9 construction types discussed in *CW* were argued to involve non-positive polarity. In fact, after *CW* went to press, I realized that an 8$^{th}$ case that I had allowed as an exception to this generalization (as do Borsley & Newmeyer, this volume), that of comparatives, is actually not exceptional at all. It is only possible to get SAI in comparatives when the main clause is positive—the positive main clause yields a negative entailment in the lower, SAI clause:

(3)a. He was faster at it than was she. -> She was *not* as fast as he was: SAI

b. *He wasn't faster at it than was she.

c. He wasn't faster at it than she was. --> She was faster than he was: no SAI

This was an oversight on my part; the finding strengthens the case for the semantic generalization.

*CW* argues that another case that looked to be exceptional, that of exclamatives, is actually strongly motivated as well. While exclamatives are not generally non-positive (*Boy, is this paper getting long*!), it turns out that exclamatives are closely related to rhetorical questions. In fact, a surprising number of exclamatives (13% in a corpus search reported in *CW*) have the overt

tag, *or what?,* and the *or what?* tag only appears with rhetorical questions (and of course it does not always appear with rhetorical questions).[5]

*CW* argues that it is the feature of non-positivity that motivates the form of SAI. It is the first main auxiliary that is stressed in order to convey positive polarity, and it is this auxiliary that is host to sentential negation (Langacker, 1987):

(1a) She *has* been working late. (emphasizing positive polarity)

(1b) She *hasn't* been working late. (host to negative marker)

Since the first main auxiliary is associated with polarity, it is motivated that it is this auxiliary that is inverted in a construction that conveys marked (non-positive) polarity.

SAI appears in a wide array of conventional constructions. It occurs in questions, counterfactual conditionals, sentences with initial negative adverbs, wishes, comparatives and exclamatives, for example. As is the case with the various senses of *bat* provided in 18a-e, each use of SAI must be learned on the basis of positive evidence, since the learner has no way to predict the exact family of constructions that share the formal property of SAI. Moreover, the properties of each construction are not-strictly predictable. For example, the Wish/Curse construction only occurs with the auxiliary *may* (Fillmore 1999):

19. May/*should/*must a thousand fleas infest his armpits!

SAI in comparatives is always optional and it is only allowed in a formal register (cf. 20a, b and 21a, b). There is also a restriction against use of the third person pronoun with SAI in comparatives (22):

20a. Fiddich is more of a slug than Sara is.
b. #Fiddich is more of a slug than is Sara. (informal register)

21a. Prof. Fiddich is more qualified for this position than Ms. Sara is.
b. Prof. Fiddich is more qualified for this position than is Ms. Sara. (formal register)

---

[5] The account in *CW* can further be used to motivate the observation that wh-exclamatives do *not* allow inversion (Jackendoff 2007: n.20):

(21)a. What a great teacher Ann is!
b. *What a great teacher is Ann!

Wh-exclamatives, unlike Y/N exclamatives, are not used as rhetorical questions. Since they are not necessarily non-positive, non-predicate focus, non-assertive or non-declarative, there is no reason to expect them to pattern with the family of SAI constructions.

22. The UIUC library is larger than is *its/Princeton's.

These idiosyncratic facts are evidence in favor of a usage-based constructionist account, since such accounts expect the existence of idiosyncracies (and subregularities).

Chapter 8 emphasizes that a purely syntactic account of SAI has no predictive power beyond stipulating the mere fact that subject-auxiliary inversion exists in English and a few Germanic languages. An account of the distributional properties of subject auxiliary inversion requires recourse to the particular functions of each of the various constructions involved.

There are likely processing and learning advantages to using the same form in more than one construction, although it remains to be seen whether these advantages exist for patterns with unrelated functions and no shared morphology (see the discussion of structural priming in chapter 6). It is ultimately necessary for any descriptively adequate account to list each of the constructions that require or allow SAI and their associated functions, since this set is not strictly predictable. *CW* argues that it may be possible to additionally relate this family of constructions via a network of related functions. But whether it is possible to do more than simply stipulate that SAI exists in English and list each of the SAI constructions remains open to debate (see other papers, this volume). Regardless, the uncontrovertably idiosyncratic facts require a usage-based constructionist account.

**9.Cross-linguistic generalizations in argument realization**

A driving question behind much of linguistic research is, what is the typology of possible constructions and what constrains it? Embracing a constructionist approach to language does not, of course, relieve us of the burden of explaining cross-linguistic generalizations (chapter 9; see also Croft 2001). Quite the contrary, it encourages explanations that go beyond a restatement of the facts. General pragmatic, processing, historical, iconic, and analogical facts, ultimately buttressed by experimental data, shed light on issues related to why languages are the way they are.

Even among generative linguists there has been a trend toward the view that many constraints on language that have traditionally been seen as requiring recourse to innate stipulations that are specific to language can actually be explained by general cognitive mechanisms. For example, the fact that that all languages seem to have noun and verb categories may be explained by the existence of corresponding basic semantic categories of entity and relation (Baker, 2004).

What are the cross-linguistic generalizations about how arguments are linked to syntactic positions, and why do they exist if the generalizations are learned on the basis of the input? Some linguists have claimed the existence of impressive universals in how arguments are expressed; however, such universals are typically claimed to hold only of some underlying level of syntactic representation. And yet such underlying levels are not supported by convincing empirical evidence (cf. *CW* chapter 2). The more modest universals that do exist are often straightforwardly

accounted for by general cognitive, attentional and processing factors (cf. also Croft 2001; Hawkins 2004).

Consider the generalization that the number of semantic arguments tends to equal the number of overt complements expressed (e.g., Lidz, Gleitman and Gleitman 2003). Clearly particular constructions circumvent this general tendency; for example, short passives allow the semantic cause or agent argument to be unexpressed. The modest empirical generalizations that are valid are captured by acknowledging the pragmatics underlying successful communication as captured in the following simple statement:

**Pragmatic Mapping Generalizations**

A) The referents of linguistically expressed NPs are interpreted to be *relevant* to the message being conveyed.

B) Any semantic participants in the event being conveyed that are *relevant* and *non-recoverable* from context must be overtly indicated.

Notice that the pragmatic mapping generalizations make no predictions about semantic participants that are recoverable or irrelevant. This is important because languages and constructions within languages treat these arguments differently. Perhaps the majority of the world's languages readily allow recoverable or irrelevant arguments to be omitted (these arguments are often, but not always, indicated by agreement markers on the verb). Other languages, such as English, generally require such arguments to be expressed (typically by pronouns); and even in English, the "deprofiled object construction" allows certain irrelevant arguments to be omitted (e.g., *Tigers only kill at night*.). Thus the original syntactic claim was too strong, while the pragmatic mapping generalizations are valid cross-linguistically and across constructions within a given language.

Dowty (1991) proposed linking generalizations that are now widely cited as capturing the observable (i.e., surface) cross-linguistic universals in how arguments are linked to syntactic relations. He observed that in simple active clauses, *if* there's a subject and an object, *and if* there's an agent-like entity and an "undergoer," then the agent is expressed by the subject, and the undergoer is expressed by the direct object. Roughly, arguments that are volitional, sentient, causal or moving are agent-like, while arguments that undergo a change of state, are causally affected or are stationary are considered undergoers. Dowty further observes that the opposite mapping appears to be possible in syntactically ergative languages. It is well known that ergative languages themselves are quite complicated and moreover, the grammatical relation of subject is not clearly valid cross-linguistically.

Therefore, clearly the relevant generalizations are again quite modest. Moreover, since the generalization only holds of active clauses, it allows for the fact that the passive construction only optionally expresses an agent argument, and when the agent is expressed, it appears as a non-subject oblique (e.g., a prepositional phrase). A fair generalization, nonetheless, can be rephrased as follows:

**The Salient Participants in Prominent Slots Generalization** (SPPS)
Actors and undergoers are generally expressed in prominent syntactic slots.

The SPPS generalization accounts for the fact that an agent argument without an undergoer, and an undergoer without an agent are also expressed in a prominent syntactic positions; this generalization has the added advantage that it follows directly from well-documented aspects of our general attentional biases.

Humans' attention is naturally drawn to agents, even in non-linguistic tasks. For example, visual attention tends to be centered on the agent in an event, during and after an action is performed (Robertson and Suci, 1980). Infants as young as nine months have been shown to attribute intentional behavior even to inanimate objects that have appropriate characteristics (e.g., motion, apparent goal-directedness) (Csibra et al., 1999); infants habituated to a scene in which a computer-animated circle jumped over an obstacle and contacted another circle, expected the first circle to take a direct route when the obstacle was removed from the scene. Thus, pre-linguistic infants attend closely to the characteristics of agents (volition, sentience, and movement) in visual as well as linguistic tasks.

The undergoer in an event is generally the endpoint of some sort of force (Talmy, 1976; Langacker, 1987; Croft, 1991). The tendency to attend closely to one particular type of endpoint, that of change of state, begins as early as 6 months. Woodward (1998) demonstrates that 6-month-old infants attend more to changes of state than to changes of motion without corresponding state-change. It has been hypothesized that *effects* of actions are the key elements in action-representations both in motor control of action and in perception (Prinz, 1990; 1997). For evidence that undergoers are salient in non-linguistic tasks, see also Csibra et al., 1999; Bekkering et al., 2000; Jovanovic et al. (to appear)). For evidence that endpoints or undergoers are salient in linguistic tasks, see Regier and Zheng (2003), Lakusta and Landau (2005), Lakusta et al. (2007). Thus the observation that agents and undergoers tend to be expressed in prominent syntactic positions follows from general facts about human perception and attention.

Several other generalizations are also discussed in *CW* that draw on existing literature. For example, it is natural to express the meaning of transfer with a ditransitive form because of simultaneous parallels between recipients and patient-objects on the one hand, and possessor-subjects on the other (see also Kemmer and Verhagen 1994). Predictable, recoverable or highly frequent information tends to be reduced in order to make expression more economical. Languages tend to develop fixed word order or case marking in order to avoid rampant ambiguity. Languages tend to have stable head orders due to diachronic processes and processing preferences.

The question is not whether anything at all is specific to human beings and/or hard wired into the brain, but whether there exist rules that are specific to human language and not a result of our general conceptual/perceptual apparatus together with our experience in the world (see Goldberg 2008b). The extant

generalizations about argument realization provide no evidence that such rules are needed.

**10: Variations on a constructionist theme**

As is true of most categories, the category of constructionist approaches includes exemplars that share a family resemblance structure (see Gonzálvez-García and Butler 2006 for an in-depth analysis of this structure). In *CW*, I devote a chapter to comparing several cousin constructionist approaches, including Cognitive Grammar, Unification (Sign-Based) Construction Grammar and Radical Construction Grammar, but the differences between these approaches are less important in my mind than the important methodological and contentful assumptions that are shared, and so I will not dwell on the differences here. Other generative approaches that are occasionally labeled "constructional" in the literature are also discussed in the book, including proposals by Borer (2001); Hale and Keyser 1997; and Marantz (1997); it is argued that while these approaches resemble constructional approaches insofar as they pair some sort of syntactic representation with some sort of semantic representation, they differ from other constructional approaches in critical ways that lead them into serious empirical problems. In particular, it is argued that the generative proposals underestimate the necessity of item-specific knowledge; a combination of morphologically specific and constructional generalizations are required to account for the richness and complexity of language.

**11. Conclusion**

What then can we conclude about the nature of generalization in language? Generalizations are best described by analyzing surface structure instead of positing an underlying level of representation. The generalizations of language, like generalizations in other cognitive domains, are formed on the basis of instance-based knowledge that is retained. Children are able to learn certain kinds of generalizations quite quickly, with skewed input like that commonly found in natural language playing a facilitory role. Associative clusters of instances encourage generalization; generalizations are constrained by the indirect negative evidence children receive involving statistical preemption of non-occurring patterns, together with conservative induction based on the degree of openness. Generalizations at the level of argument structure are made because they are useful, both in predicting meaning and in on-line processing. Classic island and scope phenomena can be accounted for by recognizing the discourse function of the constructions involved. Generalizations that appear to be purely syntactic are at least sometimes better analyzed in terms of constructions insofar as a patterns' distribution is typically conditioned by its functional role. Many cross-linguistic generalizations can be accounted for by appealing to pragmatic, cognitive, and processing facts that are independently required, without stipulations that are specific to language.

Throughout *CW*, there is an emphasis on the function of constructions and on parallels in nonlinguistic domains. Many issues remain outstanding of course. But these powerful tools hold out the promise that we may be able to ultimately account for the complexities of language without appealing to mysterious stipulations.

**References**


Abbot-Smith, Kirsten and Heike Behrens

2006 Known Constructions Influence the Acquisition of Other Constructions: the German Passive and Future Constructions. *Cognitive Science* 30, 995-1026.

Akhtar, N.

1999 Acquiring Word Order: Evidence for Data-Driven Learning of Syntactic Structure. *Journal of Child Language* 26, 339-356.

Alishahi, Afra and Suzanne Stevenson

2005 A Probabilistic Model of Early Argument Structure Acquisition. *Proceedings of the Cognitive Science Society* 82-90.

Ambridge, Ben and Adele E. Goldberg.

To appear. The island status of clausal complements: evidence in favor of an information structure explanation. *Cognitive Linguistics.*

Ambridge, Ben, Julian Pine, Caroline F. Rowland, and Chris R. Young.

In press. The effect of verb semantic class and verb frequency (entrenchment) on children's and adults' graded judgements of argument-structure overgeneralization errors. *Cognition.*

Anderson, John R.

1991 The adaptive nature of human categorization. *Psychological Review* 98, 409-29.

Abbot-Smith, Kirsten, & Behrens, Heike

2006 How Known Constructions Influence the Acquisition of Other Constructions: the German Passive and Future Constructions. *Cognitive Science, 30*(6), 995-1026.

Akhtar, Nameera

1999 Acquiring word order: Evidence for data-driven learning of syntactic structure. *Journal of Child Language, 26*(2), 339-356.

Alishahi, Afra, & Stevenson, Suzanne

2005 A Probabilistic Model of Early Argument Structure Acquisition. *Proceedings of the Cognitive Science Society*, 82-90.

Anderson, John R.

1991 The Adaptive Nature of Human Categorization. *Psychological Review, 98*(3), 409-429.

Baker, Mark

2004 *Verbs, Nouns, and Adjectives: Their Universal Grammar*. Cambridge: Cambridge University Press.

Barlow, Michael, & Kemmer, Suzanne
2000 *Usage Based Models of Grammar*. Stanford: CSLI Publications.

Borovsky, Arielle, & Elman, Jeff
2006 Language input and semantic categories: a relation between cognition and early word learning. *Journal of Child Language, 33*, 759-790.

Braine, Martin D.S., & Brooks, Patricia
1995 Verb argument structure and the problem of avoiding an overgeneral grammar. In Michael Tomasello & William E. Merriman (Eds.), *Beyond names for tings: Young children's acquisition of verbs* (pp. 353-376). Hillsdale, NJ: Lawrence Erlbaum.

Brenier, Jason, & Laura A. Michaelis
to appear Optimization via syntactic amalgam: syntax-prosody mismatch and copula doubling. *Corpus Linguistics and Linguistic Theory, 1*.

Brooks, Patricia, Michael Tomasello, Lewis, William, & Dodson
1999 Children's tendency to overgeneralize their argument structure constructions: the entrenchment hypothesis. *Child Development, 70*, 1325-1337.

Brooks, Patricia, & Tomasello, Michael
1999 How children constrain their argument structure constructions. *Language, 75*(4), 720-738.

Bybee, Joan
1985 *Morphology: A Study of the Relation between Meaning and Form*: John Benjamins Publishing Company.

Bybee, Joan L.
1995 Regular Morphology and the Lexicon. *Language and Cognitive Processes, 10*(5), 425-455.

Cameron-Faulkner, T., Lieven, E., & , M.
2003 A construction based analysis of child directed speech. *Cognitive Science, 27*, 843-873.

Chang, Franklin, Bock, Kathryn, & Goldberg, Adele E.
2003 Do Thematic Roles Leave Traces in their Places? *Cognition*.

Childers, Jane B, & Tomasello, Michael
2001 The Role of Pronouns in Young Children's Acquisition of the English Transitive Construction. *Developmental Psychology, 37*, 739-748.

Clausner, Timothy C., & Croft, William
1997 Productivity and Schematicity in Metaphors. *Cognitive Science, 21*(3), 247-282.

Deane, Paul
1991 Limits to attention: a cognitive theory of island phenomena. *Cognitive Linguistics, 2*(1), 1-63.

2003 Cooccurrence and Constructions. In Luuk Lagerwerf & Wilbert Spooren & Liesbeth Degand (Eds.), *Determination of Information and Tenor in Texts: Multidisciplinary Approaches to Discourse*: Stichting Neerlandistiek, Amsterdam and Nodus Publikationen, Muenster.

Diessel, Holger
1997 Verb-First Constructions in German. In Marjolijn Verspoor & Kee Dong Lee & Eve Sweetser (Eds.), *Lexical and Syntactical Constructions and the Construction of Meaning* (pp. 51-68): John Benjamins Publishing Company.

2004 *The Acquisition of Complex Sentences. Cambridge: Cambridge University Press.*

Elio, Renee, & Anderson, John R.
1984 The effects of information order and learning mode on schema abstraction. *Memory and Cognition, 12*(1), 20-30.

Erteschik-Shir, N.
1979 Discourse Constraints on Dative Movement. In Suzanne Laberge & Gillian Sankoff (Eds.), *Syntax and Semantics* (Vol. 12, pp. 441-467). New York: Academic Press.

Erteschik-Shir, Nomi
1998 *The Dymanics of Focus Structure*. Cambridge: Cambridge University Press.

Gibson, Edward
2000 The Dependency Locality Theory: A distance-based theory of linguistic complexity. . In A. Marantz & Y. Miyashita & W. O'Neil (Eds.), *Image, Language, Brain: Papers from the First Mind Articulation Project Symposium* (pp. 95-126). Cambridge: MIT Press.

Gibson, Edward, Desmet, Timothy, Grodner, Daniel, Watson, Duane, & Ko, Kara
2005 Reading Relative Clauses in English. *Cognitive Linguistics, 16*(2), 313-353.

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

1996 Making One's Way Through the Data. In M. Shibatani & S. Thompson (Eds.), *Grammatical Constructions: Their Form and Meaning* (pp. 29-53). Oxford: Clarendon Press.

1998 The emergence of the semantics of argument structure constructions. In Brian MacWhinney (Ed.), *The Emergence of Language* (pp. 197-212). Hillsdale, N. J.: Lawrence Erlbaum Associates.

2003 Words by Default: Inheritance and the Persian Complex Predicate Construction. In Elaine Francis & Laura Michaelis (Eds.), *Mismatch: Form-Function Incongruity and the Architecture of Grammar* (pp. 84-112). Stanford: CSLI Publications.

2008b. Universal Grammar? Or prerequisites for language language? *Brain and Behavioral Science.*

Goldberg, Adele E., & Casenhiser, Devin
2006 Learning Argument Structure Constructions. In E. Clark & B. Kelly (Eds.), *Constructions in Acquisition* (pp. 185-204). Stanford: Center for the study of Language and Information.

Gries, Stefan Th., & Wulff, Stefanie
2004 *Foreign Language learners have constructions: evidence from sorting and corpora.* Paper presented at the 3rd International Conference of Construction Grammar, Marseilles, France.

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

Hare, Mary, & Goldberg, Adele E.

1999 *Structural Priming: Purely Syntactic?* Paper presented at the Proceedings of the Cognitive Science Society.

Hudson, Richard
1990 *English Word Grammar*. Oxford: Blackwell.

Hunston, Susan, & Francis, Gill
1999 *Pattern Grammar. A corpus-driven approach to the lexical grammar of English*.

Ioup, Georgette
1975 Some universals of quantifier scope. In Kimball (Ed.), *Syntax and Semantics* (Vol. 4, pp. 37-58). New York: Academic Press.

Jackendoff, Ray
2002 *Foundations of Language*. Oxford: Oxford University Press.

Jackendoff, Ray S.
1972 *Semantic interpretation in generative grammar*. Cambridge, Mass.,: MIT Press.

Kidd, Evan, Lieven, Elena, & Tomasello, Michael
2006 Examining the contribution of lexical frequency and working memory to the acquisition of syntax. *Cognitive Development*.

Kitagawa, Yoshihisa, & Fodor, Janet Dean
2003 Default Prosody Explains Neglected Syntactic Analyses in Japanese. In William McClure (Ed.), *Japanese/Korean Linguistics* (Vol. 12, pp. 267-279). Stanford: Center for the Study of Language and Information.

Kluender, Robert
1998 On the distinction between strong and weak islands: A processing perspective. *Syntax and Semantics, 29*, 241-279.

Kluender, Robert, & Kutas, Marta
1993 Subjacency as a processing phenomenon. *Language and Cognitive Processes, 8*(4), 573-633.

Lakoff, George

1987 *Women, fire, and dangerous things : what categories reveal about the mind*. Chicago: University of Chicago Press.

Lakoff, George, & Brugman, Claudia
1987 The semantics of aux-inversion and anaphora constraints. (paper presented to LSA meeting).

Langacker, Ronald W.
1988 A usage-based model. In B. Rudzka-Ostyn (Ed.), *Topics in cognitive linguistics*. Philadelphia: John Benjamins.

Liang, Junying
2002 *How do Chinese EFL learners construct sentence meaning: verb-centered or construction-based?* Unpublished MA thesis.

Marcotte, Jean-Philippe
2005 Causative Alternation Errors and Innate Knowledge: Consequences of the 'No Negative Evidence' Fallacy. In E.V. Clark & B.F. Kelly (Eds.), *Constructions in Acquisition*. Stanford: Center for the Study of Language and Information

Michaelis, Laura A., & Francis, Hartwell S.
forthcoming Lexical Subjects and the Conflation Strategy. In N. Hedberg & R. Zacharski (Eds.), *Topics in the Grammar-Pragmatics Interface: Papers in Honor of Jeanette K. Gundel*. Amsterdam: Benjamins.

Michaelis, Laura A., & Lambrecht, Knud
1996 The Exclamative Sentence Type in English. In Adele E. Goldberg (Ed.), *Conceptual Structure, Discourse and Language* (pp. 375-398). Stanford: Center for the Study of Language and Information.

Murphy, Gregory L.
2002 *The Big Book of Concepts*. Cambridge, Mass.: MIT Press.

Pearl, Judea
1988 *Probabilistic reasoning in intelligent systems: networks of plausible inference*: Morgan Kaufman.

2000 *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.

Pollard, Carl Jesse, & Sag, Ivan A.
1994 *Head-driven phrase structure grammar*. Stanford: CSLI.

Prince, Ellen
1996 *Constructions and the Syntax Discourse Interface*.Unpublished manuscript, University of Pennsylvania.

Regier, Terry
1996 *The Human Semantic Potential: Spatial Language and Constrained Connectionism*. Cambridge, MA: MIT Press.

Ross, Brian H., & Makin, Valerie S.
1999 Prototype versus exemplar models. In R.J. Sternberg (Ed.), *The Nature of Cognition* (pp. 205-241). Cambridge, Mass: MIT Press.

Schmid, Hans-Jorg, & Ungerer, Friedrich
2002 *NP+Complement and NP+BE+Complement as constructions*. Helsinki, Finland: 2nd International Conference on Construction Grammar.

Sethuraman, Nitya
2002 *The Acquisition of Verbs and Argument Structure Constructions.* UCSD, San Diego.

Stefanowitsch, Anatol , & Gries, Stefan Th.
2003 Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics, 8*(2), 209-243.

Tenenbaum, Josh, & Griffiths, TL
2002 Generalization, similarity and Baysian inference. *Behavioral and Brain Sciences, 24*, 629-640.

Theakston, Anna
2004 The role of entrenchment in constraining children's verb argument structure overgeneralisations: a grammatical judgment study. *Cognitive Development, 19*, 15-34.

Thompson, Sandra A., & Hopper, Paul J.

2001 Transitivity, Clause Structure and Argument STructure: Evidence from Conversation. In Joan L. Bybee & Paul J. Hopper (Eds.), *Frequency and the Emergence of Linguistic Structure* (pp. 27-60): Benjamins.

Van Valin, Robert D., & LaPolla, Randy J.
1997 *Syntax: structure, meaning and function*. Cambridge: Cambridge University Press.

Wonnacott, Elizabeth, Newport, Elissa L., & Tanenhaus, Michael K.
to appear Acquiring and Processing Verb Argument Structure: Distributional Learning in a Miniature Language. *Cognitive Psychology*.

Zwicky, Arnold M.
2002, October 2002 *Seeds of Variation and Change.* Paper presented at the NWAV 31, Stanford.