Goldberg, Adele. 2006. Constructions @ Work. Oxford: OUP. [Chapters 1-3]

1

Overview

1.1 Introduction

Constructions—form and meaning pairings—have been the basis of major advances in the study of grammar since the days of the ancient Stoics. Observations about particular linguistic constructions have shaped our understanding of both particular languages and the nature of language itself. But only recently has a new theoretical approach emerged that allows observations about constructions to be stated directly, providing long-standing traditions with a framework that allows both broad generalizations and more limited patterns to be analyzed and accounted for fully. Many linguists with varying backgrounds have converged on several key insights that have given rise to a family of approaches, here referred to as *constructionist approaches*. The term *constructionist* has more than one intended association. The primary motivation for the term is that constructionist approaches emphasize the role of grammatical CONSTRUCTIONS: conventionalized pairings of form and function. In addition, constructionist approaches generally emphasize that languages are learned—that they are CONSTRUCTED on the basis of the input together with general cognitive, pragmatic, and processing constraints.

In an earlier book, *Constructions*, I focused primarily on arguments for adopting a constructionist approach to argument structure and an analysis of several argument structure constructions (Goldberg 1995). The goal of the present book is to investigate the nature of generalization in language: both in adults' knowledge of language and in the child's learning of language. That is, *Constructions at Work* addresses how and why constructions can be learned and how cross-linguistic and language-internal generalizations can be accounted for. Throughout, the function of constructions is emphasized. The present work consists of three parts:

Part I: Constructions: The chapters in Part I combine to provide the theoretical context, including an overview of constructionist approaches

(this chapter), arguments in favor of adopting a constructionist approach to argument structure and clarification of what the approach involves (Chapter 2), and an overview of evidence in favor of a *usage-based* model of grammar that includes both item-specific information and generalizations (Chapter 3). Parts II and III are aimed at accounting for how and why constructions are learned, and why the generalizations that exist take the form that they do.

Part II: Learning Constructions: After investigating what is learned (Chapter 3), Chapters 4–6 address issues surrounding how it is learned (Chapter 4), how generalizations are constrained (Chapter 5), and why constructional generalizations are learned (Chapter 6).

Part III: Generalizations: Chapters 7–9 aim to redress a perceived failure of constructionist approaches to account for language-internal and cross-linguistic generalizations without stipulation. "Island" and scope phenomena are investigated in Chapter 7, and it is argued that a combination of the function of constructions and processing demands accounts for a wide range of facts. Chapter 8 argues that recognizing the function of constructions is essential to accounting for their distribution—the focus of this chapter is English subject–auxiliary inversion. Chapter 9 investigates several cross-linguistic tendencies in argument realization and demonstrates that attention to independently needed pragmatic, processing, and cognitive processes accounts for the data without appeal to stipulations that are specific to language.

Constructionist approaches share certain foundational ideas with the mainstream "generative" approach that has held sway for the past several decades (Chomsky 1957, 1965, 1981). Both approaches agree that it is essential to consider language as a cognitive (mental) system; both approaches acknowledge that there must be a way to combine structures to create novel utterances, and both approaches recognize that a non-trivial theory of language learning is needed.

In other ways, constructionist approaches contrast sharply with the generative approach. The latter has held that the nature of language can best be revealed by studying formal structures independently of their semantic or discourse functions. Ever-increasing layers of abstractness have characterized the formal representations. Meaning is claimed to derive from the mental dictionary of words, with functional differences between formal patterns being largely ignored. Semi-regular patterns and cross-linguistically unusual patterns are generally viewed as "peripheral," with a narrowing band of data seen as relevant to the "core" of language. Mainstream generative theory

argues further that the complexity of core language cannot be learned inductively by general cognitive processes and therefore learners must be hard-wired with knowledge that is specific to language ("universal grammar").

Several basic tenets of constructionist approaches are discussed below. Each represents a major divergence from the mainstream generative approach, and a return in many ways to a more traditional view of language.

Constructions: what they are

ALL LEVELS OF GRAMMATICAL ANALYSIS INVOLVE CONSTRUCTIONS: LEARNED PAIRINGS OF FORM WITH SEMANTIC OR DISCOURSE FUNCTION, including morphemes or words, idioms, partially lexically filled and fully general phrasal patterns.[1] Examples are given in Table 1.1.

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).

Unlike mainstream generative grammar, constructionist approaches tend to emphasize the detailed semantics and distribution of particular words, grammatical morphemes, and cross-linguistically unusual phrasal patterns; the hypothesis behind this methodology is that an account of the rich semantic/pragmatic and complex formal constraints on these patterns readily extends to more general, simple, or regular patterns.

TABLE 1.1. Examples of constructions, varying in size and complexity

Morpheme	e.g. *pre-*, *-ing*
Word	e.g. *avocado, anaconda, and*
Complex word	e.g. *daredevil, shoo-in*
Complex word (partially filled)	e.g. [N-s] (for regular plurals)
Idiom (filled)	e.g. *going great guns, give the Devil his due*
Idiom (partially filled)	e.g. *jog* <someone's> *memory, send* <someone> *to the cleaners*
Covariational Conditional	The Xer the Yer (e.g. *the more you think about it, the less you understand*)
Ditransitive (double object)	Subj V Obj1 Obj2 (e.g. *he gave her a fish taco; he baked her a muffin*)
Passive	Subj aux VPpp (PP_{by}) (e.g. *the armadillo was hit by a car*)

[1] Sag, Wasow, and Bender (2003) prefer to reserve the term construction for combinations of form–meaning pairings, using the term SIGN to refer to individual form–meaning pairings.

As an example of an unusual pattern, consider the Covariational Conditional construction in Table 1.1 above (e.g. *the more you think about it, the less you understand*). The construction is interpreted as involving an independent variable (identified by the first phrase) and a dependent variable (identified by the second phrase). *The* normally occurs with a head noun but in this construction it requires a comparative phrase. The two major phrases resist classification as either noun phrases or clauses. The requirement that two phrases of this type be juxtaposed is another non-predictable aspect of the pattern. Because the pattern is not strictly predictable, a construction must be posited that specifies the particular form and function involved (Culicover and Jackendoff 1999).

Even basic sentence patterns of a language can be understood to involve constructions. That is, the main verb can be understood to combine with an argument structure construction (e.g. transitive, intransitive, ditransitive, etc.). The alternative is to assume that the form and general interpretation of basic sentence patterns of a language are determined by semantic and/or syntactic information specified by the main verb (Grimshaw 1990; Levin and Rappaport Hovav 1995; Pinker 1989). Admittedly, the sentence patterns given in (1) and (2) do appear to be determined by the specifications of *give* and *put* respectively:

(1) Chris gave Pat a ball.

(2) Pat put the ball on the table.

Give is a three argument verb and is expected to appear with three complements corresponding to agent, recipient, and theme. *Put*, another three argument verb, requires an agent, a theme, and a location, and appears with the corresponding three complements in (2). However, while (1) and (2) represent the prototypical case, the interpretation and form of sentence patterns of a language are not reliably determined by independent specifications of the main verb. For example, it is implausible to claim that *sneeze* has a three-argument sense, and yet it can appear in (3). The patterns in (4–6) are likewise not naturally attributed to the main verbs:

(3) "He sneezed his tooth right across town." (Robert Munsch, *Andrew's Loose Tooth*, Scholastic Canada Ltd., 2002)

(4) "She smiled herself an upgrade." (Douglas Adams, *Hitchhiker's Guide to the Galaxy*, New York: Harmony Books, 1979)

(5) "We laughed our conversation to an end." (J. Hart, *Sin*, New York: Ivy Books, 1992)

(6) "They could easily co-pay a family to death." (*New York Times*, January 14, 2002)

(7) "You have to consequate your children when they mess up." (Dr. Phil, forwarded by L. Gleitman, January 27, 2005)

Examples need not be particularly novel to make the point. Verbs typically appear with a wide array of complement configurations. Consider the verb *slice* and the various constructions in which it can appear (labeled in parentheses):

(8) He sliced the bread. (transitive)

(9) Pat sliced the carrots into the salad. (caused motion)

(10) Pat sliced Chris a piece of pie. (ditransitive)

(11) Emeril sliced and diced his way to stardom. (way construction)

(12) Pat sliced the box open. (resultative)

In all of these expressions *slice* means to cut with a sharp instrument. It is the argument structure constructions that provide the direct link between surface form and general aspects of the interpretation, such as something acting on something else (8), something causing something else to move (9), someone intending to cause someone to receive something (10), someone moving somewhere despite obstacles (11), someone causing something to change state (12) (Goldberg 1995).

While English has some dramatic instances in which basic argument structure constructions convey contentful meaning, examples exist in other languages as well. For example, Shibatani notes that in Croatian, the dative subject construction can be used to imply an attitudinal stance, unspecified by any particular word or morpheme in the construction (Shibatani 1999):

(13) Pil-o mi se piv-o
Drink-3SG.PAST I.DAT REF beer-NOM.3SG.NEUT
Lit. "To me, the beer drank itself": real meaning "I *felt like* drinking beer"

In French, certain verbs such as "to slide" can be used intransitively or transitively, with attendant differences in meaning:

(14) Il glisse.
He slid.

(15) Il glisse un livre a Marie.
He slid a book to Marie. (Willems 1997)

In another French construction, the main clause indirect object is interpreted as the subordinate clause possessor:

(16) Le fisc lui estime une fortune de 3 millions de francs.
The IRS to him estimates a fortune of 3 million francs
"The IRS thinks that he has a fortune of 3 million francs." (Koenig 1993)

In Maasai, an External Possessor ("possessor raising") construction allows a second object of the verb to be interpreted as a possessor of the other object:

(17) ɛ-ya al-túngání [en-kitók] [ol-coní] (Payne 1997)
3-take MSG-person.NOM FSG-woman.ACC MSG-skin/ACC
"The person/man will take the woman's animal skin."

Many languages have constructions in which no verb is expressed at all. These cases are prime examples of argument structure constructions, since their meaning cannot naturally be attributed to a (non-existent) verb. For example, Lambrecht (2004) observes a verbless construction ((*et*) NP *qui* VP) in French is used to convey a focused entire proposition:

(18) FOC [tout le monde qui part en weekend]
all the world who leaves in weekend
"Everyone is leaving for the weekend."

Russian also has a pattern where the verb can be omitted in certain motion constructions (as well as in copular constructions), particularly when the manner of motion is not relevant, and when a telic goal of motion is expressed as in:

(19) Kirill v magazin (Chidambaram 2004)
Kirill-NOM to store-ACC
"Kirill goes/will go to the store."

When a telic goal of motion is not expressed, the required meaning can sometimes be coerced by the construction as in:

(20) Kirill iz magazina (Chidambaram 2004)
Kirill-NOM from store-GEN
"Kirill just got back from the store."

The literal meaning of (20) is simply "K. from store," but because the construction requires an end point, an implicit deictic goal ("here") is inferred, resulting in the interpretation, "K. just got back (arrived here) from the store."

German has an interesting verbless construction which conveys incredulity; it involves the morpheme, *und*, normally used to convey simple conjunction:

(21) Larry und Arzt?!
Larry and doctor
"Larry, a doctor?!" (Sailer 2002)

It is always possible to posit a null verb in order to account for these verbless constructions, but such a move seems motivated only by the desire to maintain the position that the main verb determines the overall form and meaning of a sentence. The constructionist approach provides an alternative that avoids positing an unseen verb: the phrasal pattern (with or without any morphological indicators) may specify the main relational predicate of the clause.

Other patterns such as passive, topicalization, questions, and relative clauses are learned pairings of form and function—constructions, as well. Each pairs certain formal properties with a certain communicative function. The fact that various formal versions of these constructions recur cross-linguistically stems from their highly useful communicative functions.

Thus, constructions exist in every language. They are essential to an effective account of both unusual or especially complex patterns and they may be invoked to account for the basic, regular patterns of language as well.

An emphasis is placed on subtle aspects of the way we construe the world

Different surface forms are typically associated with slightly different semantic and/or discourse functions. Take, for example, the DITRANSITIVE construction, which involves the form, Subj V Obj1 Obj2. The ditransitive form evokes the notion of transfer or "giving." This is in contrast to possible paraphrases. For example, while (22) can be used to mean that Liza bought a book for a third party because Zach was too busy to buy it himself, (23) can only mean that Liza intended to give Zach the book. Similarly while (24) can be used to entail caused motion to a location (the book is caused to go to storage), the ditransitive pattern requires that the goal argument be an animate being, capable of receiving the transferred item (cf. 25–26). As is clear from considering the paraphrases, the implication of transfer is not an independent fact about the words involved. Rather the implication of transfer comes from the ditransitive construction itself.

(22) Liza bought a book for Zach.
(23) Liza bought Zach a book. (ditransitive construction)
(24) Liza sent a book to storage.
(25) Liza sent Stan a book.
(26) ??Liza sent storage a book.

In addition to semantic generalizations, there also exist generalizations about INFORMATION STRUCTURE properties of the construction, or the

way in which a speaker's assumptions about the hearer's state of knowledge and consciousness at the time of speaking is reflected in surface form. In particular, there is a statistically reliable tendency for the recipient argument to have already been mentioned in the discourse (often encoded by a pronoun), which is much more pronounced than in the prepositional paraphrases (see Chapter 7). Facts about the use of entire constructions, including register (e.g. formal or informal) and dialect variation, are stated as part of the construction as well. Constructionist approaches provide a direct way of accounting for these facts, since constructions specify a surface form and a corresponding function.

A "what you see is what you get" approach to syntactic form is adopted

Constructionist theories do not derive one construction from another, as is generally done in mainstream generative theory. An actual expression typically involves the combination of at least half a dozen different constructions. For example, the example in (27) involves the list of constructions given in (28)

(27) what did Liza buy Zach?

(28) a. *Liza, buy, Zach, what, do* constructions
b. Ditransitive construction
c. Question construction
d. Subject–Auxiliary inversion construction
e. VP construction
f. NP construction

Note that "surface form" need not specify a particular word order, although there are constructions that do specify word order. For example, the form of the ditransitive construction discussed above is characterized in terms of a set of grammatical relations. The overt order of arguments in (27) is determined by a combination of a verb phrase construction with the Question construction, the latter of which allows for the "theme" argument (represented by *What*) to appear sentence-initially. No underlying levels of syntax, nor any phonologically empty elements are posited.

Constructions are combined freely to form actual expressions as long as they are not in conflict. Unresolved conflicts result in judgments of ill-formedness. For example, the specification of the ditransitive construction that requires an animate recipient argument conflicts with the meaning of *storage* in (26) resulting in unacceptability, unless "storage" is construed to mean the people who work in storage.

1.2 Parts II and III

Parts II and III of this monograph focus on EXPLANATION. Chomsky (1965) rightly points out that any valid theory of language must achieve both descriptive and explanatory adequacy. Descriptive adequacy demands that the theory provide a full and accurate description of the language under study. Constructionist approaches excel at being descriptively adequate, since both generalizations and idiosyncratic particulars can be captured.

There has been ample discussion among linguists as to what should "count" as an explanation in linguistics (Croft 2001; Haspelmath 1999; Jackendoff 2002; Newmeyer 2003; Van Valin and LaPolla 1997). The need for lengthy discussions does not, perhaps, reflect well on the state of our field. One would hope that we would share enough of a paradigm to agree on an explanation when we see one. But since we are not yet at that stage, we enter into a brief discussion of the topic here.

Explanations are generally answers to some "how" or "why" questions. The existence of language poses at least two major questions:

1. **How do learners acquire generalizations such that they readily produce a potentially infinite number of novel utterances based on a finite amount of input?**
2. **Why are languages the way they are?**

There are different general types of explanation potentially relevant to language (Haspelmath 1999). For example, consider the following answers to the question, why does my cellphone have a button that is labeled *dial*?

A. **General–formal explanation:** demonstrates that something is an instance of a larger generalization: e.g. because all push-button phones have a button labeled dial.
B. **General–functional explanation:** identifies the purpose or function: e.g. because people use the word out of habit, and because it is short and fits easily on small instruments.
C. **General–historical explanation:** identifies general constraints on historical change that give rise to the phenomenon: all phones used to have rotary dials, and we often retain words that are no longer synchronically motivated, e.g. we also *hang up* our cell phones, *turn* the channel on TVs, and *cut and paste* our *papers* on the computer.

Mainstream generative grammar has traditionally focused on question (1) and has provided generalizing–formal types of explanations (such as A).

formal generalizations about language are hypothesized to be part of "universal grammar," a genetic endowment of knowledge that is specific to language. It is often assumed that that there are no theoretically interesting reasons for *why* grammars are the way they are beyond formal generalizations. In fact, it is grammar's essentially arbitrary nature, it is argued, that makes it unlearnable, such that critical aspects of it must be hard-wired into our genetic make-up.

Clearly, explanations that generalize (i.e. are independently motivated) are a critical part of explanation for everyone, in any field, since an explanation that refers only to the specific question at hand runs the risk of being ad hoc. General constructs that are independently needed are always preferable. It is for this reason that cognitive and functional linguists prefer explanations that rely on semantic and pragmatic facts, since these sorts of facts are independently required for the sake of interpretation in a way that (underlying) syntactic representations or phonetically null features are not. Moreover, in the same spirit of seeking general explanations, cognitive and functional linguists tend to seek out generalizations that apply beyond language whenever these can be justified; a goal is to posit as little that is specific to language as possible.

Part II: Learning Generalizations

Constructions are understood to be LEARNED on the basis of the input and general cognitive mechanisms. Part II focuses on several interrelated issues related to how and why constructions are learned. In particular, Chapter 3 documents the need for a usage-based model of our linguistic knowledge that allows for input-driven inductive learning. Chapter 4 reports experimental evidence that indicates that constructions can in fact be learned on the basis of the input, and further explores empirical evidence for parallels in the learning of non-linguistic categories. Chapter 5 explores the question of how generalizations are constrained, that is, issues surrounding partial productivity of constructions. Chapter 6 provides motivation for why constructions are learned: exploring the issue of what advantage constructions provide.

Usage-based Models There is ample evidence from research in non-linguistic categorization that ITEM-SPECIFIC KNOWLEDGE EXISTS ALONGSIDE GENERALIZATIONS. Drawing on parallels between general categorization and linguistic knowledge, Chapter 3 emphasizes the need for both item-specific knowledge and generalizations in language. Patterns are stored if they are sufficiently frequent, even when they are fully regular instances of other

constructions and thus predictable (Barlow and Kemmer 2000; Bybee and Hopper 2001; Bybee 1995; Diessel 2001; Goldberg 1999; Langacker 1988; Losiewicz 1992; Thompson and Fox 2004; Tomasello 2003).

For example, there is ample motivation to recognize that we record information about how particular verbs are used in particular argument structure patterns. Evidence comes from several sources: (*a*) issues related to the partial productivity of constructions, (*b*) evidence that children are quite conservative in their productive use of argument structure patterns, and (*c*) evidence that the frequencies with which particular verbs appear in particular argument structure patterns influences speakers' online comprehension.

A usage-based model is needed to account for facts beyond argument structure as well. For example, 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). For example, it's much more idiomatic to say *I like lima beans* than it would be to say *Lima beans please me.* The idiomatic nature of language is made clearly evident by cross-language comparison. For example, in English, a punch in the face can cause a black eye. In German, it's a blue eye: *ein blaues Auge.* An English speaker may boast that she slept like a log or a dog, but a German speaker is more likely to sleep like a woodchuck or marmot (*wie ein Murmeltier schlafen*).[2] An English speaker might complain that someone thinks of themselves as God's gift to the world, whereas in French the complaint would be that the prima donna believed himself sprung from Jupiter's thigh (*se croire sorti de la cuisse de Jupiter*); while an English speaker might bore someone with endless conversation, a French speaker has to avoid holding someone's leg (*tenir la jambe à quelqu'un*). 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).

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 that have long been used for representing non-linguistic generalizations (Goldberg 1995, 2003; Hudson 1990; Lakoff 1987; Pollard and 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.

[2] Examples from <http://german.about.com/library/weekly/aa030899.htm>.

Exceptional patterns are captured by low-level constructions. For example, the "What's <X> doing <Y>?" construction, which has a fixed form and connotes some sort of unexpectedness, captures a subregularity in the grammar of English. It inherits from several other more general constructions, including the Left Isolation, the Subject–Auxiliary Inversion, the Subject-Predicate and the Verb-Phrase constructions (Kay and Fillmore 1999). Language-specific generalizations across constructions are captured via inheritance networks.

Most construction grammars these days are usage-based, due to the sort of evidence to be reviewed in Chapter 3.

How and why constructions are learned Crucially, all linguists recognize that a wide range of semi-idiosyncratic constructions exists in every language, constructions that cannot be accounted for by general, universal, or innate principles or constraints. These include the types of examples given in Table 1.2. Generative linguists argue that these constructions exist only on the "periphery" or "residue" of language—that they need not be the focus of linguistic or learning theorists. Constructionists on the other hand have zeroed in on these constructions, arguing that whatever means we use to learn these patterns can easily be extended to account for so-called "core" phenomena. In fact, by definition, the core phenomena are more regular, and tend to occur more frequently within a given language as well. Therefore if anything, they are likely to be easier to learn. Since every linguist agrees that the "peripheral," difficult cases must be learned inductively on the basis of the input, constructionists point out that there is no reason to assume that the more general, regular, frequent cases cannot possibly be.

TABLE 1.2. Productive or semi-productive constructions that are unusual cross-linguistically and must be learned on the basis of the input

Construction Label	Example (reference)
Time *away* construction	*Twistin' the night away* (Jackendoff 1997b)
What's X doing Y?	*What's that fly doing in my soup?!* (Kay and Fillmore 1999)
Nominal Extraposition	*It's amazing the difference*! (Michaelis and Lambrecht 1996b)
Mad Magazine construction	*Him, a doctor*?! (Lambrecht 1990)
N P N construction	*house by house; day after day* (Williams 1994)
Stranded preposition construction	*Who did he give that to?*
Omitted determiners (and vocatives)	*I don't think, Mac/*cabby, that this is the best way to go.* (Zwicky 1974)

In fact, constructionists argue that language *must* be learnable from positive input together with fairly general cognitive abilities since the diversity and complexity witnessed does not yield to nativist accounts (Culicover and Jackendoff 1999; Elman et al. 1996; Tomasello 2003, forthcoming). Research in this area is quickly gaining momentum. A number of constructionists have made good on the promise to explain how particular constructions are learned (Diessel and Tomasello 2001; Israel 2002). It turns out that the input is not nearly as impoverished as is sometimes assumed (Pullum and Scholz 2002; Scholz and Pullum 2002); analogical processes can be seen to be viable, once function as well as form is taken into account (Goldberg 1999; Israel 2002; Tomasello 2003); there is good reason to think that children's early grammar is quite conservative, with generalizations emerging gradually (Lieven, Pine, and Baldwin 1997; Tomasello 2000, 2003); and the ability to record transitional probabilities and statistical generalizations in the input has proven a powerful means by which to learn certain types of generalizations, including word and phrasal boundaries (Saffran, Aolin, and Newport 1996; Saffran 2001a). In Chapter 4, we demonstrate that novel phrasal constructions can indeed be learned, and learned quickly. This chapter emphasizes parallels with the learning of non-linguistic categories, providing experimental results that demonstrate certain facilitory factors. Chapter 5 addresses the flip side of the coin: how generalizations are constrained. It is argued that children are exposed to constant indirect negative evidence that helps them to recover from overgeneralizations.

In suggesting that language is learned as a type of categorization, we undertake the obligation to explain why generalizations exist. We investigate this question from several perspectives. In Chapter 6, we focus on the advantage both in terms of predictive value and in terms of priming that the learning of constructional generalizations provides. In Part III, we address the existence of generalizations across languages and/or across constructions within a single language.

Part III: Explaining Generalizations

Information Structure and Syntax: 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 (Erteschik-Shir 1979; Erteschik-Shir 1998a; Ioup 1975; Kluender 1998; Kluender and Kutas 1993; Van Valin and LaPolla 1997). These include constraints on long-distance dependencies ("Island constraints") and scope phenomena.

Subject–Auxiliary Inversion: A Natural Category: 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 (Newmeyer 1998). In Chapter 8, we critically evaluate this claim with respect to a case that has been cited as a clear example of a purely formal generalization: subject–auxiliary inversion. We observe that the purely formal generalization fails to predict any of the special properties of the family of constructions that require subject–auxiliary inversion. At the same time, recognizing the family of related functions that the relevant constructions serve, enables us strongly to motivate the distribution of the pattern.

Cross-linguistic Generalizations: A driving question behind much of linguistic research is, what is the typology of possible constructions and what constrains it?

The constructionist approach takes a somewhat different view of what is universal about language than mainstream generative approaches. As Tomasello (2003) observes, what is truly remarkable is the degree to which human languages differ from one another, given that all languages need to express roughly the same types of messages. Constructionist approaches anticipate such fairly wide variability across languages (Croft 2001; Foley and Van Valin 1984; Garry and Rubino 2001).

Nonetheless, there are certain cross-linguistic generalizations that require explanation. For example, constructionist approaches agree that there are recurring semantic prototypes ("conceptual archetypes" in Langacker's terminology) across languages, owing to the fact that humans are overall, more alike than different: we are all born with the same basic conceptual apparatus, with the same basic communicative demands, and we all live in the physical world with forces of gravity, bodies, and night and day (cf. also Lakoff 1987; Webelhuth and Ackerman 1998). Certain semantic/pragmatic functions are so relevant and useful that they can be found in language after language (Croft 2001).

Cross-linguistic generalizations that relate form and function are explained by appeal to general cognitive constraints together with the functions of the

constructions involved. Constructionists turn to grammar-external explanations such as universal functional pressures, iconic principles, and processing and learning constraints to explain such empirically observable cross-linguistic generalizations. For example, certain generalizations about how arguments tend to be expressed cross-linguistically can be explained by appeal to iconic and analogical processes (Givón 1991; Givón and Ute Language Program 1980; Haiman 1985; Lambrecht 1994). Constraints on long-distance dependency constructions (traditional "island constraints") appear to yield to processing explanations that take into account the function of the constructions involved (Erteschik-Shir 1998a; Kluender 1998; Kluender and Kutas 1993). Processing accounts have also been suggested to account for word-order options (Hawkins 1994; Yamashita and Chang 2001). Discourse-pragmatic motivations underlie other sorts of generalizations, such as correspondences between the number of semantic arguments and the number of complements expressed.

Even among generative linguists there has been a trend toward the view that many generalizations about 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 most languages seem to have noun and verb categories may be explained by the existence of corresponding basic semantic categories (Baker 2004). Hauser, Chomsky, and Fitch go so far as to suggest that the only innate ability specific to language that may be required are the recursive interfaces to phonology and semantics, and they allow that even these may turn out not to be specific to language (Hauser, Chomsky, and Fitch 2002). Indeed, a recursive aspect of semantics is required for our (non-linguistic) theory of mind: the fact that I can imagine what you think about what someone else believes. Recursive structure is also evident in familial relationships (the child of a parent can become a parent herself) and in other non-linguistic domains such as stacking and weaving. Thus the ability to use recursion in language may have been an exaptation from possible precursors in other domains.

Constructionist approaches

There is a growing body of work within the constructionist framework, broadly construed: Barðdal (1999); Bates and Goodman (1997); Bencini and Goldberg (2000); Boas (2000); Booij (2002); Choi (2003); Chung (2001); Croft (2001); Davidse (2000); Davies (2005); Dominey and Inui (submitted); Fillmore, Kay, and O'Connor (1988); Fillmore, Michaelis, and Sag (in progress); Fried (2002); Gleitman et al. (1996); Goldberg (1992, 1995); Hovav

and Levin (1998); Iwata (2000); Jackendoff (2002); Kaschak and Glenberg (2000); Kay (2002a); Kay and Fillmore (1999); Lambrecht, (1994); Michaelis (1994, 2004); Michaelis and Lambrecht (1996a); Mughazy (2002); Pizer (1994); Riehemann (1997); Rowlands (2002); Rudanko (2002); Sag (1997); Schmid and Ungerer (2002); Schmid (2001); Schultze-Berndt (1998); Shibatani (1999); Toivonen (2002); Verhagen (2002); Williams (1994); Zadrozny and Ramer (1995); Zhang (1998); Zwicky (1994).

There are of course fine points of divergence even within these approaches. For example, particular constructionist approaches differ in the types of inheritance that are allowed (default or complete), in the type of semantics that is emphasized, and the degree of emphasis on usage and psychological reality. Several differences are discussed in Chapter 10. Also in Chapter 10, certain less closely related approaches that are occasionally labeled "constructional" in the literature are discussed (Borer 1994, 2001; Hale and Keyser 1998; Marantz 1997). It is argued that while these approaches resemble constructionist approaches insofar as they pair some sort of syntactic representation with some sort of semantic representation, they differ from other constructionist approaches in critical ways. It is also argued that the differences inherent in these approaches lead to serious empirical problems.

What makes a theory that allows constructions to exist a "construction-based theory" is the idea that the network of constructions captures our grammatical knowledge of language *in toto*, i.e. **it's constructions all the way down.**

To summarize, constructionist approaches demand answers to the question of how knowledge of language comes to be in the mind of a learner, and also to the question of *why* languages are the way they are: why the generalizations that exist, exist. These are the questions that are the focus of Parts II and III. But before we turn to these questions, we first must address the issue of what level of generalization requires explanation. This is the focus of Chapter 2.

2

Surface generalizations[1]

A mathematician, a physicist, an engineer, and a linguist are trying to decide if all odd numbers are prime. The mathematician says, "one's prime, 3's prime, 5's prime, 7's prime, 9's not prime, so no." The physicist says, "one's prime, 3's prime, 5's prime, 7's prime, 9's not prime, but maybe that's experimental error." The engineer says, "one's prime, 3's prime, 5's prime, 7's prime, 9's prime . . ."

The linguist says, "one's prime, 3's prime, 5's prime, 7's prime. Aha! We have a universal generalization. Nine doesn't seem to be prime, but it MUST *be prime at some underlying level of representation!"*

(Joke told by Arnold Zwicky during his Presidential Address at the Linguistic Society of America, 1992)

Since the earliest days of generative grammar, there has existed a strong tendency to consider one construction in relation to a particular rough paraphrase. Initially this was a result of the emphasis on transformations that derived one pattern from another. While today there exist many non-derivational theories for which this motivation no longer exists, the traditional outlook has not completely lost its grip, as can be seen from continuing focus on partial or incomplete generalizations such as the "dative" construction or the "locative" alternation. This chapter argues that it is profitable to look beyond alternations and to consider each surface pattern on its own terms. The arguments in this chapter extend those presented in Goldberg (1995) in addressing the advantages of non-derivational accounts of argument structure (cf. also Michaelis and Ruppenhofer 2001); at the same time, the analysis of particular constructions discussed is consistent with that offered in my earlier work.

It is observed that when considering instances of the *same* surface pattern involving *different* words, similarities should be attributed to the surface

[1] I am grateful to Tsuguro Nakamura, Laura Michaelis, and Hideyuki Sugiura for helpful discussion on this topic and also to Tsuguro and an anonymous reviewer for extensive comments on an earlier draft. This chapter is based on Goldberg (2002).

pattern and differences to the different verbs and arguments involved. A specific proposal by Baker (forthcoming) to derive certain constructions from their rough paraphrases is critiqued below (Section 2.4) in order to demonstrate that what appear to be arguments in favor of derivations are often in fact arguments in favor of attention to surface structure instead.

2.1 What is meant by surface form?

In this section we clarify what is meant by *surface form*. Surface form need not specify a particular word order, nor even particular grammatical categories, although there are constructions that do specify these features. Adopting the notation of Goldberg (1992, 1995) we might characterize the ditransitive construction as in fig. 2.1.

The first line provides the semantics of the construction. The ditransitive involves a predicate with three arguments; these three arguments are labeled "agent," "recipient," and "theme" for convenience but there is no assumption that these thematic role labels are drawn from a universal or limited set. Instead the roles are determined by the meaning of the construction. In this case the main predicate is "CAUSE-RECEIVE" or more informally "give," and the three argument roles correspond to the three major entities involved in the semantics of giving.

As is the case with other constructions, including words and morphemes, constructions typically allow for a range of closely related interpretations. The "CAUSE-RECEIVE" predicate associated with the ditransitive construction is subject to systematic variation depending on which verb class it interacts with. For example, the construction can be used to convey "intention to cause to receive" when used with verbs of creation; "refuse to cause to receive" when used with verbs of refusal, etc. See Bryant (2005); Goldberg (1992, 1995); Kay (2001); Koenig and Davis (2001); Leek (1996); for details and slightly differing analyses.

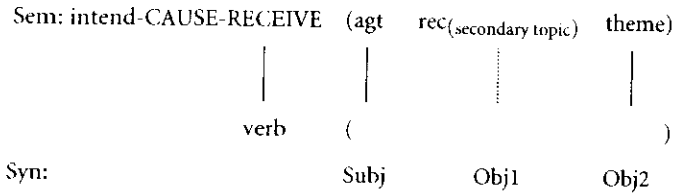


FIGURE 2.1. The ditransitive construction

As indicated in Fig. 2.1 by the lines between the argument roles of the construction and the role array of the verb, the verb and its own arguments are integrated (unified, "fused") with the predicate and arguments of the construction. Solid lines indicate that the argument role of the construction must fuse with an independently existing participant role of the verb. Dashed lines indicate that the argument role of the construction may be contributed by the construction without a corresponding role existing as part of the inherent verbal meaning. That is, a corresponding participant role of the verb may exist, but need not.

Information structure properties of constructions are specified by subscripts. The specification noted is that the recipient argument is a secondary topic. This statistical generalization as well as important syntactic implications of this generalization are discussed in depth in Chapter 7 (Section 7.3).

Finally, the linking of roles to grammatical relations is provided. See Goldberg (1995: ch. 4) for arguments that both generalizations and exceptional mappings can be captured by positing construction-specific linking generalizations when constructions are related within an inheritance hierarchy.[2]

As discussed in Chapter 1, it is important to realize that reference to form in the definition abstracts away from specifics of surface form that can be attributed to other constructions. That is, an actual expression typically involves the combination of many different constructions. For example, the expression in (1) involves the eleven constructions given in (2):

(1) A dozen roses, Nina sent her mother!

(2) a. Ditransitive construction
b. Topicalization construction
c. VP construction
d. NP construction
e. Indefinite determiner construction
f. Plural construction
g. *dozen, rose, Nina, send, mother* constructions

[2] By appealing to grammatical relations instead of grammatical categories in the syntax of this construction, we do not intend that grammatical categories are irrelevant in general, contra what is assumed in a critique by Newmeyer (forthcoming). In the present case, grammatical relations are found to be more perspicuous because they serve to distinguish the ditransitive from the construction involved in (i):

(i) She considered him a fool.
(ii) She considered him crazy.

Expression (i) is an instance of a construction that has the grammatical relations: Subj V Obj PRED (see Gonzálvez García 2000). Both predicates and second objects can appear as NPs. PRED, however, can also be realized as an AP, as in (ii).

Constructions are combined freely to form actual expressions as long as they can be construed as not being in conflict (invoking the notion of construal here is intended to allow for processes of accommodation or *coercion*; see Kadmon 2001; Michaelis 2004). Thus, the same ditransitive construction is involved in active declarative form as well as in topicalized, clefted, or questioned forms. That is, the recipient argument is an Object whether or not it appears directly after the verb or whether it appears as a distantly instantiated topicalized NP. It is, for example, the (non-echo) question construction that determines the fact that the wh-word appears sentence-initially in English.[3]

Constructional approaches share with mainstream generative grammar the goal of accounting for the creative potential of language (Chomsky 1957, 1965). That is, it is clear that language is not a set of sentences that can be fixed in advance. Allowing constructions to combine freely as long as there are no conflicts, allows for the infinitely creative potential of language. At the same time, constructional approaches generally recognize that grammars don't generate sentences, speakers do. That is, a speaker is free to creatively combine constructions as long as constructions exist in the language that can be combined suitably to categorize the target message, given that there is no conflict among the constructions (cf. Fillmore 1975; Langacker 1987a: ch. 2 for discussion).

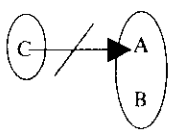
2.2 The Surface Generalization Hypothesis

Many theoretical approaches today eschew the need for any kind of transformation or derivation (Bresnan 1982, 1994; Fillmore et al. forthcoming; Johnson-Laird 1968; Lakoff 1987; Langacker 1987a, 1991; Pollard and Sag

[3] Given the syntactic specifications of the ditransitive construction given in fig. 2.1, a separate but related construction is required to account for passives of ditransitives since such passives do not involve the same linking of grammatical functions to roles. Supporting this idea that there exists a passive ditransitive construction is the fact that the actual form of the passive ditransitive is not strictly predictable. At one time in the history of English, only impersonal passive ditransitives were allowed (Denison 1993). In some languages, both the recipient and patient arguments can passivize, whereas in English only the recipient argument can be passivized (Alsina and Mchombo 1990; Polinsky 1998). In addition, as discussed in Chapter 8, the information structure properties of the passive ditransitive are distinct from that of the active ditransitive. The fact that there is something non-predictable about the passive ditransitive entails that a construction be posited. If it were possible to predict the specifics of passive-ditransitive expressions in some way, an alternative route would be possible. The alternative would be to define the ditransitive construction more abstractly such that it would not specify that there are two objects overtly realized, nor the specifics of the mapping between thematic roles and grammatical functions; instead the only syntactic or linking specification would be that there is an extra object (Kay 1997). In this way, it would be possible to unify the highly abstract "extra object" construction with passive without positing an additional ditransitive-passive construction. See also work by Bryant that aims to provide underspecified characterizations of argument structure constructions such that they unify with passive, active, raising constructions to yield surface forms (Bryant 2005).

1987). A compelling reason to avoid positing derivations in favor of an emphasis on surface form is simply that there are typically powerful generalizations surrounding particular surface forms that are more broad than those captured by derivations or transformations. We refer to these broader generalizations as SURFACE GENERALIZATIONS. The present chapter focuses on the domain of argument structure; the surface formal and semantic/pragmatic generalizations in this domain are captured by ARGUMENT STRUCTURE CONSTRUCTIONS: pairings of form and function that are used to express basic clauses. Several case studies are considered including the "dative" construction and the "locative alternation." It is argued that these traditional divisions underrepresent the generalizations that exist. We address the question of how to account for paraphrase relations, as well as how to account for various differences between instances of the same argument structure construction, in Section 2.8. In this section we review an important historical precedent for the form of argument made here.

Despite being the most influential architect of transformations and later, derivations, Chomsky (1970) put forward one of the most well-known and widely accepted arguments against deriving one subset of data from another. His argument was based on Surface Generalizations. In particular, he demonstrated that NPs based on "derived" nouns (i.e. nouns that have verbal counterparts) have exactly the syntax of NPs based on underived nouns. In particular they both have the same internal and external syntax. Both types occur with the full array of determiners, often pluralize, and take complements marked with *of*. Both types can appear as the subject of passives or can be distantly instantiated by a question word. To avoid an account in which this is mere coincidence, Chomsky reasoned, we need to recognize that both types are base-generated as nouns instead of attempting to derive certain NPs from clausal counterparts (Lees 1960). With Williams (1991), we might call this the "**target syntax argument**": it is preferable to generate A directly instead of deriving it from C if there exists a pattern B that has the same target syntax as A and is clearly not derived from C.



Williams (1991) makes a parallel "target semantics argument." He observes that the meanings of NPs based on underived nouns fall into the same set of categories as the meanings of NPs based on "derived" nouns. For example,

extent, temporal duration, and evaluative states can be predicated of both "derived" and underived nouns (1991: 584):

Extent

(3) a. The destruction of the city was complete. Potentially derived
b. The carnage was complete. Underived

Temporal duration

(4) a. The destruction of the city took four hours. Potentially derived
b. The war took four hours. Underived

Evaluative state

(5) a. The destruction of the city was horrible. Potentially derived
b. The war was horrible. Underived
(not just the fact of the war, but the way the war was)

At the same time, Williams observes, the range of NP meanings is distinct from the range of S meanings, as seen in examples (6)–(8) (1991: 585):[4]

Extent

(6) *That the city was destroyed was complete.

Temporal duration

(7) *That the city was destroyed took four hours.

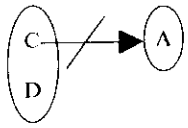
Evaluative state

(8) That the city was destroyed was horrible.
(can mean that the fact that the city was destroyed was horrible, but not that the way it was destroyed was horrible)

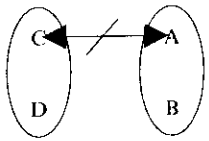
In short, given that the syntax and semantics of derived nouns are like those of underived nouns, and unlike the syntax and semantics of clauses, it is clearly simpler to allow the nouns to be base-generated as nouns as opposed to deriving them from clause structures.

Beyond target syntax and target semantics arguments are what are referred to below as "input" syntax and semantics arguments. In particular, one should not attempt to derive A from C if there exists a pattern D that has the same syntax and semantics as C and yet cannot serve as input from which to derive A.

[4] In this, Williams echoes Langacker (1987b) insofar as the semantics of Ns are argued to be distinct from that of Ss.



When both target syntax/semantics and input syntax/semantics hold, the relationship between possible input and possible output is symmetric, providing an argument against deriving either type of pattern from the other.



The arguments put forth by Chomsky (1970) (and Williams 1991) have been robust. For more than three decades, the field has resisted the temptation to derive deverbal NPs from clauses. What is less widely acknowledged is that parallel arguments hold in the domain of argument structure. These arguments support the idea that each argument structure pattern is best analyzed on its own terms, without relying on explicit or implicit reference to a possible alternative paraphrase. It is argued that such reliance effectively puts blinders on, and limits a theory's ability to state the full extent of the relevant generalizations.

We might label the hypothesis that the target syntax and target semantics arguments and the input syntax and semantics arguments hold in general for argument structure patterns, the *Surface Generalization Hypothesis.*

Surface Generalization Hypothesis: there are typically broader syntactic and semantic generalizations associated with a surface argument structure form than exist between the same surface form and a distinct form that it is hypothesized to be syntactically or semantically derived from.

Support for the Surface Generalization Hypothesis provides substantial motivation for the assumption that the syntax of argument structure should be represented without recourse to derivations. It also suggests that it is possible to overplay the importance of alternative forms (paraphrases).[5]

[5] Bolinger, an early advocate of the Surface Generalization Hypothesis put the problem with ignoring semantic differences between alternative formal patterns this way:

[It is often considered normal] for a language to establish a lunacy ward in its grammar or lexicon where mindless morphs stare vacantly with no purpose other than to be where they are....

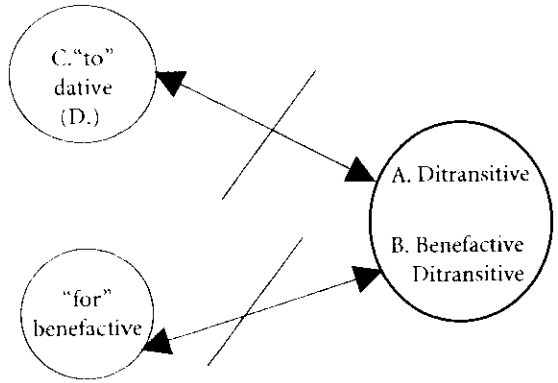
2.3 The Ditransitive construction

Many generative theories derive the two ditransitive or double object expressions in (9) and (10) from distinct input expressions on the left, which correspond to their rough paraphrases (Baker 1988; Larson 1988):

(9) Mina bought a book for Mel.→ Mina bought Mel a book.

(10) Mina sent a book to Mel.→ Mina sent Mel a book.

Even certain constructionist approaches treat the two examples on the right as instances of two independent constructions (e.g. Jackendoff 1990; Kay 2001). However, both instances of the ditransitive share many properties with each other and differ systematically from their paraphrases (see also Langacker 1991; Oehrle 1975). That is, there are good reasons to group the two "outputs" together as distinct from the "inputs" as follows:



Let us focus on the generalization boldfaced in the diagram above: namely that ditransitives paraphrasable with "to" form a class with those that are paraphrasable with "for." Similarities between the two proposed types of ditransitive expressions (A and B in the diagram above) begin with their shared surface form; in its simple active form, the ditransitive involves an active verb followed by two objects. In both cases questioning the recipient argument is less than fully acceptable (11a,b, in Table 2.1); adverbs may not separate the two object arguments in ditransitives (12a,b); both types of

{C}ontemporary linguistics has carried the fantasy to new heights, and expanded it with a new version of an old vision, that of synonymy: not only are there mindless morphs, but there are mindless differences between one construction and another. (Bolinger 1977)

See also Haiman (1985); Johnson-Laird (1967, 1968); Langacker (1987a); Wierzbicka (1988).

ditransitives prefer the recipient argument to be more given or topical than the theme argument (cf. 13a,b; see Chapter 7 for discussion). Semantically, both so-called *for* ditransitives and so-called *to* ditransitives require that the recipient argument be construed to be animate (14a,b) (Green 1974; Oehrle 1975; Partee 1965/1979).[6] More generally, as noted in Chapter 1, the ditransitive evokes the notion of "giving" in various ways, depending on the verb class involved (15a,b). Each of these parallels are represented in the left-hand column of Table 2.1. The paraphrases are shown to be distinct in each of these ways, in the right-hand column.

Other interpretations for the ditransitive can also be systematically related to the notion of giving, in that they may imply that transfer will occur if certain satisfaction conditions evoked by the main verb occur (16a), that transfer will *not* occur (16b), or that the antonymic relation of giving, that of taking away, occurs (16c).[7]

Ditransitives, whether paraphrasable with "to" or "for," pattern alike and differently from their prepositional paraphrases:

TABLE 2.1. Ditransitives pattern alike (left) and differently than their prepositional paraphrases (right)

Ditransitives : Subj V Obj Obj2 (paraphrasable with "to" or "for")		Paraphrases
(11)	a. ??Who did Mina buy a book?	Who did Mina buy a book for?
	b. ??Who did Mina send a book?	Who did Mina send a book to?
(12)	a. *Mina bought Mel yesterday a book.	Mina bought a book yesterday for Mel.
	b. *Mina sent Mel yesterday a book.	Mina sent a book yesterday to Mel.
(13)	a. ??Mina bought Mel it.	Mina bought it for Mel.
	b. ?? Mina sent Mel it.	Mina sent it to Mel.
(14)	a. ??Mina sent that place a box.	Mina sent a box to that place.
	b. ??Mina bought that place a box.	Mina bought a box for that place.
(15)	a. Mina bought Mel a book. (Mina intends to give Mel the book)	Mina bought a book for Mel. (the book could be intended for Mel's mother, bought by Mina because Mel was too busy to buy it)
	b. Mina sent Mel a book. (Mina intends to give Mel the book)	Mina sent a book to storage.

[6] See Goldberg (1992, 1995: 146–7) for arguments that the first Object in *The paint job gave the car a higher sales price* is based on a Causal Events as Transfers metaphor. The constraint that the recipient must be animate holds of the source domain of the metaphor.

[7] Goldberg (1995: 150) argues that even instances of such relatively marked examples such as *Cry me a river* can be related to the notion of giving via a metaphorical extension.

(16) a. Mina guaranteed/offered Mel a book. (If the guarantee or offer is satisfied, Mel will receive a book)
b. Mina refused Mel a book. (Mina caused Mel not to receive a book)
c. Mina cost Mel his job. (Mina causes Mel to lose his job)

It has been suggested that the existence of variable meanings undercuts the claim of a unified construction (Nakajima 2002). The criticism stems from the belief that the concepts of, for example, giving, not giving, and taking away cannot naturally be classed together. However, it is clear that both the negation and the antonym of a particular concept are closely associated with that concept. For example, a concept and its antonym typically serve as strong associates for one another in psycholinguistic studies (Meyer and Schvaneveldt 1971): e.g. *hot* primes *cold*, *high* primes *low*, and *giving* primes *taking away*. Negated sentences typically presuppose that the corresponding positive assertion has been asserted or might be believed in the particular context of use (Horn 1989). In this way we can see that giving, not giving, and taking away *are* in fact closely associated concepts.

The existence of a corresponding passive has been thought to differentiate ditransitives into two types; it has been claimed that only those with paraphrases involving *to* can be passivized (Fillmore 1965; Kay 2001). While it may be true that ditransitives that have paraphrases with *to* show a statistical tendency to passivize more easily than those that have paraphrases with *for*, the generalization is far from clear-cut as many have observed (Culicover and Wexler 1973; Erteschik-Shir 1979; Oehrle 1975). For example, the following examples appear to be equally acceptable, or if anything, (17a) is more acceptable than (17b) despite the fact that only (17b) is paraphrasable with *to*:

(17) a. Mel was cooked a fine dinner by the new chef. (cf. The new chef cooked a fine dinner for Mel.)
b. Mel was tossed a blanket by the babysitter. (cf. The babysitter tossed a blanket to Mel.)

Thus we see that ditransitive expressions pattern alike on a number of syntactic and semantic dimensions regardless of their potential paraphrases. Still, perhaps there are empirical facts that would indicate that a derivational relationship is important. A particular proposal is discussed below.

2.4 A Derivational Proposal

Many generative grammar proposals invoke underlying structure to capture paraphrase relations, although there has been a trend toward recognizing surface-level constructional generalizations, even within this general framework (Harley 2002; Marantz 1997). While it is perhaps impossible and certainly beyond the scope of the present chapter to try to critique all possible derivational accounts, it is worth examining at least one in more detail to understand both the possible allure of derivations and the underlying empirical issues that arise. So as to be certain not to argue against a straw man, I present recent arguments by Baker (forthcoming), a leading proponent of derivational accounts of argument structure.

Baker assumes that all and only sentences with the same truth-conditional semantics must be derived from a common source (the UTAH hypothesis). Thus Baker does not try to defend a derivational account for *all* argument structure alternations. Recognizing the semantic difference between variants of the "load/spray" alternation, for example, he allows that each variant should be base-generated; the two surface forms then reflect two different possible construals of an event (following e.g. Dowty 1991; Goldberg 1995; Pinker 1989; Rappaport and Levin 1985).

As a prime example of an alternation that is claimed to support the derivational approach, Baker offers the ditransitive and its prepositional paraphrase.[8] The assumption is that there is no semantic difference between the ditransitive and its prepositional paraphrases. However, as we have already seen, only in the ditransitive is the requirement that transfer be intended operative (18a). In paraphrases with *for* a larger set of benefactive relationships are possible (Goldberg 1992, 1995; Green 1974; Oehrle 1976), as is evident in (18b):

(18) a. *Bill baked Chris a cake, but never intended for her to have the cake.
b. Bill baked a cake for Chris, but never intended for her to have the cake—instead he did as a favor for Chris because Chris was too busy to bake it herself.

Thus if we accept that so called *to* ditransitives and *for* ditransitives should be treated under the same general rubric as argued above and assumed by Baker as well, then we can see that the basic assumption of synonymy underlying the derivational proposal is flawed from the outset.

But perhaps there are compelling syntactic motivations for a derivational account. Baker suggests that the recipient ("goal") argument has several special restrictions that distinguish it from other types of objects: it cannot be the subject of a depictive predicate (19a), it cannot be the non-head of a compound (19b), it cannot undergo "wh-movement" (19c), or heavy NP shift

[8] Levin and Rappaport Hovav (2003) similarly suggest that the "load/spray" alternation be accounted for by positing two distinct semantic representations, while the ditransitive and prepositional paraphrase, they argue, share the same semantic representation.

(19d), and it must take wide scope with respect to the second NP (19e) (Baker forthcoming: 19):

(19) a. *She gave Mary$_i$ the towel undressed$_i$. (Williams 1980: 204)
b. *spy-telling (cf. secret telling) (Baker forthcoming: 17)
c. ??Who did she give the towel? (Erteschik-Shir 1979)
d. *She gave the towel the woman she just met. (Stowell 1981)
e. The teacher assigned one student every problem: unambiguous scope assignment (Larson 1990: 604)

The derivational account proposes that these restrictions (several of which we will see are only apparent) can be accounted for if one assumes the recipient argument originates within a PP even in the double object construction. In the case of (19a), the logic of the argument is clear; the recipient argument appears to pattern just like the prepositional argument of the paraphrase in disallowing predication by depictive predicates (*hungry, wrapped*):

(20) a. *John gave the meat to Mary$_i$ hungry$_i$. (prepositional paraphrase)
b. *She gave Mary$_i$ the meat hungry$_i$. (ditransitive)

(21) a. John gave the meat$_i$ to Mary wrapped$_i$. (prepositional paraphrase)
b. John gave Mary the meat$_i$ wrapped$_i$. (ditransitive)

Thus a restriction against depictive predicates applying to recipient arguments would seem to follow from a generalization that depictive predicates only apply to subjects and objects; they cannot be used to predicate properties of prepositional complements (Williams 1980). If the recipient argument is underlyingly a prepositional phrase, then the fact that it patterns just like the prepositional argument of the paraphrase would be expected. However, contra the derivational account, depictive predicates *may* apply to certain ditransitive recipients:

(22) a. They gave her$_i$ communion awfully young$_i$.
b. The guard gave him$_i$ a pat-down naked$_i$.

The prepositional paraphrases of these expressions differ in acceptability:

(23) a. They gave communion to her awfully young.
b. ??The guard gave a pat-down to him naked.

Moreover, Pylkkanen (2003) notes that in other languages depictive predicates readily apply to the recipient argument of a ditransitive. Thus the idea that recipients disallow depictive predication is not well supported empirically.

The second observation is that the recipient argument of both the ditransitive and the prepositional paraphrase resists "incorporation" into a synthetic compound:

(24) *spy-telling, *child-reading (Baker forthcoming: 17)

The logic of this argument is somewhat less clear. It is true that the generalization does not depend on whether the argument is construed to be a recipient of the ditransitive or a goal argument of the prepositional paraphrase—indeed, it is impossible to tell the difference. There is likely to be a semantic explanation for the resistance of prototypically animate arguments to appear in compounds (Nunberg, Sag, and Wasow 1994). In any case, the restriction against incorporation is not likely to be an effect of the recipient argument being part of a PP, as the derivational account implies, because other arguments regularly expressed as PPs do readily incorporate, including, for example, locations (*river-fishing; sky-diving*).[9] Another reason why the restriction is not likely to be syntactic is that certain counterexamples exist, including *deity-offering*.

Moreover, the logic of the argument for a derivational account is deeply flawed in the case exemplified by the generalizations in (19c,d,e): instead of the recipient argument patterning like the goal prepositional argument in the paraphrase, it patterns *differently* from it. In particular, the prepositional goal argument can readily be questioned (with preposition stranded or not: (25a,b)), whereas the recipient argument of the ditransitive cannot be (25c):

(25) a. To whom did you give the meat?
b. Who did you give the meat to?
c. ??Who did she give the towel?

Similarly, the prepositional dative can be involved in heavy NP shift, whereas the recipient argument of the ditransitive cannot be ((19d), repeated as (26)):

(26) *She gave the towel the woman she just met. (Stowell 1981)
(cf. John gave to Mary the meat he just bought at the store.)

As Baker acknowledges, it is not clear on his account what actually accounts then for the ill-formedness of (25c) and (26). Various stipulations are offered and the question is left open (p. 16).

In addressing the fact that the recipient argument of the ditransitive must take wide scope over the theme argument (cf. (19e) repeated as (27)), Baker again observes that the same is *not* true for the prepositional dative:

[9] Baker of course recognizes this fact (Baker 1988), but it nonetheless serves to undermine the explanation as to why recipients resist incorporation.

(27) The teacher assigned one student every problem. (not ambiguous: only wide scope reading of "one student")
(cf. The teacher assigned one problem to every student. (ambiguous))

It is acknowledged that, in fact "Unfortunately, there is no satisfactory analysis available for the scope-freezing effect in double object constructions" (Baker forthcoming: 18).[10] Still he concludes, "Nevertheless, we can be optimistic that, whatever the final analysis is, it will support the idea that only the dative double object construction has an underlying structure that does not match its surface form" (Baker forthcoming: 18).

To summarize, let us return to the generalizations observed by Baker (cf. examples (19a-e)). The first two generalizations do not hold: the recipient argument *can* be the subject of a depictive predicate, and the recipient argument can at least occasionally be the non-head of a compound. Other generalizations, namely the resistance of the recipient argument of the ditransitive to appearing in long-distance dependency relations and its preference for a wide-scope interpretation with respect to the theme argument are not properties that motivate a derivational account of the ditransitive, because they serve to distinguish the ditransitive recipient from the recipient/goal argument of the prepositional paraphrase.[11]

[10] See Bruening (2001) for a proposal to account for the scope-freezing effect that depends on the syntactic structure of ditransitives. The fact that, as we shall see in Ch. 8, inverse scope is sometimes possible in the ditransitive immediately undermines this account, since the structural proposal would predict that inverse scope is categorically ruled out. In fact Breuning had claimed to have "established that wide scope for the second object is never available in double object... constructions, regardless of context or quantifier involved" (2001: 239).

[11] A final argument Baker offers is that while the NP PP variant has an unaccusative counterpart (cf. (i), (ii)), the NP NP does not (cf. (iii),(iv)):

(i) The ring passed to Mary.
(ii) The beer opened for Max.
(iii) *Mary passed the ring.
(iv) *Max opened a beer. (on the reading that Max is a recipient, not an agent)

Once again, this would suggest that the ditransitive and the prepositional variant are distinctly different. Moreover it is clear that Baker's example of *pass* and *open* in (i–ii) are extremely atypical. In general, the agent is very much required in active sentences:

(v) 1. *The milk gave to Mary.
2. *Salad fixed for Mary.
3. *The ball threw to Mary.
4. *The story told to Mary.
5. *The letter sent to Mary.
6. *The house built for Mary.

It is true that there do exist certain verbs that can be expressed with <theme goal> arguments, or alternatively with <agent theme goal> arguments, such as *pass* and *open* (also *move*), but there also exist certain verbs that appear with <recipient theme> arguments, or alternatively with <agent recipient theme> arguments. These include *get, rent,* and a non-standard use of *learn*:

(vi) 1. She got the book.
2. He got her the book.

(vii) 1. She rented an apartment.
2. Her mother rented her an apartment.

(viii) 1. She learned how to cook.
2. She learned me how to cook. (non-standard English)

The facts are clearly variable depending on particular lexical items.

Thus Baker's description of the special properties of the ditransitive construction do not provide motivation for a derivational account; instead they beg for an explanation as to what makes the ditransitive construction *distinct* from its prepositional paraphrase. We return to answer this question in Chapter 7, where we will see that the information structure independently associated with the ditransitive construction predicts its behavior with respect to long-distance dependencies and scope assignment. It seems that the only thing that the respective paraphrases share with the ditransitives is the quite rough paraphrase relations themselves. There is no empirical motivation to derive ditransitives from prepositional paraphrases, nor is there motivation to treat ditransitives that admit of distinct paraphrases as more than minimal variants of each other. The robust generalizations are surface generalizations.

2.5 The Caused-Motion construction

Beyond target syntax and target semantics arguments are *input* syntax and semantics arguments: it is preferable to avoid deriving A from C if there exists a pattern D that has the same target syntax and semantics as C and yet cannot serve as input from which to derive A. By widening our focus beyond those expressions that may serve as paraphrases of ditransitives, we see that each paraphrase expression itself is a small part of a much broader generalization. For example, although only (28a) can be paraphrased by a ditransitive, it patterns together with (28b,c,d) both syntactically and semantically; in fact, all of the expressions in (28) can be captured by a single "caused-motion" construction (Goldberg 1995; cf. also Pinker 1989).

(28) a. Mina sent a book to Mel.
b. Mina sent a book to Chicago.
c. Mina sent a book toward the front of the room.
d. Mina sent a book through the metal detector.

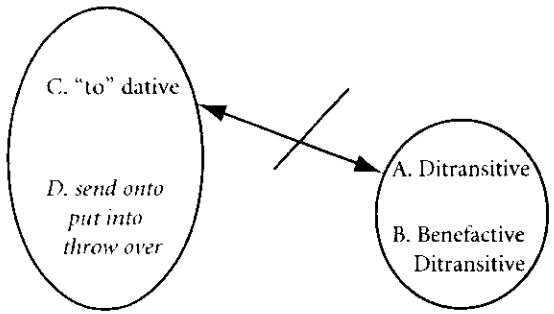
Although Baker (forthcoming), as we saw above, argues in favor of a derivational account of ditransitives, he recognizes that examples such as those in

(28) should be treated alike, noting that "it seems artificial to say that the PP in [examples like (28a)] is not a locational path" (Baker forthcoming: 31; cf. also Marantz 1997).

Similar extensions of meaning that we saw above for the ditransitive likewise exist in the case of the caused-motion construction, even though the verb classes involved are distinct:

(29) a. Mina coaxed Mel into the room. (if coaxing is successful, Mel moves into the room)
b. Mina helped Mel into the room. (Mina helps Mel move into the room)
c. Mina blocked Mel out of the room. (Mina causes Mel not to move into the room)

These facts motivate treating the caused-motion construction as a general construction, independent of the ditransitive.



2.6 Load/Spray

Similar arguments can be made for other types of argument structure patterns that are often only considered in terms of alternations (Anderson 1971; Fraser 1971; Hook 1983; Rappaport and Levin 1988). Consider the following examples in (30) and (31):

(30) Pat loaded the wagon with the hay.

(31) Pat loaded the hay onto the wagon.

It has been suggested that the *with* variant is derived from the *into* variant (e.g. Rappaport and Levin 1988). Let us consider the "input" syntax and semantics first. The "into" variant can be seen to be an instance of the much broader caused-motion construction already discussed. That is, each of the examples in (32) shares the same surface syntax: each has a DO and prepositional oblique phrase. The semantics are closely related as well; in each case the subject argument serves to cause the motion of the DO argument along the path or to the location specified by the oblique argument:

(32) a. Pat loaded the hay onto the wagon.
b. Pat put the hay on the wagon.
c. Pat shoveled the hay into the wagon.

The (b) and (c) forms of (32) cannot serve as input to any locative alternation, as can be seen in the ill-formedness of the following examples:[12]

(33) a. *Pat put the wagon with hay.
b. *Pat shoveled the wagon with the hay.

We thus see that the input syntax and semantics arguments hold for the *into* variant of the so-called locative alternation.

We now turn to the putative "output" syntax.[13] Consider just the following examples that have been independently classified as instances of the "locative" construction by Pinker (1989) and Levin (1993):

[12] Recognition of the fact that *load onto* type expressions are instances of the more general caused-motion construction serves to solve a certain paradox in the acquisition literature. It has often been observed that children are more likely to make overgeneralizations such as those in (i) than they are to overgeneralize the pattern with *with* as in (ii):

(i) She filled the water into the cup. (relatively common)
(ii) She poured the cup with water. (rare)

The explanation for this has been thought to be mysterious because it has been claimed that far fewer verbs appear in the *into* variant than the *with* variant (Gropen et al. 1991). The overall frequency of the *into* variant was thought to be less than the *with* variant as well. However, once we recognize that the *into* variant is actually part of a much larger generalization, the caused-motion construction, it becomes clear that the frequencies that matter are the frequencies associated with that broader generalization as compared with the causative-plus-instrumental-adjunct pattern. Sethuraman (2002: 146) has calculated just these statistics in the (Bates et al. 1988) corpus of speech between twenty-seven mothers and their twenty-eight-month-old children. The children produced a total of forty-two caused motion tokens compared with two transitive + *with* tokens. Mothers produced 199 caused-motion tokens compared with twenty-five transitive + *with* tokens. If we extrapolate from these patterns is seems that the caused-motion construction is 8–20 times more frequent than the causative + *with* adjunct construction. Figures for the *type* frequencies involved in the causative + *with* variant are not available, but the number could not possibly be higher than the token frequencies (since each unique type requires a new token), and is likely much lower. The type frequency of the caused-motion construction in children's speech is sixteen; in the mothers' speech it is forty. Again the *token* frequencies for the *with* construction are two and twenty-five, respectively. Since type frequency is correlated with productivity (Bybee 1985, 1995), the fact that children more readily overextend the caused-motion construction than the causative + *with* phrase is to be expected.

[13] Arguing that the *with* variant is an instance of a broader generalization is somewhat more controversial than the other cases discussed so far, primarily because *with* has a remarkably wide range of uses, but see Goldberg (2002) for discussion.

(34) a. Pat loaded the wagon with the hay.
b. Pat sprayed the wall with paint.
c. They covered the wall with posters.
d. Pat adorned the tree with lights.
e. They tiled their bathroom with blue tiles from Mexico.
f. They stained the wood with an all-weather protector.
g. He speckled the canvas with dots.
h. He wrapped the present with tin foil.

The examples in (34) are arguably all licensed by the combination of two constructions: a causative construction and an independent construction headed by *with*. In this way we account for the well-known fact that the DO in these examples is necessarily interpreted as affected in some way; e.g. the truck must be interpreted to be full or otherwise affected in *Pat loaded the truck (with hay)*. The same is not true for *Pat loaded hay onto the wagon* (Anderson 1971), which only entails that some hay is put on the wagon. That is, the affected status of the DO is rendered completely non-mysterious and requires no ad hoc stipulation (cf. also Rappaport and Levin 1988; Gropen et al. 1991).[14]

Adopting then the idea that the examples in (34 a–h) admit of a causal analysis, notice none of the examples (34 c–h) permit the alternation typically discussed as being relevant to *load* and *spray* as shown in (35c–h) (see also Pinker 1989):

(35) a. Pat loaded the hay onto the wagon.
b. Pat sprayed paint onto the wall.
c. *They covered posters onto the wall.
d. *Pat adorned lights onto the tree.

[14] One would have to be quite an ardent lumper to try to class all of these uses of *with* under a single sense. In (i) is a sampling of its various uses. Again, this is not the claim of the present chapter: we do not deny the existence of constructional homonymy. It is suggested, however, that it is important not to assume massive ambiguity without seeking out broader surface generalizations.

(i) 1. Elena traveled with Maya.
2. Elena traveled with a hat on.
3. Aliza traveled with great enthusiasm.
4. People associate one variant with another.
5. Be sure to mix the butter with sugar.
6. The foundation provided the school with funding.
7. Pat loaded the wagon with hay.
8. The garden swarmed with bees.
9. The detective entered the room with a key.
10. Pat broke the window with a hammer.
11. Pat watched the bear with a telescope.

e. *They tiled blue tiles from Mexico onto their bathroom
f. *They stained an all-weather protector onto the wood.
g. *He speckled dots onto the canvas.
h. *He wrapped tin foil onto the present.

Thus in accord with the target syntax argument, it is preferable to generate (35a,b) directly instead of deriving them from (34a,b) due to the fact that there exist (34c–h) that have parallel syntax and semantics and cannot be derived from (35c–h).

Other surface structure generalizations

Applying parallel reasoning, it can be demonstrated using the same input syntax/semantics arguments and target syntax/semantics arguments that the *for* paraphrase of certain ditransitives (e.g. (36a)) patterns together with (36 b,c) syntactically and semantically; each are instances of a transitive construction together with a benefactive adjunct construction. The shared syntax and semantics of these phrases argue for treating them alike (see Goldberg 2002).[15]

(36) a. Mina sent a book for Mel.
b. Mina sent a book for the library.
c. Mina sent a book for her mother's sake.[16]

[15] An objection might be raised against the proposal that all *for*-benefactive phrases should be treated as a natural class. It might be argued that because more than one can co-occur, they cannot play the same role in the sentence:

(i) Mina sent a book for Mel for her mother's sake.

That is, Fillmore (1968) long ago observed that only one semantic role of each type may occur in a single clause. We do not find two distinct agents or patients co-occurring in a single clause:

(ii) *Bob melted the butter by Paul.

(iii) *The butter was melted the ice.

But Fillmore's constraint only holds of certain semantic roles, namely those that can be identified as arguments. Adjuncts can freely be added as long as they do not imply a semantic contradiction; in particular they must be construed to have concentric semantic scope such that one more narrowly specifies another. Consider the following sentence with multiple temporal adjuncts:

(iv) Mina met Bob in the morning yesterday at 11 o'clock.

Notice the hour (11 o'clock) must occur within the part of the day (morning) which is in turn within the day (yesterday). It cannot be claimed that the temporal phrases must be interpreted syntactically as a single complex adjunct because they need not be continuous:

(v) a. Yesterday Mina met Bob in the morning at 11 o'clock.
b. At 11 o'clock in the morning Mina met Bob yesterday.
c. Yesterday Mina met Bob in the morning by the beach at 11 o'clock.

[16] It should be made clear that we are not claiming that *all for* phrases encode benefactives. Clearly there are other uses of the preposition *for* in English which may not be related, for example, those in

Other "alternations" have been analyzed similarly as independent surface generalizations, including the following:

(37) a. Bees are swarming in the garden.
b. The garden is swarming with bees. (Fried forthcoming; Jackendoff 1990; Salkoff 1983)

(38) a. Tom is similar to Bill.
b. Tom and Bill are similar. (e.g. Gleitman et al. 1996)

(39) a. She broke it.
b. It broke. (e.g. Van Valin 1990)

(40) a. They considered him to be a fool.
b. They considered him a fool. (e.g. Borkin 1974; Gonzálvez-García 2001)

Target and input syntax and semantics arguments argue against derivations. The formal patterns involved are more profitably viewed as constructions on their own terms. Each of these constructions can be seen to be much more general than is often recognized when only instances that alternate in certain ways are considered.

2.7 The role(s) of the verb

In this section, we address the question of how to account for the overlap in meaning in paraphrases and we examine why the overt interpretation of instances of the same construction may differ, and may allow distinct ranges of paraphrases. One key to these questions lies in the recognition that there is more to the interpretation of a clause than the argument structure construction used to express it. The overall interpretation is arrived at by integrating the argument structure construction with the main verb and various arguments, in light of the pragmatic context in which the clause is uttered.

(i) and (ii). Prepositions are typically highly polysemous and sometimes homonymous (see Brugman 1988; Lakoff 1987; Lindner 1981; Jackendoff 1990).

(i) The statue stood for three hours.
(ii) He exchanged the socks for a belt.

That is, there do exist instances of *constructional homonymy*: a single surface form having unrelated meanings. In order to identify which argument structure construction is involved in cases of constructional ambiguity, attention must be paid to individual verb classes. In fact, in order to arrive at a full interpretation of any clause, the meaning of the main verb and the individual arguments must be taken into account (see Chapter 6). What is being proposed is simply that if a constituent looks like a benefactive phrase and acts like a benefactive phrase, then there is no reason to be shy about calling it a benefactive phrase.

There is a growing recognition that it is important to recognize a distinction between the frame semantics associated with a verb and the set of phrasal patterns or argument structure constructions that are available for expressing clauses (Gleitman et al. 1996; Goldberg 1992, 1995, forthcoming; Rappaport Hovav and Levin 1998; Iwata 2000; Jackendoff 1997b, 2002; Kay 2001; Pinker 1994). Following Goldberg (1992, 1995) the slots in the argument structure constructions are referred to as "argument roles." That is, phrasal constructions that capture argument structure generalizations have argument roles associated with them; these often correspond roughly to traditional thematic roles such as *agent, patient, instrument, source, theme, location*, etc. At the same time, because they are defined in terms of the semantic requirements of particular constructions, argument roles in this framework are more specific and numerous than traditional thematic roles (see also Jackendoff 1990, 2002).

Argument roles capture surface generalizations over individual verbs' participant roles. That is, each distinct sense of a verb is conventionally associated with rich frame semantic meaning that in part specifies certain *participant roles*: the number and type of slots that are associated with a given sense of a verb. A subset of those roles, namely those roles which are lexically *profiled*, are obligatorily expressed, or, if unexpressed, must receive a definite interpretation.[17] Lexical profiling, following the general spirit of Langacker (1987a, 1991), is designed to indicate which participant roles associated with a verb's meaning are obligatorily accessed, functioning as focal points within the scene, achieving a special degree of prominence. Fillmore (1977) similarly notes that certain participant roles are obligatorily "brought into perspective" achieving a certain degree of "salience." The notion of lexical profiling is intended to be a semantic one: it is a stable aspect of a word's meaning, and can differentiate the meaning difference between lexical items—cf. *buy* versus *sell* (Fillmore 1977) or *rob* versus *steal* (Goldberg 1995). Participant roles may be highly specific and are often unique to a particular verb's meaning; they therefore naturally capture traditional selectional restrictions.

Two general principles can be understood to constrain the ways in which the participant roles of a verb and the argument roles of a construction can be put into correspondence or "fused": the **Semantic Coherence Principle** and

[17] This generalization is true for English. In many other languages profiled arguments are omissable as long as they are given and non-focal in the context. At the same time, lexically profiled roles are expressed by a small set of core grammatical relations, when they are expressed in these languages as well.

the **Correspondence Principle** (see Goldberg 1995 and Goldberg forthcoming for further discussion).

The Semantic Coherence Principle ensures that the participant role of the verb and the argument role of the construction must be semantically compatible. In particular, the more specific participant role of the verb must be construable as an instance of the more general argument role. General categorization processes are responsible for this categorization task and it is always operative. This principle follows from the idea that argument structure constructions are learned by generalizing over the semantics of instances of the pattern used with particular verbs (e.g. Tomasello 1992, 2000; Goldberg 1999; this volume, Chapter 4).

As is the case with lexical items, only certain argument roles are profiled. In the case of simple sentences, only roles expressed in formally prominent positions are considered prominent. Such positions receive a special status in most theories as the set of "terms" which correspond to "core," "nuclear," or "direct" arguments. In English, profiled argument roles are realized as Subj, Obj, or the second object in ditransitives. These positions are afforded a high degree of discourse prominence, being either topical or focal in the discourse (see Keenan 1976, 1984; Comrie 1984; Fillmore 1977; Langacker 1987a for arguments to this effect.). Specifically the Correspondence Principle states that profiled participant roles of the verb must be encoded by profiled argument roles of the construction, with the exception that if a verb has three profiled roles, one can be represented by an unprofiled argument role (and realized as an oblique argument). The Correspondence Principle is a default principle, which is at root iconic.

The intuition behind the Correspondence Principle is that lexical semantics and discourse pragmatics are in general aligned. That is, the participants that are highly relevant to a verb's meaning (the profiled participant roles) are likely to be the ones that are relevant or important to the discourse, since this particular verb was chosen from among other lexical alternatives. In particular, the Correspondence Principle requires that the semantically salient profiled participant roles are encoded by grammatical relations that provide them a sufficient degree of discourse prominence: i.e. by profiled argument roles. As a default principle, the Correspondence Principle can be overridden by particular constructions that specify that a particular argument be deemphasized and expressed by an oblique or not at all. Passive, for example, is a construction that overrides the Correspondence Principle and insures that a normally profiled role (e.g. agent) be optionally expressed in an oblique *by* phrase. See Goldberg (forthcoming) for discussion of other constructions that serve to override the Correspondence Principle.

2.8 Accounting for paraphrase relations

We are now in a position to address the question of how the overlap in meaning between alternants is accounted for. The shared meaning can be attributed directly to the shared verb involved. That is, the verb evokes the same frame semantic scene and the same profiled participant roles. For example, if we assign the participant roles of *load* the labels loader, loaded-theme, and container, we can see that these roles line up with the roles in the caused motion construction and causative + *with* constructions as follows:

(41) Caused motion (e.g. *Pat loaded the hay onto the truck*)

CAUSE-MOVE	(**cause**	**theme**		path/location)
|	|	|		|
Load	(**loader**	**loaded-theme**		**container**)

(42) Causative + with constructions (e.g. *Pat loaded the truck with hay*)

CAUSE	(**cause**	**patient**)	+	INTERMEDIARY		(instrument)
|	|	|				|
Load	(**loader**	**container**				**loaded-theme**)

All three of *load*'s roles are profiled. This includes the loaded-theme role even though that role is optional. This is because when that role is optional, it receives a definite interpretation as indicated by the strangeness of the following mini-conversation (see Fillmore 1986 for tests to distinguish definite from indefinite omission):

(43) She loaded the truck. #I wonder what she loaded onto it.

Because all three roles are profiled, one of the roles may be expressed as an oblique argument, in accordance with the Correspondence Principle. The Semantic Coherence Principle insures that only semantically compatible roles may be fused. As indicated above, the loaded-theme role of *load* may either be construed to be a type of theme as in (41) or an intermediary as in (42). The container role can either be construed to be a path/location as in (41) or a patient role as in (42). Construing the verb's roles as instances of different argument roles is what results in the different semantic construals of the two constructions.

On this view, there is no need to say that the *with* phrase itself designates a theme relation (cf. e.g. Jackendoff 1990). Instead, the fact that the hay is interpreted to be loaded onto the truck even in the *with* variant is attributed, not to the argument structure construction, but to the specifications of the verb *load*.

2.9 Arguments and Adjuncts

Recognizing that the verb has its own profiled participant roles that are distinct from the argument roles associated with an argument structure construction allows us to recognize four possibilities.

The most common, prototypical case is one in which the profiled participant roles of the verb line up isomorphically with the argument roles of an argument structure construction. This is represented in cell (a) in Table 2.2. Another familiar case is one in which a non-profiled role is expressed by an adjunct construction as represented in cell (d).

In other cases, there is a mismatch between the roles of the verb and the argument structure construction. Sometimes an argument role may not correspond to an independent obligatory participant role of the verb (c). For example, when the ditransitive construction is combined with verbs of creation, the recipient role is associated only with the construction; we do not need to assume that verbs of creation lexically specify a potential recipient. The same is true for certain verbs of motion as well. *Kick*, for example, only has two profiled participant roles; the recipient argument in *She kicked him the ball* is added by the construction.

The fourth logical possibility is that a profiled participant role of the verb is expressed by what is normally considered to be an adjunct phrase. As suggested in the (b) cell of the table, it seems appropriate to identify the *with* phrase that appears with *load* as an instance of this type. Clearly other

TABLE 2.2. Possible routes to argument status

	Role of argument structure construction	Not a role of argument structure construction
Profiled/obligatory participant role of verb	**(a) ARGUMENT of verb and construction** *He* devoured *the artichokes.* *She* gave *him* a letter. *She* put *the package on the table.*	**(b) ARGUMENT contributed by the verb** She loaded the wagon *with hay.*
Not a profiled/ obligatory participant of verb	**(c) ARGUMENT contributed by construction** He baked *her* a cake. She kicked *him* a ball. She sneezed *the foam off the cappuccino.*	**(d) Traditional ADJUNCT** He baked a cake *for her.* She broke the window *with a hammer.* She swam *in the summertime.*

instances of the same construction (including what are usually referred to as instruments) normally function as adjuncts (in being omissible, able to appear sentence-initially or after a clear adjunct such as *yesterday*, etc.). However, we have seen that the loaded-theme participant role of *load* is a profiled role. Because the *with* phrase codes a profiled role but is expressed by a phrase that is normally an adjunct, we might expect the behavior of this argument to fall somewhere in between that of traditional arguments and traditional adjuncts. In (44) we see that this is the case. While placing a clear adjunct before the *with* phrase is not crashingly bad in (44a); it is slightly less felicitous than the corresponding example in (44b).

(44) a. ?Pat loaded the wagon yesterday with hay.
b. Pat broke the window yesterday with a hammer.

To summarize, we need not be blind to potential differences between uses of a construction with particular verbs. We need to account for verb meaning anyway, so it makes sense to look to verb meaning to determine whether differences in interpretation or in the range of possible paraphrases can be straightforwardly accounted for by it.

2.10 Conclusion

Recognizing surface generalizations surrounding argument structure (i.e. argument structure constructions) is important in that it leads to the recognition of generalizations in language that might otherwise be overlooked. But it is equally important to bear in mind that the meaning of a clause is more than the meaning of the argument structure construction used to express it. Individual verbs as well as particular arguments and context must be factored into the equation. In accounting for similarities among alternative expressions and dissimilarities among instances of the same argument structure construction, careful attention must be given to the verb which is the same in the former and different in the latter.

The arguments in this chapter should not be taken to imply that possible paraphrase relations play *no* role in the learning, processing, or representation of language. The essentially structuralist observation that the semantic interpretation of one linguistic construct tends to be affected by the existence of possible alternatives, receives empirical support from a number of studies (e.g. Lambrecht 1994; Lambrecht and Polinsky 1997; Moore and Ackerman 1999; Spencer 2001; McCawley 1978).

In Chapter 5 it is argued that the statistical use of paraphrases in actual discourse contexts is critical to unlocking Baker's paradox of partial product-

ivity (Brooks and Tomasello 1999; Goldberg 1993, 1995; Pinker 1984; Regier 1996). Paraphrase relations can also be seen to be relevant to online choices made in production (see Chapter 6, Section 6.10).

However, it is less clear that one particular paraphrase should have a privileged status, nor that it is profitable to analyze one phrasal pattern solely by implicit or explicit reference to another. It has been argued here that by carefully examining a fuller range of surface phenomena, broader generalizations, surface generalizations, in the form of argument structure constructions, are revealed.

3

Item-specific knowledge and generalizations

There is always a tension between being a "lumper" and being a "splitter." As a biologist once put it, "splitters see very small, highly differentiated units—their critics say that if they can tell two animals apart, they place them in different genera... and if they cannot tell them apart, they place them in different species. Lumpers, on the other hand, see only large units—their critics say that if a carnivore is neither a dog nor a bear, they call it a cat" (Simpson 1945).

Language contains both large generalizations and idiosyncratic facts, and therefore we unavoidably find those who favor lumping, and those who favor splitting. The constructionist approach to grammar offers a way out of the lumper/splitter dilemma: the approach 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. In this chapter we consolidate evidence that such a usage-based model is required to account for the synchronic state of grammar. Before turning to the case of language, let us review some relevant findings in the non-linguistic category literature.

3.1 Exemplar-based knowledge in categorization generally

There is a good deal of evidence in the field of non-linguistic categorization that information about specific exemplars is stored. In a classic dot-recognition study, for example, Posner and Keele (1968) asked subjects to classify random dot patterns of low or high distortion from the average. One group had small distortions from the average; the other group had larger distortions from the same average. Subjects were subsequently tested on dot patterns with more distortion than either training set. The subjects in the high-distortion

condition performed better on the task, indicating that they stored more than the simple average (prototype) of the instances on which they were trained.

In related tasks, it has been demonstrated that an instance that is more similar to recently studied instances will be classified more accurately than another which is equally similar to a prototype (Whittlesea 1987). And a recently studied exemplar will tend to be classified more accurately even if less similar to a prototype (Medin and Schaffer 1978).

Other work indicates that people are able to use statistical properties of the features of stored exemplars, including the range of values for each feature and correlations among features (Rips 1989). For example, if you learn of two bird species, one with three times the size of beak of the other, which do you think is more likely to sing? Most people are able to predict that the bird with the smaller beak is more likely to sing, presumably because of correlations between small size and small beak, and small size and ability to sing among bird exemplars. Since these correlations were not likely consciously taught or recognized, it seems that the generalization is made on the fly on the basis of stored exemplars.

Exemplar-based models of categorization have been proposed in order to capture these sorts of empirical facts (Medin and Schaffer 1978). In these models, a category is represented by a collection of instance representations. Classification of new instances is based on their similarity to the stored exemplars. It is fair to say that until very recently such models dominated work on categorization in cognitive psychology.

Exemplars are somewhat abstract

It is important to realize that exemplar-based models of categorization do not do away with abstraction completely. Generally attributes that are more relevant to the task at hand are more likely to be noticed. Any aspect of an exemplar that is not recorded because the learner failed to (unconsciously) notice it, is obviously not stored. This represents a degree of abstraction over the actual input: if a given stimulus, S, has attributes a, b . . . z, but the person witnessing S only records attributes a, b, c, and d, the resulting representation will be more abstract than S, in that it will not specify attributes e–z. Because of this *selective encoding*, what is actually recorded is not a fully specified memory of an encounter, but rather a partial abstraction over what was encountered. In addition, human beings' knowledge erodes over time—the (unconscious) forgetting of attributes (and entire exemplars) also renders our representations more abstract than a collection of actual veridical reproductions of stimuli.

3.2 Generalizations over exemplars

At the same time that much item-specific knowledge is recognized in non-linguistic categorization, there is a growing recognition that exemplars alone do not account for our intricate knowledge of generalizations. That is, exemplar models fail to explain how exactly items cohere as a category. As Ross and Makin put it, "the exemplar view seems to take away the 'categoriness' of categories" (Ross and Makin 1999: 8). We have certain knowledge about the category *bird* that extends beyond the individual exemplars we have experienced. We know that birds as a class lay eggs, have feathers, and typically fly. We know that some types of dinosaurs may have evolved into birds (without necessarily knowing which particular types of dinosaurs or birds they may have evolved from or into).

Experimental evidence indicates that our knowledge of instances leads to generalizations. For example, Ross, Perkins, and Tenpenny (1990) devised the following category-learning experiment. They initially showed all subjects descriptions of two members of some Club Y:

Member A: likes ice cream
Buys nails

Member B: likes to read Westerns
Buys a swimsuit

Subjects in the C condition were then asked to decide whether a new person was a member of the same club, given the following description:

C: likes sherbet
Buys wood
Buys a towel

Other subjects, in the D condition, were asked to decide whether a different set of features described a member of the club or not.

D: likes to read Cowboy and Indian stories
Buys wood
Buys a towel

Afterwards, subjects in both groups were asked to rank the following features on scale of 1–7 as to their relevance for the category, Members of Club Y:

Plays tennis
Buys a chisel
Has children
Buys sunglasses

Ross, Perkins, and Tenpenny found that subjects who saw C rated "buys a chisel" as more relevant to the category of members than subjects who saw D. Conversely, subjects who saw D rated "buys sunglasses" as more relevant to members than those who saw C. The interpretation of this data is that subjects in the C condition were more reminded of exemplar A, because of the inclusion of the thematic relationship between the first features (liking ice cream and liking sherbet). This in turn led to more attention to the second feature as well, leading subjects to generalize to a superordinate category involving "carpentry" which led to the rating that buying chisels was relevant. In contrast, subjects in condition D were reminded of B, leading them to generalize a "beach" category that in turn led to their ranking of "buys sunglasses" as being more relevant to the category than it was to subjects in condition C.

In recognition of data such as these, there are a growing number of psychological models of categorization that combine exemplar-based knowledge with some type of generalizations. For example, on the *exemplar-based abstraction* view, categorizations are made using exemplars, but the effect is abstraction based on similarity that is additionally stored (Medin and Edelson 1988; Ross, Perkins, and Tenpenny 1990; Spalding and Ross 1994). Abstractions are created locally, on the basis of small numbers of exemplars (even just two). "These abstractions will often be far more specific than [an abstract schema] would be... However, if these abstractions are used later in classifying another instance, a still more general abstraction may be made, consisting of the commonalities between the first local abstraction and the new instance" (Ross and Makin 1999).

Another concrete model of categorization that involves both instances and abstractions is Anderson's Rational Model (Anderson 1991). In this model, exemplars are grouped together in clusters. Each cluster has a central tendency that represents that cluster (a mini-prototype). The model determines whether to add a new exemplar to an existing cluster or start a new cluster by comparing a new instance to all existing clusters. Assignment is also affected by the size of the clusters and the likelihood that instances group together. Groups may be of varying size, determined by a variable "coupling" parameter.

In an impressive cross-discipline convergence, certain linguists as well as these cognitive psychologists are embracing combination models that involve both instances and abstraction over those instances (Barlow and Kemmer 2000; Langacker 1987a; Taylor 1995). These theories all acknowledge that while we record a great deal about individual instances of categories, we also discern

meaningful relationships among members that make categories cohere together as categories. Let us now turn to the domain of language.

3.3 Item-specific knowledge in adult grammar

Language learning must involve memories of individual examples because the end state of grammar is only partially general (Bybee 1985; Bybee and McClelland 2005; Culicover 1999; Daugherty and Seidenberg 1995; Lakoff 1970; Plunkett and Marchman 1993).

Phonology/Morphology

At the level of phonology, it is clear that very specific aspects of usage events can be and are retained as part of our cognitive representations. For example, there is evidence that particulars of phonetic realizations are retained even if they are predictable by phonological generalizations. For example, what counts as a voiced consonant has been shown to be different in different languages, indicating that speakers retain more specific information than simply [+voiced] (Pierrehumbert 2000). Individual words that are used with higher frequencies tend to be more reduced than other words. For example, Losiewicz (1992) has shown more reduction in the final syllable of *needed* (frequent) than *kneaded* (less frequent). Bybee (2000) observes that *every*, a very high-frequency word, has come to be pronounced as a two-syllable word /evri/, where as low-frequency words like *mammary* or *summery* are pronounced with three syllables. Words of intermediate frequency, such as *camera*, *family*, and *memory*, allow for variation between pronunciations with two or three syllables. These facts suggest that we index patterns according to their frequencies.

Gahl and Garnsey (2004) demonstrate that phonological reductions (/t,d/ deletions) are more likely in high-probability constructional contexts than in low-probability contexts. In particular, when experimental subjects are asked to produce a sentence in which there is a match between the overt construction used and a verb's statistical bias (as determined by corpora and sentence completion norming studies), they tend to abbreviate the verb form more than when the construction to be produced does not match the bias of the verb involved.

Booij (2002c) has argued that predictable allomorphemic variation is lexically represented, in the case of highly frequent words, since the variation may be retained even when the regular generalization is lost (see also Baayen, Burani, and Schrender 1997 for evidence that highly frequent, regular morphology is lexically represented).

Grammatical category

Idiosyncratic facts about more high-level generalizations such as grammatical categories are retained as well. Consider our knowledge of adjectives. Prototypical adjectives, such as *red, hot, dry,* and *big,* modify referents and can appear prenominally or predicatively (e.g. after copula verbs such as *seem),* as in (1)–(2). However, neither of these formal properties is necessary. The adjective *mere* can only appear pronominally (3)–(4), while the adjective *aghast* can only appear predicatively (5)–(6):

(1) a red book

(2) The book seems red.

(3) a mere child

(4) *That child seems mere.

(5) *an aghast man

(6) The man seemed aghast.

Even the semantic property of modification does not hold of all adjectives. *Occasional* does not modify or quantify the cigarette in (7), but the act of smoking:

(7) She smoked an occasional cigarette.

There is also a little construction in English that allows adjectives to appear postnominally, [(all) *things* AP], as in *She loves all things linguistic.* Clearly we must learn the distributional properties of these words and constructions individually. Their distribution does not follow from general facts about adjectives.

English Serial Verb Constructions

To illustrate the partial productivity and idiosyncrasy evident in argument structure patterns within the adult grammar, we briefly consider here a few related, quite understudied constructions. These constructions serve further to illustrate the type of partially idiosyncratic and partially general knowledge that language learners retain.

The first construction, the VVingPP construction, was brought to my attention by Ray Jackendoff (personal communication, 2002). This construction involves a motion verb followed by a verb in progressive form and a directional complement, as in (8):

(8) The toddler went screaming down the street.

The VVingPP construction clearly has its own peculiar constraints. The directional is an argument of the main verb, not of the second verb. In fact, unlike the paraphrase of motion predicates involving a subordinate clause (illustrated in (9)), the progressive verb in the VVingPP construction may not appear with its own arguments (10):

Subordinate manner clause

(9) a. Bill went down the street whistling a tune.
b. Bill took off toward the cops screaming at the thief (all the while).

VVingPP

(10) a. *Bill went whistling a tune down the street.
b. *Bill took off screaming at the thief toward the cops.

The main verb in the VVingPP construction is not very productive. Acceptable examples involve intransitive motion verbs with a very general meaning, namely *come, go, run,* and *take off.* With these verbs, the Ving slot is quite open:

(11) Bill went singing/grinning/waving/laughing down the street.

Other intransitive motion verbs are unacceptable:

(12) a. *Bill raced whistling down the street.
b. *Bill walked whistling down the street.

Transitive verbs *take* and *bring* are also acceptable to varying degrees, depending on the choice of Ving:

(13) Bill took him kicking into the room.

(14) Bill brought him kicking and screaming into the room.

(15) ?? Bill took him whistling into the room.

(16) ?? Bill brought him grinning into the room.

The progressive form of the complement bears its normal semantics such that the activity described must be construed as obtaining over a period of time or as being iterative:

(17) a. Bill jumped off the bridge. ≠
b. Bill went jumping off the bridge. (VVingPP)

While (17a) is interpreted as a one time, telic action, (17b) is necessarily interpreted as iterative.

The syntax of the active construction appears to be [Subj [VVingPP]]. Notice that [VingPP] is not a constituent in that it cannot appear as a fronted unit:

(18) Down the hill Bill went screaming.

(19) ??Screaming down the hill Bill went.

This parse also captures the idea that the PP is an argument of the main verb, since it is a sister to the verb. In addition, the second verb has an adverbial meaning and its distribution mimics to some extent that of adverbs, insofar as it appears without arguments and is sister to the verb as well.[1]

Thus the VVingPP construction appears to be a serial verb construction of English, despite the fact that English does not allow serial verbs in general. The VVingPP construction represents a special form with its own special semantic and syntactic constraints: a conventionalized construction that must be learned from the input the learner receives. The construction is represented below, in Table 3.1.

TABLE 3.1. The VVingPP ("Took-off-Screaming") Construction

Sem:	Move	in a Manner	along a Path
	|	|	|
Syntax:	$V_{\in \{go,\ come,\ run,\ take\ off\}}$	Ving	(Oblique)

In fact a close look at the data reveals a family of related constructions in English. The Took-off-Screaming construction is superficially similar to another conventionalized construction, which we can label the *Go*VPing construction, or more informally, the "Don't go sticking your nose" construction. There are several important reasons to distinguish the two constructions. Unlike the main verbs in the VVingPP construction, *go* in the *Go*VPing construction is not interpreted as a motion verb and therefore does not license a directional. Moreover, the V of the VP expresses its arguments, as indicated by the use of "VP" instead of a simple V. For example, the direct object in (20) is a complement of *read* not *go*:

(20) You shouldn't go reading the newspaper all day.

Further differentiating the *Go*VPing construction from the VVingPP construction formally is the fact that the *Go*VPing construction only allows the main verb *go* (cf. (21) and (22)), and it prefers that *go* appear in bare form (23):

[1] This may be a candidate for the sort of "syntactic mimicry" described by Francis (2005). However, adverbs and not V-ing forms can appear preverbally:

(i) The barrel quickly rolled down the hill.
(ii) * Bill screaming went down the hill.

(21) Pat'll go telling Chris what to do, you'll see.

(22) *Pat'll come telling Chris what to do.

(23) ??Pat went telling Chris what to do.

The semantic properties of the two constructions also differ. Unlike the VVingPP construction, the *Go*VPing construction may refer to an instantaneous action:

(24) Don't go spilling your drink!

(25) Don't go jumping off the bridge now!

There is additionally a semantic constraint only associated with the *Go*VPing construction: it implies that there is something negative about performing the action designated by the complement. For example, (21) requires a context in which the speaker disapproves of Pat's telling Chris what to do.

TABLE 3.2. The *Go*VPing, "don't go sticking your nose" Construction

Pragmatics: The action designated by VP is construed negatively by the speaker		
Sem:		Action type
		|
Syntax:	*go*	$[\text{Ving}\ldots.]_{VP}$

Finally, while the VVingPP construction is part of Standard English, the *Go*VPing construction is restricted to informal speech.

There is yet another distinct construction in which both verbs are in bare form: the $Go\text{VP}_{bare}$ construction. In this construction, *go*, *come*, and *run* are all acceptable (but *take off* is not). Unlike the *Go*VPing construction, the verbs in this construction retain their usual motion interpretations:

(26) Go tell your sister to come here.

(27) Won't you come sit with me?

(28) Would you run get me a pencil?

The motion is interpreted as facilitating the action designated by the VP, where the main assertive event is encoded by the VP. Tensed forms of the first verb (*go*) are unacceptable:

(29) *She came sat/sit with me.

(30) *He goes bring/brings the paper.

The negative implication associated with the *Go*VPing ("don't go sticking your nose") construction is absent from the $GoVP_{bare}$ construction:

(31) She had better go tell her what to do.

(32) Go do your homework!

This "go tell it to the mountain" construction can be represented as in Table 3.3.

TABLE 3.3. The $GoVP_{bare}$ "go tell it to the mountain" Construction

Sem: MOVE in order to do action	
Syn: $V_{\langle go,\ come,\ run \rangle}$	VP_{bare}

Thus there are clearly three separate constructions in English. Each must be described on its own terms, with its particular syntactic, semantic, and pragmatic constraints. None is completely general. The idiosyncratic properties of each construction provide an argument that its specifics must be learned on the basis of generalizing over particular examples.

Conventionality and Redundancy

> My legal name is Alexander Perchov. But all of my many friends dub me Alex, because that is a more flaccid-to-utter version of my legal name. Mother dubs me Alexi-stop-spleening-me!, because I am always spleening her... because I am always elsewhere with friends, and disseminating so much currency, and performing so many things that can spleen a mother. Father used to dub me Shapka, for the fur hat I would don even in the summer month. He ceased dubbing me that because I ordered him to cease dubbing me that. It sounded boyish to me, and I have always thought of myself as very potent and generative.
>
> Jonathan Safran Foer, *Everything is Illuminated* (Boston: Houghton Mifflin Co., 2002).

The narrarator of the passage above is clearly intended to be a non-native speaker. How can we tell? It is because much of the phrasing used and combination of lexical choices are non-conventional, even if fully grammatical.

It is in fact 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. Conventions of telling time and reporting height, for example, differ from language to language, forming regular subpatterns that must be learned from the input. Pawley and Syder (1983) point out that this "native-like selection" must ultimately be accounted for as part of a native speaker's knowledge of language (see also Lamb 2002). To do so, however, clearly requires that a certain amount of redundant information must be represented, since the conventional patterns would in any case be generated by the grammar.

Further evidence for some amount of redundancy in language comes from the fact that very typically, a fully general linguistic pattern is instantiated by a few instances that are highly conventional. In such a case, it is clear that both generalizations and instances are stored. For example, conventional instances of the constructions just discussed are given below:

Idiom	don't go sticking your nose in where it doesn't belong
General pattern	*go*VPing
Idiom	go kicking and screaming <path>
General pattern	VVingPP
Idiom	go tell it to the mountain
General pattern	*go*VP_{bare}

A few more general examples of this phenomenon are given below, along with a reference to work on the more general pattern:

Idiom:	Boys will be boys.	
General pattern:	$NP_{animatepl}$ will be $NP_{animatepl}$	(Wierzbicka 1988)
Idioms:	The bigger they come, the harder they fall. The more the merrier.	
General pattern:	The Xer the Yer	(Fillmore, Kay, and O'Connor 1988; Culicover and Jackendoff 1999)
Idioms:	What's this fly doing in my soup? What's a nice girl like you doing in a place like this?	
General pattern:	What's X doing Y?	(Kay and Fillmore 1999)
Idiom:	X worked x's way through school.	
General pattern:	subj V poss way PP	(Goldberg 1995; Israel 1996; Jackendoff 1990)
Idiom	Give me a break	
General pattern	Subj V Obj1 Obj2	(Goldberg 1992; Green 1974; Oehrle 1975)

Item specific facts about argument structure

In this section, we review evidence that speakers have exemplar-based knowledge of verb-specific patterns such as that given in (33):

(33) <actor> *put* < thing > < location >

There exists abundant evidence that children are very conservative in their early argument structure productions. That is, they stick closely to the forms they have heard used with particular verbs (Akhtar and Tomasello 1997; Baker 1979; Bates and MacWhinney 1987; Bowerman 1982; Braine 1976; Brooks and Tomasello 1999; Gropen et al. 1989; Ingram and Thompson 1996; Lieven, Pine, and Baldwin 1997; MacWhinney 1982; Olguin and Tomasello 1993; Pinker 1986; Schlesinger 1982; Tomasello 1992). For example, Olguin and Tomasello (1993) taught twenty-five-month-old children four novel transitive verbs, each in a different syntactic pattern: both participants expressed, agent only, patient only, or neither argument expressed. Children almost always reproduced the same exact pattern they had heard. Of course, in order to restrict their usage to formulations that they have heard or used in the past, they need to record what they have heard.

In a remarkably comprehensive diary study, Tomasello (1992) observed that by far the best predictor of his child's use of a given verb on a particular day was her use of the same verb on the previous few days, not, as might be expected, her use of other verbs on the same day. Tomasello and his colleagues have discussed this verb-centered conservatism under the rubric of *verb islands*, since children readily substitute new nouns into the frames (Akhtar and Tomasello 1997; Clark 1996; Gropen, Epstein, and Schumacher 1997; Tomasello 1992; Tomasello et al. 1997). A simple example of this type of conservatism comes from the diary data of a child, Aliza, collected by the author. Aliza routinely omitted prepositions before the age of approximately twenty-one months. During that time, she produced *come me* and *play me* to mean "come with me" and "play with me," respectively. Beginning at 1;9.9, and subsequently, Aliza reliably produced *come with me*. Nonetheless, she still continued to produce *play me*, without the preposition, for another two months.

There is evidence that adults retain much verb-specific knowledge as well. Verbs are occasionally quite idiosyncratic in the types of argument structure patterns they appear in (Bresnan 1982; Chomsky 1965; Pollard and Sag 1987). For example, the near synonyms *help* and *aid* differ in their distribution:

(34) a. Pat helped her grandmother walk up the stairs.
b. *Pat aided her grandmother walk up the stairs.

(35) a. ??Pat helped her grandmother in walking up the stairs.
b. Pat aided her grandmother in walking up the stairs.

Even though this sort of example may be rare, a learner cannot possibly know which patterns will turn out to be productive and which will not be on initial encounter. Thus it is clear that all early-learned instances must be stored. Unless we posit some sort of house-cleaning device to erase this early scaffolding, these early-learned forms would continue to be stored.[2]

Psycholinguistic studies have demonstrated that speakers are influenced by the relative frequencies with which they have heard particular verbs used in various argument structure constructions (Ford, Breslan, and Kaplan 1982; Jurafsky forthcoming; MacDonald, Pearlmutter, and Seidenberg 1993). For example, knowledge that *believed* is more likely to appear with a clausal complement than with an object complement influences speakers' online interpretation of potentially ambiguous sentences (Garnsey et al. 1997; Trueswell, Tanenhaus, and Kello 1993). The relative frequencies play a role despite the fact that both possibilities are fully grammatical, as in the examples (36a–b):

(36) a. Pat believed the speaker might cause a riot.
b. Pat believed the speaker.

Newmeyer (2003) rightly cautions that frequency information can sometimes be overinterpreted. For example, he argues that the choice of syntactic constructions depends on their differing meanings—that the probabilities of use are a function of their meanings, not some inherent statistical property of the structure. This idea is supported by the fact that statistics are known to vary, sometimes rather dramatically, across different corpora (Roland and Jurafsky forthcoming).

Hare, McRae, and Elman (2004, 2003) explore the idea that semantics drives distribution in some detail. They confirm that subcategorization possibilities are conditioned by a verb's and a construction's senses. For example, the verb *find* must occur with a direct object if it is used to mean "locate," whereas it is biased toward appearing with a sentential complement when it means "realize." However, statistical factors still play a role. *Find* can appear with a direct object, even when it is used to mean "realize;" online reading times indicate that the statistical preference for the sentential complement plays a role, even when the verb's sense is controlled for (see also Argaman 2002; Roland and Jurafsky forthcoming). Findings of this sort

[2] Even proponents of "dual route models" of morphology who argue that regular forms are created on the fly by rules while irregular forms are stored, have observed that at least some high-frequency regulars are stored redundantly with their stems (Pinker 1999; Pinker and Jackendoff 2005).

demonstrate that detailed verb-specific knowledge about frequencies of usage influence adult grammar.

We now turn to evidence for the existence of higher-level generalizations, beyond generalizations over particular arguments of a given verb. In Chapter 4 it is argued that argument structure generalizations are *based on* verb-specific patterns.

3.4 Argument Structure Generalizations

To perhaps most linguists, it goes without saying that languages contain generalizations. But if we take item-based knowledge seriously, it raises the question as to whether only individual tokens are stored without any generalization. This possibility has in fact been raised by certain researchers, insofar as they seem to make the claim that the totality of what is stored are specific usage events (Boas 2000; Thompson and Fox 2004; Verhagen 2002). This view deserves pause, in part because, as noted above, a similar suggestion had great currency, and still has its adherents within cognitive psychology in theories known as *exemplar models* of categorization (e.g. Medin and Schaffer 1978).

Still, there is ample evidence that generalizations are essential to language. If generalizations were not necessarily made, we would expect to find languages whose argument structure patterns varied arbitrarily on a verb-by-verb basis. For example, we might expect to find one semantically transitive verb expressed by SVO (Subject Verb Object) word order, another expressed by SOV order, and a third verb expressed by VSO order:

(37) a. Pat saw Chris.
b. Pat Chris kissed.
c. Hate Pat Chris.

But in fact languages are much more regular. Semantically similar verbs show a strong tendency to appear in the same argument structure constructions. *Help* and *aid* cited above are unusual; more typically, verbs that are closely related semantically do appear in the same argument structure constructions (Fisher, Gleitman, and Gleitman 1991; Goldberg 1995; Gross 1975; Levin 1993; Pinker 1989). Newly formed creoles quickly generalize patterns beyond individual verbs (Sandler et al. 2005).

Further evidence that children generalize the patterns they use stems from the fact that they occasionally produce spontaneous overgeneralizations, as in the following examples, from Bowerman (1982) and the author's own diary data:

(38) a.	Will you have me a lesson?	Bowerman (1982) Christy 4;0
b.	She came it over there.	Bowerman (1982) Christy 3;4
c.	I'll hockey over there. (to mean she'll move over there in her roller skates, carrying a hockey stick)	Aliza 6;7
d.	She unlocked it open.	Zach 3;0
e.	Circle it back up!	Zach 4;8

It is also clear that adults continue spontaneously to generalize argument structures patterns (Aronoff 1976; Clark and Clark 1979; Pinker 1989). The attested examples in (39) provide examples of such adult novel productions:

(39) a. Once you resort to higher-level predicates, you can just lambda your way out of practically anything. (reported by John Moore, May 1995)
b. He concentrated his hand steady. (reported by Georgia Green; found in Russell Atwood's *East of A*, New York: Ballantine Books, 1999).
c. They haven't found the time to play him a whole lot of minutes. (Pinker 1989: 154)
d. Mary presented as an attractive, neatly dressed woman. (Pinker 1989: 155)
e. I'll just croak my way through, I guess. (reported by Mike Tomasello, May 1996)
f. Diane hasn't Botoxed and siliconed herself into some kind of weird creature. (reported by Hana Filip, March 2004)

The successful manipulation and comprehension of nonsense verbs in experimental settings also demonstrates that speakers are in fact able to make generalizations (Akhtar and Tomasello 1997; Gropen et al. 1989; Maratsos et al. 1987; Naigles 1990).

KNOWLEDGE OF LANGUAGE IS KNOWLEDGE. Speakers classify the instances they hear into categories. Verb-centered categories are categorized together, ultimately resulting in general, abstract argument structure constructions.

3.5 When do generalizations emerge?

It is sometimes suggested that children are almost totally unaware of argument structure generalizations until the age of three or three-and-a-half (Tomasello 2000, 2003). However, it might be expected that generalizations

emerge gradually from early on. In fact, in diary records kept of my children, there exist a number of early overgeneralizations:[3]

(40) up and down the neigh (requesting that I raise and lower a toy horse) (Aliza 1;8.21)

(41) come Abbi (requesting that I make toy dog, Abbi, come to Aliza) (Aliza 1;8.2)

(42) you jump me to the sky (Aliza 2;1.23) (asking me to help her jump on the bed)

(43) you mad to the pig? (Aliza 2;1.23)

(44) I reach my hands up (Aliza 2;1.26)

Aliza's first overgeneralizations appeared when she was only twenty months. At that point she used approximately twenty-five verbs, including the transitives *eat, get, give, have, hear, help, hit, hold, make, put, read, ride, take, tickle,* and *wipe*. Thus she had had an opportunity to make some tentative generalizations over the verbs that had already been learned. Her overgeneralizations grew greatly in number and frequency over the next year and a half.

Other early instances of overgeneralizations come from Zach. This data was collected less systematically and there may have been even earlier generalizations. No comprehensive record of his vocabulary was kept (readers with more than one child may understand this lapse):

(45) It noises. (Zach 2;4) (In answer to "what's that?" about a top when its electronic noise wouldn't turn off)

(46) Could you reach me up here, Kitty Cat? (Zach 2;4) (playing that one toy cat was talking to another, while raising them to the bookshelf)

(47) Hold me what I want. (Zach 2;5) (to mean, give me what I want by holding me so that I can reach it: creative extension of the ditransitive construction)

(48) I becamed to be Spiderman. (Zach 2;7) (cf. I wanted to be Spiderman)

It is conceivable that both children happen to fall on the very early end of the curve of when generalizations emerge as reported by Tomasello (2000). However, there also exists experimental work that suggests that in fact,

[3] Note that omission of prepositions cannot account for these novel uses because (40) involves the novel use of particles as verbs, and (41) cannot be paraphrased with any existing preposition: the only candidate, *with* as in *come with Abbi*, could not have been intended because I was bending over Aliza, putting her in her car seat when she asked me to go bring her the toy dog.

argument structure generalizations over verb-centered instances emerge gradually from very early on.

Strong evidence that young children can use more general, abstract argument structures (as well as verb-specific knowledge) is provided by Akhtar (1999). In this study, thirty-six two-, three-, and four-year-olds (mean ages: 2;8, 3;6, and 4;4) were taught meanings for novel verbs that were modeled for the children using non-English word orders as well as canonical English order, for example:

(49) a. Elmo the car gopping. (SOV)
b. Dacking Elmo the car. (VSO)
c. Elmo blicking the car. (SVO: canonical English order)

The children's spontaneous use of these verbs and responses to queries of "what happened?" were recorded. The two- and three-year-olds matched SOV or VSO patterns roughly half the time and changed the order to SVO roughly half the time.[4] The fact that the children produced the non-English orders at all is striking evidence that they are able to learn patterns on a verb-by-verb basis; at the same time, the fact that the children ever produced the unmodeled SVO order that corresponds to the regular English pattern indicates that the children recognized the regular English pattern as a generalization over the particular instances they had heard. The generalization played a greater role in the productions of the four-year-old participants, in that children at this age were overwhelmingly more likely to correct to SVO than to match the modeled order.

Akhtar also ran a control condition in which familiar verbs were modeled in the same non-English orders:

(50) a. Elmo the car pushing.
b. Hitting Elmo the car.

In this case, the two- and three-year-olds as well as four-year-olds were significantly more likely to correct to SVO than to use the order that was modeled. This is expected since the argument structure for these particular verbs had already been learned. Thus the children in the experimental condition were not simply mimicking the experimenter's usage blindly, since they did not do so when the verbs involved were already known and their patterns of usage were already familiar.

[4] The children were more likely to switch to English order if their utterances were produced with pronouns, e.g. *He dacked it.* This may be because pronouns are so frequent and appear in relatively fixed positions in the input (see Pine, Lieven, and Rowland 1998 for discussion).

Abbott-Smith, Lieven, and Tomasello (2001) replicated Akhtar's study with children that were slightly younger. They found that the youngest children (2;4) corrected non-canonical orders only half as often as the children did at 2;8 in Akhtar's study. Still, even children aged 2;4 corrected the non-canonical orders roughly 25 per cent of the time. The children also used the novel verb that was heard in a grammatical order more often than they used a novel verb that had been heard in an ungrammatical order. This is yet another indication that the generalizations are learned gradually, beginning from very early on (cf. also Ninio 1999, 2005).

By the time children are four years old, it is clear that they readily form generalizations over lists of attested instances (represented at some level of abstraction). Thus, for example, English speakers are aware of a productive transitive construction and readily extend it with appropriate verbs, in addition to being aware of which particular verbs they have heard used in the construction before.

3.6 Representations of constructions capture predictive aspects of the constructions

As was the case with non-linguistic categorization, selective encoding and imperfect memory ensure that our exemplars are somewhat abstract. We do not store an unlimited number of complete utterance representations; rather what we retain are instances at some level of abstraction. That is, we do not passively retain a huge mental corpus, consisting of all the strings we have ever heard, as a computer might do. Instead we constantly parcel out meaning, form abstractions, and generalize over the instances we hear.

For example, in learning constructions whose primary function involves information structure and discourse pragmatics, not semantics and not phonology—constructions such as relative clauses, questions, clefts—learners' representations are likely to abstract quickly over the particular semantics involved. For example, learners' representation of the English relative clause construction may contain almost no information about the *semantic* content that happens to have appeared in actual usage events. Upon hearing, *the man who sat on every hill*, we do not retain the phrase, *sat on every hill* as part of the representation of the relative clause, let alone the phonetic character of the way the word *every* was pronounced (although we do retain this information as part of our representation of the word *every;* recall the discussion in 3.3).

Learners *must* be attempting to assign functions to different parts of the utterances—without this tendency to analyze and decompose strings, we would be unable to assign any meaning to particular words unless the words happened to be used in isolation. This process is not well understood, but it must involve the notion of predictive value. The pronunciation of particular words helps users identify the word (comprehension) and produce the word as others do (production); therefore specific information about pronunciation is stored with particular words. But in other cases, such specific information is undoubtably abstracted away. There is no reason to think that the phonology of determiners is stored with particular verbs that happen to have been heard used with those determiners (the fact that *the* appears in *The man left* is not a fact about the verb *leave*). The co-occurrence of these two features is recognized to be contingent on the particular circumstance. This idea remains to be fleshed out, but it may ultimately help to explain the fact that constraints generally tend to be local (applying to immediate daughters and only less often to granddaughters or more distantly related constituents).

The tendency to seek out predictive correlations is clearly not specific to language. Accounts of our ability to parcel out responsibility in complex events has been discussed under the rubric of Bayesian causal networks (Pearl 1988, 2000). We constantly seek out causal connections within events: the door opens because the handle is turned, not because I happened to be wearing jeans. The cat screams because her tail was pulled, not because it happens to be raining outside. We are expert at identifying predictive correlations (see Chapter 6 for more discussion of this idea).

3.7 Usage-based Models of language

We have seen that in the case of language, as in categorization generally, there is solid evidence that both item-specific knowledge and generalizations coexist. A number of researchers have emphasized the need for both types of knowledge. Langacker (1987a), for example, warns against the "rule *vs.* list fallacy," criticizing the prevailing tendency to view productive generalizations and elaborated lists as being in competition (e.g. Marcus et al. 1995; Pinker 1999). He defines a *usage-based* approach that allows both instances and generalizations to be captured as follows:

> Substantial importance is given to the actual use of the linguistic system and a speaker's knowledge of this use; the grammar is held responsible for a speaker's knowledge of the full range of linguistic conventions, regardless of whether these conventions can be subsumed under more general statements. [The usage based model is] a non-reductive approach to linguistic structure that employs fully articulated schematic networks and emphasizes the importance of low-level schemas. (Langacker 1987a: 494)

It is clear that knowledge about language must be learned and stored as such whenever it is not predictable from other facts. Thus evidence that a word or pattern is not strictly predictable provides sufficient evidence that the form must be listed as a construction in what is sometimes called a "constructicon," in allusion to an expanded lexicon (e.g. Jurafsky 1996). At the same time, unpredictability is not a necessary condition for positing a stored construction. There is evidence from psycholinguistic processing that patterns are also stored if they are sufficiently frequent, even when they are fully regular instances of other constructions and thus predictable (Bybee 1995; Bybee and Hopper 2001; Losiewicz 1992; Pinker and Jackendoff 2005). We must recognize that patterns are stored as constructions even when they are fully compositional under these circumstances. Thus the present approach advocates a usage-based model of grammar (Barlow and Kemmer 2000; Bybee and McClelland 2005; Bybee 1995; Goldberg 1999; Langacker 1988). Grammars are usage-based if they record facts about the actual use of linguistic expressions such as frequencies and individual patterns that are fully compositional alongside more traditional linguistic generalizations. Most construction grammars are usage-based, due to the sort of evidence reviewed in this chapter.[5]

Usage-based proposals have been formulated in slightly differing ways by different researchers, but the essential point that both instances and generalizations over instances are stored remains the same. Culicover (1999) suggests two general properties of language learners: they are *conservative* in that they do not generalize significantly beyond the evidence in the input, and they are *attentive* in that they seek out generalizations that are consistent with the evidence presented. Insofar as learners have to record information about the input in order to be conservative with respect to it, Culicover's proposal is essential a version of the usage-based model. Israel (2002) echoes a similar theme. He suggests that language learners seek out both *local* consistency and *global* consistency. Local consistency makes learners aim to be conservative and stick closely with the local instances that they have witnessed. Global consistency makes learners seek out generalizations among instances so that the overall system coheres.

[5] Unification Construction Grammar (see Chapter 10), on the other hand, is not uniformly usage-based. That is, according to UCG, constructions are only posited if there is something not strictly predictable about either their form or their function. Fully compositional expressions are not stored even if they are highly frequent unless some aspect of their high frequency such as a register difference is non-predictable. For example, *walked* is not redundantly stored because it is fully predictable from the *walk* lexeme composed with the productive "past tense morphemic" construction.

Some may argue that linguists are not obliged to address the facts outlined in this chapter that argue for a usage-based model of grammar (Newmeyer 2003). However, it must be borne in mind that many of the facts that have been discussed have long been taken to be within the purview of all mainstream linguistic theories. For example, facts about an individual verb's complement taking possibilities or "subcategorization" frames have been assumed to be an essential part of linguistic theorizing since Chomsky (1965). More generally, if our aim is ultimately to characterize grammar in such a way that it is consistent with what we know about the use of language, then, other things being equal, a grammar that accounts for all the facts outlined in this chapter is preferable to one that does not. Moreover, viewing language in terms of a usage-based model allows us the chance to bridge naturally to an empirically grounded theory of how language can be learned. This is the subject of Chapters 4 and 5.