5 Partial Productivity

If you invent a verb, say green, which refers to an intended act of communication by speech and describes the physical characteristics of the act (say a loud, hoarse quality), then you know . . . it will be possible to green, to green for someone to get you a glass of water, to green to your sister about the price of doughnuts, to green "Ech!" at your enemies, to have your green frighten the baby, to green to me that my examples are absurd, and to give a green when you see the explanation.

Arnold Zwicky (1971)

5.1 Introduction

It has been a long-standing puzzle that many constructions are used somewhat productively (as implied by the above quotation), yet resist full productivity. This chapter addresses the issue of partial productivity for the most part by examining the ditransitive construction as an example. In section 5.5.2, other constructions, which can be seen to be either more productive or less productive, are considered.

The ditransitive construction can be used somewhat productively; that is, the construction can be extended to new and hypothetical verb forms (e.g., Wasow 1981). For example, the new lexical item fax can be used ditransitively as in (1):

(1) Joe faxed Bob the report.

Also, hypothetical lexical items are readily adapted to the ditransitive syntax. As Marantz (1984) notes, if we define a new verb, shin, to mean "to kick with the shin," it is quite natural for us to allow this new verb to be used ditransitively, as in (2):

(2) Joe shinned his teammate the ball. (p. 177)


At the same time, the ditransitive pattern is not completely productive within any generally defined class of verbs. Seemingly closely related words show distinct differences as to whether they allow ditransitive syntax. The following contrasts exist in many dialects:

(3) a. Joe gave the earthquake relief fund $5.
    b. *Joe donated the earthquake relief fund $5.

(4) a. Joe told Mary a story.
    b. *Joe whispered Mary a story.

(5) a. Joe baked Mary a cake.
    b. *Joe iced Mary a cake.

Brown and Hanlon (1970) have argued that children are neither corrected nor miscomprehended more often when they speak ungrammatically, so that they have no recourse to "negative evidence" that could allow them to either unlearn or avoid learning the above type of ungrammatical sentences (cf. Braine 1971; Baker 1979).

The standard solution to the no-negative-evidence problem in the case of vocabulary learning is to assume that there is indirect negative evidence in the form of attested input, assuming a principle that synonymy is avoided (cf. discussion in chapter 3). That is, a child may overgeneralize the past tense construction to produce comod as the past tense of come, but upon learning that come is synonymous, the child will expunge comod from her vocabulary, since she will assume that the language does not have two terms comod and come which are synonymous. Such indirect evidence is not forthcoming in an obvious way in the case of alternative syntactic patterns. It is not likely that the child simply expunges (6) upon hearing (7) because many verbs (e.g., give) do occur in both forms.

(6) *He whispered the woman the news.
(7) He whispered the news to the woman.

Moreover, as noted above, experimental evidence shows that children do not learn how to use syntactic patterns entirely conservatively, that is to say, solely on the basis of the input. If properly primed, they are willing to extend their use of verbs to previously unheard but related patterns.

An apparent paradox arises then, since if speakers have a productive mechanism that allows them to extend the use of the ditransitive syntax to new and novel verbs, it is not clear what prevents them from overgeneralizing to produce the above ill-formed examples (3b–5b).

This paradox is often sidestepped in linguistic theories. Thus, whether relation-changing lexical rules are intended to be purely redundant generaliza-
tions over stored items in a fixed lexicon, or rather generative rules which produce new forms productively, is often not made entirely clear.

Jackendoff (1975), for example, states that his lexical rules are intended only to account for existing regularities (both morphological and semantic) within the lexicon. These rules are represented by two-way arrows which encode the symmetric relation "is lexically related to." This aspect of Jackendoff's account is crucial, since he argues explicitly against Lakoff's (1965) proposal that productive rules generate "hypothesised lexical entries." However, Jackendoff also suggests that "after a redundancy rule is learned, it can be used generatively, producing a class of partially specified possible lexical entries" (p. 668).

Bresnan (1982) also attempts to find a middle ground between nonproductive and fully productive rules. While the lexical rules of LFG are explicitly conceived of as "redundancy rules," the metaphor of a lexically changing process is pervasive. The following is Bresnan's early description of the passive lexical rule:

\[
\text{Passivization in English} \quad \text{Functional change:} \quad (\text{SUBJ}) \rightarrow \emptyset / (\text{BY OBJ}) \\
(\text{OBJ}) \rightarrow (\text{SUBJ}) \\
\text{Morphological change:} \quad V \rightarrow V_{pass}
\]

The use of single-headed arrows and the word "change" indicate that the rule is a generative relation-changing rule. In fact, the notion of a "redundancy rule" itself is slightly oxymoronic, since a redundant statement of regularity is not in any normal sense rule-like.

In the remainder of this chapter, a resolution of the paradox of partial productivity is suggested, involving two types of learning mechanisms. The first is a certain type of indirect negative evidence; the second mechanism, presumably working in tandem with the first, draws largely on work by Pinker (1988) and Levin (1993) and the related experimental evidence of Groen et al. (1989).

5.2 INDIRECT NEGATIVE EVIDENCE

I do not attempt to survey the full range of efforts to suggest that some type of indirect negative evidence is available here (see Bowernan 1988 and Pinker 1989 for detailed discussion of the problem and critiques of many possible solutions), but there is one possibility raised in Pinker 1981, 1984, and then rejected in Pinker 1989 that deserves further study.

Since we have assumed that no two constructions are entirely synonymous both semantically and pragmatically (cf. chapter 3), it should be possible to find contexts in which a given construction is the most preferred. If the pre-ferred form is not used, then the child is able to tentatively infer that that form is disallowed. The inference would have to be tentative, since it is unrealistic to expect speakers to systematically use the most felicitous form in all contexts. However, if the situation repeats itself several times, the child's tentative hypothesis may become a fairly strong conviction. In this way, children would have the opportunity to unlearn certain overgeneralizations.

A simple case that may illustrate this is lexical and periphrastic causatives. It is well known that lexical causatives are used for cases of direct causation, whereas periphrastic causatives may be used for indirect causation (cf. 7.4.2). Therefore, for example, after seeing a magician make a bird disappear, the child may expect to hear a lexical causative as in (8), given that the causation is direct:

8. *The magician disappeared the bird.

Instead, however, the child may hear a periphrastic causative:

9. Look! The magician made the bird disappear.

The child may now tentatively hypothesize that the lexical causative is unavailable. That is, since the causation is direct, the lexical causative would be preferable if it were an option.

As another example, consider a child's strategy in determining whether a given verb can occur in the ditransitive construction. As noted by Ereschik-Shir (1979), and discussed in section 3.4.2 above, the ditransitive and its propositional paraphrase with no differ in the information structure of the clause. In particular, the ditransitive construction requires that the recipient argument be nonfocused (or "non-dominant" in Ereschik-Shir's terminology) and the transferred entity be focused ("dominant"). Propositional paraphrases prefer the opposite information structure: the recipient tends to be focused, the transferred entity nonfocused. Both of these generalizations are motivated by the fact that focused information tends to come at the end of the nuclear clause.

If the recipient is unfocused and the transferred entity is focused, we find the ditransitive more acceptable than the propositional paraphrase:

10. a. Sally gave him a brand-new red Volkswagen.
    b. Sally gave a brand-new red Volkswagen to him.

If the recipient argument is focused and the transferred entity nonfocused, we find the reverse situation:

11. a. Sally gave that to a charming young man.
    b. Sally gave a charming young man that.
When using verbs which can occur in both constructions, speakers are free to exploit the difference in pragramatic structure. There is, in fact, evidence that children are sensitive to these pragramatic factors (Gropen et al. 1989).

Indirect evidence would arise, then, from situations in which the discourse context matches a certain form but the speaker nevertheless uses a less feliciitous form. For example, speakers use the prepositional form for *whisper* even when the information to be conveyed more closely matches the information structure of the ditransitive construction. Thus, if a child hears (12) instead of (13), when the latter might be expected given the fact that the news is the focused information, the child will infer that the ditransitive form is not a possibility for *whispers.*

(12) Sally whispered some terrible news to him.
(13) *Sally whispered him some terrible news.

Pinker raises this possibility in several places (Pinker 1981, 1984:400). However, he ultimately rejects the idea that this mechanism could be sufficient for learning to disallow particular forms, for two reasons (Pinker 1989:16): The first objection he raises is that children's sensitivity to discourse contexts is statistical, not absolute. That is, children do not treat discourse effects as a determinant factor in choosing alternate argument structures; they are more likely to use the argument structure with the better-suited pragmatics, but they do not always do so (Gropen et al. 1989). However, the very fact that children are more likely to use the construction with better-suited pragmatics is sufficient to show that they do have an implicit knowledge of the information structure and are able to attend to it. For instance, it is possible that a child wouldn't notice the first time that *whisper* was used with the focus on the transferred entity, or the second time. But eventually, the child would presumably notice: at that time she would be able to use the input evidence to form the hypothesis that *whisper* cannot be used in the ditransitive form.

Even if we strengthen Pinker's first objection and assume that adult's sensitivity to discourse contexts is also only statistical and not absolute, we do not undermine this strategy. That is, even if we concede that neither adults nor children can be assumed to always use the most felicitious form, this strategy is not ruled out. All that is required is that the child be capable of recognizing a statistical correlation in the input data (Kapur 1993).

The second objection Pinker raises stems from his assumption that other focusing devices such as pronouns, cleft constructions, and contrastive stress can be used to override the default differences in information structure between alternative argument structures. On that view, adult speakers would be able to compensate for using less-preferred argument structures by overlaying these less-preferred argument structures with various focusing devices (e.g., pronouns, focus constructions, and stress), thus altering the information structure encoded by particular argument structures as a default. Assuming that speakers make use of these strategies, the input would be for the most part optimal, and children would have no reason to infer that the speaker in a given situation would have used a different argument structure if he could have.

However, this suggestion is ultimately not persuasive, since focusing devices are not able to alter or override the information structure of the clause but instead are required to obey the independently existing information structure of the clause. For example, pronouns are preferred in nonfocus positions:

(14) a. She gave it to a woman. >
   b. *She gave a woman it.

(15) a. She gave her a brand-new house. >
   b. #She gave a brand-new house to her.

Similarly, focus has been argued to only pick out arguments that are in focusable positions as defined for a given construction. For example, as mentioned above, Ertelchik-Shir (1979) has argued that the recipient argument of the ditransitive construction is not available as focus because the construction requires that argument to be nonfocused (or "non-dominant"): (16)

(16) a.??Who did you give the book?
   b.??It was Mary you gave the book.
   c.??Was it Mary you gave the book?

Finally, stress is also more felicitous on arguments that are in focus position, thereby generally emphasizing the information structure rather than overriding it. Thus example (17a) is more felicitous than (17b):

(17) a. She gave that to A WOMAN SHE JUST MET. >
   b. #She gave A WOMAN SHE JUST MET that.

Therefore, focusing devices might be well limited to giving the child additional evidence for the information structure of the clause, rather than serving to dilute other evidence by providing ways for the adult speaker to circumvent the information structure associated with a particular argument structure. Since two constructions generally differ either semantically or pragmatically, the hypothesis that indirect negative evidence is inferred from hearing a verb in a less-than-optimal construction deserves further study.

5.3 Circumscribing Verb Classes

Pinker (1989), arguing against any negative evidence (direct or indirect), ultimately provides a different, compelling resolution of the paradox of partial
productivity. A broad-range rule is proposed to capture the necessary conditions for a verb’s occurrence in additional syntactic frames. In the ease of the ditransitive, Pinker posits a broad-range rule that states in effect that a "prospective possessor" must be involved—that is, the first object referent must be understood to be a prospective possessor. This general rule does not provide sufficient conditions, however, there being many verbs that can be understood to involve a prospective possessor which do not allow ditransitive syntax (e.g., donate, contribute, pull, shout, choose, credit, say).

Drawing on work by Green (1974) (and Levin 1985 and Rappaport & Levin (1985) for the locative alternation), Pinker suggests that sufficient conditions are determined by a set of narrow-range rules which classify verbs into narrowly defined semantic classes. The specific classes that Pinker proposes are the following (cf. also Groen et al. 1989; Levin 1993):

1. Verbs that inherently signify acts of giving: e.g., give, pass, hand, sell, trade, lend, serve, feed

2. Verbs of instantaneous causation of ballistic motion: e.g., throw, toss, flip, slap, poke, fling, shoot, blast

3. Verbs of sending: e.g., send, mail, ship

4. Verbs of continuous causation of accompanied motion in a deliberately specified direction: bring, take

5. Verbs of event having (involving a commitment that a person will have something at some later point): e.g., offer, promise, bequeath, leave, refer, forward, allocate, guarantee, allot, assign, advance, award, reserve, grant

6. Verbs of communicated message: e.g., tell, show, ask, teach, pose, write, spin, read, quote, cite

7. Verbs of instrument of communication: e.g., radio, e-mail, telegraph, wire, telephone, teletype, fax

8. Verbs of creation: e.g., bake, make, build, cook, sew, knit, toss (when a salad results), fix (when a meal results), pour (when a drink results)

9. Verbs of obtaining: e.g., get, buy, find, steal, order, win, earn, grab

It may seem that if we admit the possibility of indirect negative evidence as suggested above, there is no need to adopt Pinker's suggestions that narrowly defined semantic classes also play a role in the acquisition of argument structure. However, circumscribing narrowly defined classes of verbs to be associated with a particular construction will allow us to account for extremely low-frequency or novel non-alternating verbs (since the assumed notion of indirect negative evidence presupposes hearing the verb in a non-optimal construction on several occasions). For example, taking Zwicky's example of a novel verb green, defined as a manner-of-speaking verb referring to speech of a loud, hoarse quality, speakers presumably know that they cannot say (18):

(18) *He greened her the news.

This knowledge cannot be attributed to any kind of indirect negative evidence, because the verb is novel; speakers would not have had a chance to unlearn or avoid learning it in this use. Other situations in which the type of indirect negative evidence suggested above would not be an aid to acquisition might include cases in which the construction in question is so low-frequency that the child can never with any modicum of confidence expect its occurrence, and cases in which there is no construction which is closely enough related semantically to the target construction so that the child would be able to infer that the speaker would have used the target construction if possible. Moreover, the generalizations Pinker describes are real; it is necessary to account for the fact that verbs which are used in particular argument structures do often fall into similarity clusters (cf. Levin 1993). There is no reason not to believe that children exploit multiple sources of evidence for learning argument structure; it is suggested here that they make use of narrowly defined verb classes as well as appealing to some degree of indirect negative evidence as described above.

Before continuing with Pinker's argument, we might make several small comments on the particular set of subclasses he proposes, listed above. The fifth subclass, "verbs of future having," can be seen to conflate three distinct subclasses. Some of the verbs are used in expressions which imply that the subject argument actually acts to cause the first object argument to receive the second object argument at some later point in time (e.g., bequeath, leave, forward, allocate, assign). In other cases, only if the conditions of satisfaction (Searle 1983) associated with the act denoted by the predicate hold does the subject argument cause the first object argument to receive the second object argument at some later point in time (e.g., promise, guarantee, owe). Finally, some verbs are used in expressions which imply that the subject argument only enables the first object argument to receive the second object argument (e.g., permit, allow) (cf. discussion in section 3.3.2).

The sixth class, "verbs of communicated message," should be understood to include verbs whose inherent semantics involves a communicative act, in order to distinguish this class from similar verbs such as say, assert, claim, and doubt which might be described as verbs of propositional attitude. Understood in this way, several of the verbs listed by Pinker seem to be misclassified; for example, pose and spin do not obviously fall into the class of "verbs of communicated message," and accordingly (at least in my dialect) are not readily dativizable.
b. *Bill spun her a fairy tale.

Both this class and the seventh class, "verbs of instrument of communication," should be classified as metaphorical classes since they are based on a systematic metaphor that involves understanding communicated information as being linguistically packaged and exchanged between interlocutors (Reddy 1979).

Finally, at least one additional subclass should be added to the list, namely, verbs of refusal such as refuse, deny. Expressions involving these verbs, like (20a, b), imply that the subject argument refuses to cause the first object argument to receive the second object argument.

(20) a. Bill refused Joe a raise.
b. The committee denied him a promotion.

In any case, we need only accept the spirit of Pinker's analysis—that there is a need to identify narrowly defined semantic subclasses—in order to accept his conclusion that this type of narrow circumscription allows us to capture the fact that subclasses of verbs which refer to the same kinds of general events as the ones listed, but do not fall into any of the above particular classes, fail to dativize. His examples of such nondativizing classes are as follows:

1. Verbs of fulfilling (X gives something to Y that Y deserves, needs, or is worthy of): *I presented him the award; *I credited him the discovery.

2. Verbs of continuous causation of accompanied motion in some manner: *I pulled/carried/pushed/schlepped/lifted/lowered/hauled John the bag.


5. Verbs of choosing: *I chose/picked/selected/favored/indicated her a dress.

Groen et al. (1989) provide experimental evidence to show that speakers are sensitive to certain morphophonological constraints. In particular, verbs with particular morphemes such as per-, con-, mid, same and polysyllabic verbs with non-initial stress are disallowed from participation in the ditransitive construction. These constraints largely coincide with distinctions between Latinate and native vocabulary, and between specialized and more basic vocabulary; however, we clearly would not want to ascribe recourse to etymological

information to children, and the experiments in support of these particular constraints controlled for semantic information. Therefore, the constraints are stated in terms of morphophonology. They are used to explain the following:

(21) Chris bought/purchased/obtained/collected him some food.
(22) Jan told/explained/announced Chris a story.

However, the constraints do not apply to every narrowly defined class of verbs. Verbs of future having, in particular, are not subject to them:

(23) Chris assigned/allotted/guaranteed/bequeathed him the tickets.

The class of instrument-of-communication verbs and the class of creation verbs also include verbs which are exceptions to the morphophonological constraints:

(24) Chris e-mailed/radioed/arpanetted him a message.
(25) Chris xeroxed/thermoluxed/urnoffed him a copy.

Groen et al. suggest that each of the verbs in (24–25) is classified, independently of the morphological criteria, as a special kind of complex stem having a noun or name as its root. They cite evidence that tacit knowledge of a word's stem being from another syntactic category allows it to be treated specially with respect to morphological processes (cf. Pinker & Prince 1988). To account for these cases, we can state the generalization that a verb from any class which is understood to have a noun or name as its root is not constrained by the morphophonological constraints.

The narrowly defined subclasses of verbs together with the morphophonological constraints provide a high degree of predictive power. A new or nonce verb which falls into one of the recognized narrow classes of verbs and which, if applicable, obeys the morphophonological constraints is automatically licensed to be used ditransitively (but see the next section). Verbs in conflict with these requirements are ruled out. This circumscription of narrow domains in which the ditransitive is productive goes a long way toward accounting for the apparent paradox that Pinker set out to resolve: that the ditransitive syntax can be extended to new and novel verbs, but at the same time is not available to all verbs of any broadly defined class.

5.4 Exceptions

The above generalizations are compelling, and in fact every researcher who has studied the semantics of the ditransitive construction in any detail has found it necessary to classify verbs which occur in the construction as belonging to narrowly defined subclasses as a descriptive device (cf. Green 1974;
Oehrle (1976; Wierzbicka 1986). Still, there are various kinds of exceptions to the generalizations just described. First, there are a couple of members in some subclasses which, do occur, yet the subclasses are not fully productive. Second, there is at times a certain degree of variability in judgments for verbs which are supposedly within the same narrowly defined class. Finally, there are exceptional verbs such as envy and forgive which do occur in the ditransitive construction although they do not entail the relevant semantics.

Each of these cases is discussed in turn. In section 5.3, an interpretation of the nature of the verb classes is suggested which can naturally account for all of these seemingly problematic phenomena.

5.4.1 Unproductive Subclasses

The small classes of verbs of permission (permit, allow) and verbs of refusal (refuse, deny) are unique in not forming productive subclasses:

(26) Sally permitted/allowed/*let/*enabled Bob a kiss.
(27) Sally refused/denied/*prevented/*disallowed/*forbade him a kiss.

These classes actually have a slightly different status in the theory proposed by Pinker (1989), because the verbs in these classes do not alternate with prepositional paraphrases. Thus on Pinker's account, these semantically related verbs are not eligible to undergo the lexical rule. However, since we are not postulating a lexical rule, we cannot appeal to the same solution. We need another way to account for their lack of productivity.

5.4.2 Differences in Judgment within Classes

An expected source of idiosyncrasy stems from the fact that the determiner of the narrowly defined class which a given verb belongs in is not always entirely clear-cut. For example, I have suggested that bequeath falls into the dativizing class of verbs of future having, along with leave, forward, allocate, etc. However, it seems that on semantic grounds it might be equally plausible to instead classify bequeath in the nondativizing class of verbs of fulfilling (X gives something to Y that Y deserves, needs, or it worthy of), along with present, credit, entrust, donate, etc. Because of these two classification possibilities, we would expect bequeath in fact to dative in some dialects and not to dative in others. In general, in the case of verbs that on the basis of their meaning may fall into one of two classes, one which can appear ditransitively and one which cannot, we would expect to find some dialectal variation as to whether these verbs can be used ditransitively.

Another source of lexical idiosyncracy is evidenced by the fact that speak-
The subjects in these cases are not causal, and no reception is involved. However, these predicates have illuminating semantic histories. *Forgive* and *envy* historically had senses that were closely related to *give*. *Forgive* used to mean "to give or grant" (OED). *Envy* used to mean "to give grudgingly" or "to refuse to give a thing to" (OED). This of course is not evidence that *forgive* or *envy* are part of the synchronic semantic pattern outlined above. But the historical facts do suggest that these predicates were at least at one time associated with this sort of pattern. Correspondingly, these facts also suggest that a construction can occasionally be frozen without continuing reference to the original semantics.

However, it seems reasonable that syntactic change should tend toward patterns that are more transparent to the speaker. If the construction with the semantics outlined here is psychologically real, then it would be natural for odd cases of ditransitives involving *forgive* and *envy* to drop out of use. I myself find archaic-sounding sentences involving *forgive* and *envy* much more acceptable than modern-sounding sentences; for example:

(36)  
(a) She forgave him his sins.
(b) *She forgave him his goof.*

(37)  
(a) She envied him his vast fortune.
(b) *She envied him his extensive stock portfolio.*

And in fact, other speakers are even less accepting of these constructions. In attempting to explain the idea of positive exceptions to a class of undergraduate cognitive science students, I wrote sentence (36a) and (37a) on the board. In response an audible groan arose from the class. When asked what was wrong, the students said they didn't find those sentences acceptable (this judgment was held by more than half of them). Thus it seems indeed that *envy* and *forgive* are dropping out of the language (at least among undergraduate students) just as we would expect if the semantics associated with the ditransitive pattern were synchronically real.

Nonetheless, *envy* and *forgive* have been exceptions for some time, and have been learned as such by generations of speakers. Thus an adequate account of grammar must allow for some degree of lexical idiosyncrasy, despite the powerful effect of semantic motivations (cf. also Lakoff 1965; Fillmore 1977b; Rosen 1984; Mithun 1991; Dowty 1991). Note that these cases are unproblematic from the point of view of learning, since the child has positive evidence that the verbs in question are used in the ditransitive construction, and can therefore learn them on an instance-by-instance basis as idioms.

### 5.5 Accounting for the Exceptions: A Usage-Based Account

#### 5.5.1 Productivity Defined by Verb Clusters

In the preceding sections, we have seen that even after embracing the idea of narrowly defined verb classes to account for the partial productivity of the ditransitive construction, there remains a residue of lexical idiosyncrasy. There are small subclasses which are not productive, varying degrees of acceptability within seemingly productive subclasses, and positive exceptions to the semantic generalizations, such as *envy* and *forgive*.

This idiosyncrasy is in fact expected if one considers certain experimental findings. Groepen et al. (1989) suggest that speakers tend to be conservative in their use of lexical items. Specifically, they show that people tend to use lexical items in the same constructions in which they have heard those items used by others, but that they can, if properly primed, extend the uses to new patterns. This phenomenon would be impossible if people did not store in memory the specific syntactic patterns that a word is heard used with (see also Bybee 1985 and Langacker 1987a for particular usage-based models of grammar). This being the case, a certain degree of lexical idiosyncrasy is to be expected.

However, the existence of some degree of lexical idiosyncrasy should not be taken as counterevidence against the existence of narrowly defined semantic subclasses of verbs that occur in the ditransitive construction. Although the exact formulation of these classes has differed, their existence as such has been recognized by every researcher who has looked in any detail at the verbs occurring in this construction. And, as has just been discussed (and is spelled out in more detail in Pinker 1989), the existence of such classes helps to explain the phenomena of partial productivity.

These two facts—that there are narrowly defined productive verb classes, and that at the same time we find scattered positive exceptions and varying degrees of acceptability within these narrowly defined classes—can be reconciled by recognizing verb classes to be implicitly represented speaker-internally as generalizations over learned instances. Because memory is associative, similar verbs used in the same constructions are classified together by general categorization processes. Therefore the claim is that speakers attempt to categorize learned instances.

Narrowly defined verb classes, then, are implicitly represented as clusters of semantically related verbs known to occur with a given construction. New or previously unclassified verb forms are attracted to existing clusters on the basis of similarity to existing cases. However, judgments of similarity are notoriously variable across speakers and contexts, and two activities can almost always be said to be similar in some respect. Therefore, in order to adequately
defend the idea that the use of new and novel senses is determined by similarity to existing cases, one must be able to define the similarity metric which is to be used as the basis of comparison. On the present account, the characterizations of the verb classes themselves can be viewed as providing a similarity metric. For example, if one of the verb classes associated with the ditransitive is "verbs of ballistic motion," then we can consider shin to be relevantly like kick in that it is a verb of ballistic motion.

The determination of which verb classes are relevant, or alternatively, what features of similarity are important, requires empirical crosslinguistic study, and I do not claim to provide an account here (but cf. Pinker 1989 and Levin 1993 for discussion and suggestions, and Mufwene 1978 for an early discussion on this subject). Only by looking at which distinctions are made crosslinguistically can we determine what the semantically (or morphophonologically) relevant aspects of verb meaning are that determine the basis of the clustering into subclasses.

On this view, frequency is expected to affect the classification of new verbs. Two kinds of frequency information need to be distinguished. On the one hand there is token frequency, which refers to the number of times a given instance (e.g., a particular word) is used in a particular construction; on the other hand there is type frequency, which refers to the number of distinct words that occur in a particular construction. MacWhinney (1978) and Bybee (1985) have argued that it is the type frequency of a particular process (or a particular construction) that plays a crucial role in determining how likely it is that the process may be extended to new forms: the higher the type frequency, the higher the productivity.

To see the relevance of the type/token frequency distinction for productivity, consider the following example cited by Bybee (1985:132–133). She notes that Guillaume (1927) documented the fact that French-speaking children most frequently overgeneralize the use of first-conjugation suffixes with verbs of other conjugations. He also observed the number of verbs of each conjugation used spontaneously in children’s speech. Bybee cites the following table, which shows the number of occurrences of each conjugation class and the number of verbs used from each class:

<table>
<thead>
<tr>
<th>Conjugation Class</th>
<th>Number of Uses</th>
<th>Number of Verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td>First (changer)</td>
<td>1,060 (36.2%)</td>
<td>124 (76.0%)</td>
</tr>
<tr>
<td>Second (fair)</td>
<td>173 (6%)</td>
<td>10 (6.1%)</td>
</tr>
<tr>
<td>Third (vendre)</td>
<td>1,706 (57.8%)</td>
<td>29 (17.9%)</td>
</tr>
</tbody>
</table>

Although more than half of the number of tokens ("uses") of verbs were of the third conjugation, the number of different verbs that occurred in this class was much smaller than the number that occurred in the first conjugation. Correspondingly, the first conjugation was seen to be much more productively used.

The proposal to implicitly represent verb classes as similarity clusters can perhaps be made more clear by the following rough-and-ready representation (which does not take morphophonological similarity into account):

![Figure 5.1](image)

Each circle represents a lexical entry; the entries are projected onto two dimensions, with semantically closer verbs being represented by physically closer circles. By way of demonstration, one or more instances within a given cluster have been labeled. The circles representing bake and cook, for example, are close together to indicate their being in the same narrowly defined class.

Type frequency can be discerned by considering the number of circles in a given cluster. Clusters containing more circles are more likely to be productive. Subclasses with only two members such as the verbs of refusal (deny, refuse) are expected not to be productive, because of their low type frequency.

The idea that verbs are represented this way in an associative memory is of course inspired by recent connectionist representations. However, the diagram need not be construed as necessarily presupposing a connectionist model of memory; all that is needed is an associative memory (e.g., as proposed in work in the domain of morphology of Pinker and Prince (1991)).

This view of the way new verbs are attracted to learned instances makes
several predictions. First, it predicts that subclasses with few members will not contain enough instances to create a similarity class, and so will not be productive. Secondly, it predicts the possibility of differences in judgments within similarity classes. Such differences will result from (1) the degree of similarity between the case being judged and other cases within the subclass, and (2) the relative type frequency that the relevant cluster displays. It is not necessary (or possible) to exhaustively list all the verbs that can potentially occur in a given construction. Novel cases are analogized to previously learned cases on the basis of their similarity to these familiar cases and the type frequency of these familiar cases.

Occasional positive exceptions (such as envy and forgive for the ditransitive construction) are tolerated because speakers simply associate the words with the constructions idiomatically. There is no danger of productive extensions from these outliers because they, like subclasses of fewer than two members, do not constitute a cluster, and therefore do not attract novel cases.

The representation in figure 5.1 entails that the knowledge that certain verbs are used in a particular construction is part of a speaker’s competence. However, it is not necessary that each new entry be stored as an additional member of a cluster, throughout the speaker’s life. It is possible that once a critical number of instances in a particular cluster is learned—insuring that novel instances that fall into the class will be included—new cases are no longer stored in memory since they would provide only entirely redundant information. It is also possible that learned instances are not necessarily stored as discrete, clearly individuated cases; rather, the edges of learned instances that form a cluster may blend into each other, delimiting an area in semantic space without specifically retaining each individual instance within.

5.5.2 Varying Degrees of Productivity

The ditransitive provides a good example of a construction with associated verb classes. The degree of productivity of other constructions can be seen to form a cline between those constructions which are not fully productive even within narrowly defined verb classes and those which approach full productivity as long as general constraints are obeyed. An example of the first case, that of very limited productivity (at least in some dialects), is the resumptive construction. As discussed in chapter 8, there is a large degree of idiosyncrasy as to which verbs can occur with which resumptives. Notice the following contrasts:

(38) a. She shot him dead. >
   b. ??She blasted him dead.

(39) a. She cried herself to sleep. >
   b. She cried herself asleep. >
   c. ??She wept herself to sleep/asleep.

(40) a. ??He ate himself as sleep.
   b. ??He cried himself sick.

At the same time, it is clear that resumptives are not entirely idiomatic and do occasionally occur productively. For example:

(41) a. “I cried myself well-nigh blind.” (Tennyson, “Grandmother X” (1884); cited by Visser (1963))
   b. “Drive your engine clean” (Mobil ad; cited by Rappaport Hovav & Levin 1991)
   c. “She could wonder herself crazy over the human eyebrow.” (R. L. Stevenson, “Virginius Puerispete” (1881))

The particular factors which underlie the limited productivity of this construction must include semantic factors of the type outlined in chapter 8. In addition, morphophonological factors, such as the ones Gropen et al. (1989) found to be relevant in the case of the ditransitive, and the token frequency of the analogical source may need to be taken into account. The role of each of these factors remains an issue for further research.

An example at the opposite end of the continuum is the way construction discussed in chapter 9. This construction appears to be almost entirely productive. The following examples come from the Oxford University Press corpus:

(42) a. “But he consummately ad-libbed his way through a largely secret press meeting.”
   b. “... nasty gossip about me now sludging its way through the intestines of the society I know...”
   c. “... their customers snorted and injected their way to oblivion and sometimes died on the stairs.”
   d. “... they hoped they too could massage their way to keeping power.”
   e. “Lord Klang craftily joked and blustered his way out of trouble at the meeting.”

As discussed in chapter 9, the few non-occurring verbs (such as the vanilla action verbs go, walk, move) can be accounted for by general semantic constraints on the construction. Interestingly, the token frequency of this construction is low, with one example occurring approximately every 40,000–56,900 words in the Lund Corpus of conversational texts and the Wall Street Journal. This provides support for the idea that productivity has little to do with token frequency and more with type frequency.
The range of differing productivity illustrated by these examples is exactly what we would expect given the usage-based model of grammar described above. That is, learned and thus stored resultative cases are few and only dot the semantic landscape; little or no clustering of examples is attested. Therefore novel extensions sound novel and are not fully idiomatic, unlike such extensions of the ditransitive cases as fix or Marantz's novel verb shin.

At the same time, attested way construction examples seem to span the spectrum of semantic space, given the general constraints imposed by the construction. Since the construction has such a high type frequency of attested verbs, novel verbs are freely used in it.

5.5.3 Productive Links

As discussed in chapter 3, relations between constructions, represented by various types of inheritance links, are also objects in our system. Different instances of a given link occur with different type frequencies, just as different instances of a particular construction occur with different type frequencies.

For example, the causative–inchoative relation, which is represented by a kind of subsumption inheritance link, occurs between the caused-motion and intransitive motion constructions, the resultative and the intransitive resultative constructions, and the simple causative and simple inchoative constructions. Therefore this particular link would be said to have a type frequency of at least three.

Some of the polysemous extensions we have seen occur in both the ditransitive construction and the caused-motion construction. The type frequency of each of these polysemous links is increased with every construction which is extended in the same way. Because productivity is directly correlated with type frequency, the higher the type frequency, the more likely a particular inheritance link will exist between pairs of new constructions that are relevantly similar to the pairs of existing constructions which the inheritance link already relates. In the limiting case, a link will apply fully productively, yielding extensions every time a novel construction is encountered, as long as that construction satisfies the particular semantic characteristics of the existing instances. In this case, the link between the two constructions is quite analogous to a rule, in that the existence of one form can be used to predict the existence of the other form.

For example, the passive construction, discussed briefly at the end of chapter 2, is instantiated by many different particular versions, each corresponding to an active construction with the relevant semantics (the active construction must have at least two arguments, with one being higher on the role hierarchy than the other). Because the type of link between active and passive constructions occurs between so many different active and passive pairs, it has an extremely high type frequency. Therefore the passive link is, in effect, rule-like in its application.

5.6 Conclusion

The account proposed here to explain the partial productivity of constructions involves two types of learning mechanisms. The first is a type of indirect negative evidence, based on the hypothesis that every construction contrasts with every other construction. Therefore, upon hearing a verb in a construction that must be considered non-optimal given the current context, the learner tentatively hypothesizes that the verb cannot occur in what would be the preferred construction. The reasoning is roughly, "If that construction could have been used, I guess it would have been used; therefore maybe it can't be used." Upon witnessing the verb in a non-optimal construction, given the context, a number of times, the learner's hypothesis that the verb cannot occur in the optimal construction is strengthened. This strategy was first proposed by Pinker (1981); his later rejection of this strategy (Pinker 1989) was argued here to have been unwarranted.

The second learning mechanism, presumably working in tandem with the first, draws largely on recent work by Pinker (1989) and the related experimental evidence of Groen et al. (1989). Specifically, the need to circumscribe narrowly defined semantic subclasses characterized by local productivity is acknowledged.

The account proposed here differs somewhat from Pinker's and Groen's in that on the present account, the narrowly defined subclasses are understood to be clusters defined by semantic and morphophonological similarity that are conventionally associated with the construction, as opposed to subclasses that are conventionally allowed to undergo a lexical rule. Moreover, on the account presented here, the verb classes are interpreted as implicit generalizations over learned instances in order to account for small nonproductive subclasses, differences in judgments even within narrowly defined classes, and the existence of positive exceptions such as envy and forgive. Viewing verb classes as clusters of cases in an associative memory also allows us to assimilate other constructions which involve markedly more or less productivity. In particular, it was shown that we might actually expect the fact that the resultative construction is productive only to a limited degree, whereas the way construction is almost fully productive.

It may seem that by allowing the knowledge of whether a verb is used in a particular construction to be stored, we undermine the existence of the construction as an independent entity. That is, if we need to posit the fact that kick
can be used with the ditransitive construction as a separate piece of grammatical knowledge, why not instead posit a new sense of \textit{kick}, along the lines suggested by semantics-changing lexical rule accounts (cf. the discussion in chapter 4)?

The reason for postulating constructions is analogous to the reason why other researchers have wanted to postulate a lexical rule; in order to capture generalizations across instances. Moreover, it is claimed here that what is stored is the knowledge that a particular verb \textit{with its inherent meaning} can be used in a particular construction. This is equivalent to saying that the composite fused structure involving both verb and construction is stored in memory. By recognizing the stored entity to be a \textit{composite} structure, we gain the benefits described in chapter 1 over a lexical rule account. For example, we avoid implausible verb senses such as \textit{"to cause to receive by kicking."} It is the \textit{composite} structure of verb and construction that has this meaning. We also allow other syntactic processes to refer to the inherent lexical semantics of the verb. Thus we do not lose the information conveyed by the verb, because the verb is not changed into a new verb with a different sense.

6 The English Ditransitive Construction

6.1 Introduction

The ditransitive construction has already been discussed with respect to its polysemy (chapter 2) and its partial productivity (chapter 5). In this chapter, I concentrate on particular semantic constraints and metaphorical extensions of the construction. Highly specific semantic constraints are associated directly with the ditransitive argument structure, revealing a more specific semantic structure than is generally acknowledged. In particular, the central sense is argued to involve transfer between a volitional agent and a willing recipient. Several systematic metaphors are identified and associated with the construction, showing that expressions such as \textit{Mary gave Joe a kiss} and \textit{Mary's behavior gave John an idea}, which are often assumed to be idiosyncratic, are instances of a large and productive class of expressions that are based on systematic metaphors.

Before getting to those constraints, however, evidence that a construction is indeed required in this case is reviewed.

6.2 The Existence of the Construction

Following the program laid out in previous chapters, we need to show that aspects of the syntax or semantics of ditransitive expressions are not predictable from other constructions existing in the grammar. First, to see that the construction contributes semantics not attributable to the lexical items involved, consider the verb \textit{bake} when used ditransitively:

(1) Sally baked her sister a cake.

This expression can only mean that Sally baked a cake with the intention of giving it to her sister. It cannot mean that Sally baked the cake so that her sister wouldn't have to bake it; nor can it mean that Sally baked the cake as a demonstration of cake-baking, or that she baked a cake for herself because her sister wanted her to have one. Unless we associate the \textit{"intended transfer"} aspect of meaning to the construction, we are forced to say that \textit{bake} itself means something like \textit{"X intends to cause Y to receive Z by baking."} This \textit{"transfer sense"} of \textit{bake} would be posited only to avoid attributing aspects of the semantics to the construction. The positing of such ad hoc verb senses which only occur in a particular construction was argued against extensively in previous chapters.