8 The Child as Word Learner

SUSAN CAREY

What is learned when a word is added to a child's vocabulary? Minimally, the child learns that it is a word: he enters it into his mental lexicon. Also he must learn its syntactic properties-its part of speech and its syntactic subcategorizations. He must learn its place in lexical structure-its relations to other words. He must learn its semantic properties, its roles in determining entailments, and its referential properties. Finally, the conceptual underpinnings that determine its place in the child's entire conceptual system must also be learned. Some inkling of the complexity of the semantic and conceptual representations that are mapped onto words is provided in Chapters 2 and 5, and the expanded role of lexical entries in syntax is the focus of Chapter 1. In Chapter 7 it is argued that these recent developments make the acquisition of syntax less of a mystery. But just those developments make more prodigious the child's feats as a word learner.

By age six the average child has learned over 14,000 words.[1] On the assumption that vocabulary growth does not begin in earnest until the age of eighteen months or so, this works out to an average of about nine new words a day, or almost one per waking hour. So we have a puzzle. Learning even a single new word involves representing a great deal of information, yet the child learns an average of nine words a day. As we shall see, part of the resolution of this puzzle is that the learning of every word involves long-term developmental processes. In this chapter, I will focus on those processes, including the nature and significance of the incomplete representations a child develops along the way.

The Process of Learning a Single New Word

Where does the process of word learning begin? In the earliest phases of vocabulary learning there is often much concentrated teaching and

Research reported in this chapter was partially supported by a grant from the Public Health Service (GM 21796) to The Rockefeller University and a grant from NIH (2-R01-HD5168) to Massachusetts Institute of Technology. I would like to thank Katherine Miller, George Miller, and Ned Block for comments on an earlier version of this manuscript.

1. The estimate of 14,000 words includes inflected and derived words and is based on comprehension vocabulary. For root words only, the estimate falls to around 8,000, or roughly five new root words a day (Templin 1957; see also Miller 1977).

drill on routines like "bye-bye" and on naming objects (see Nineo and Bruner 1977 for some illustrative examples.) By the time the child is over two years old, however, such drill is not typical. The child must learn most of the 14,000 words from hearing other people use them in normal contexts. If so, only two sources of data are available to him: the linguistic context in which the word occurs and the situation in which it is used. Of course, it is the child's representation of each that affects what he will learn. The child's already existing linguistic and conceptual systems determine those representations; the word-learning process must start there. The only way to begin to account for the child's wizardry as a word learner, given the sheer weight of how much there is to be learned, is to grant that the child brings a great deal to the "original word game" (to use Roger Brown's phrase).

Suppose, for example, that somebody were to show you a picture of an unfamiliar action being performed upon an unfamiliar substance in an unfamiliar container, and you were told the picture depicted how *to sib.* Your knowledge that *to sib* is a verb form and that verbs can refer to actions would lead you to the hypothesis that *sib* is a verb naming the strange action. Told that the picture depicted *a sib* or *some sib, you* would arrive at different hypotheses about the meaning of *sib.* Brown (1957) presented four-year-olds with just such an experience and found that they formed the same kinds of hypotheses. Thus, four-year-olds must have conceptual distinctions between actions, things, and stuff, in terms of which they can represent the picture. They must also have the syntactic distinctions between verb, count noun, and mass noun, and must know some surface syntactic indicators for each. This example illustrates only some of the knowledge brought to bear on learning the new word *sib* from just a few exposures. That four-year-olds have already developed that knowledge is perhaps not surprising. But when did they acquire it?

Following up on Brown's demonstration, Katz, Baker, and McNamara (1974) demonstrated that much younger children rely on such knowledge in learning a new word. At the age of seventeen months, girls (but not boys) already have a distinction between common and proper nouns and they already know and can exploit in word learning a syntactic indication of whether a given noun is common or proper. The syntactic cue is the presence or absence of an article. For instance, if I say to you, pointing to a dog, "Look, there's Corgi," you rightly assume that *Corgi* is the dog's name. If I say, "Look, there's a corgi," you assume that *corgi* is the name of a breed of dog, or kind of animal. Katz, Baker, and McNamara showed also that their young subjects could exploit a conceptual distinction between things likely to have individual names and things not likely to. Dolls belong in the former category; boxes, in the latter.

Their demonstration was elegant and simple. They introduced children to a new doll. To those in one group they said, "See what I've brought you. This is Dax." For this group, *Dax* was syntactically marked as a proper noun. Children in a second group heard, "See what I've brought you. This is a dax." Another, similar doll was also present; in subsequent play, the child was asked to "show Dax (or "a dax," depending upon which group she was in) to Mommy," or "pick up Dax," or "feed Dax," and so on. The result was that if the child was in the common-noun group, she used both dolls interchangeably - the one the experimenter had called *a dax* and the other similar one present. If she was in the proper-name group, she picked only the doll that had been called *Dax.* This observation establishes that girls as young as seventeen months knew two lexical categories for nouns and could exploit a relevant syntactic cue in placing a newly learned noun in the correct category. Another two groups of baby girls were needed to demonstrate the role played by knowledge of what kinds of things are likely to be given proper names. The experiment was repeated, with fancily decorated boxes instead of dolls. In this case there was no difference between the two groups of children. Whether the box had been introduced as *a dax* or *Dax,* the child treated the word as a common noun and picked both boxes equally often when asked to "show Mommy Dax" or "show Mommy a dax." They did not expect an individual box to be given its own name.

These little girls learned the new word, mapped it onto a referent, understood novel sentences containing it, and learned its syntactic category (in the doll case) from just a few exposures to the word. This is word-learning wizardry. Recognition of the knowledge brought to the word-learning task gives a momentary illusion of reducing the mystery of the efficiency of the process. But this is only an illusion, for now we must account for how such knowledge had been acquired by the age of seventeen months. I will not speculate about that mystery.

How general is the finding that young children pick up new words with such ease? There have been relatively few studies. Braine (1971) reversed some aspects of the above two demonstration experiments.

He introduced two nonsense syllables *(niss,* a kitchen utensil the child was allowed to play with, and *seb,* finger walking that typically ended in tickling) without any sentential context. The words were clearly spoken, in isolation, while Braine was sebbing, or indicating a niss, and were "rapidly taken up into the speech" of his barely two-year-old daughter. She always used *niss* in contexts where other nouns appeared; she used *seb* both in contexts where verbs appeared *(Seb Teddy,* for example) and in contexts where nouns appeared *(More seb).*

The contexts of the Brown (1957) study and the Katz et al. (1974) study were different from the use to which I am putting them here. Both concerned the issue of semantic correlates to syntactic subcategorizations - the question whether such correlates could provide the basis of the child's abstraction of part of speech. Braine's (1971) study spoke to this issue; he presents arguments against the semantic basis, in development, of syntactic subcategorizations. In my use of these data, I am arguing that the child's knowledge of syntactic cues to part of speech as well as his knowledge of whatever semantic regularities there may be both play a role in his efficiency at learning new words.

In many studies using the technique of teaching nonsense words, the child appears far from a word-learning wizard. (See, for example, Werner and Kaplan 1950; Klatzky, Clark, and Macken 1973; Nelson 1973.) In all these studies, however, the child faced various unusual departures from normal word-learning contexts. In Werner's study, no natural nonlinguistic context was given. The child was to learn the meaning of a new word from a series of linguistic contexts alone, such as "A wamplum is long and skinny" and "A wamplum is used to poke things." This may be an interesting problem-solving paradigm, but it hardly mimics the natural word-learning situation. In the studies by Nelson and by Klatzky, Clark, and Macken, several new words were taught at once. In Nelson's case 16 new nouns *(snorkel, compass, handcuff,* for example) were taught to two-year-olds. All the words were taught in the same sessions, increasing the difficulty of keeping the words straight-of remembering what was a compass and what a snorkel. The Klatzky, Clark, and Macken study introduced an additional difficulty: the words taught were not possible lexical items in English. The child was shown four sticks, varying in length, the experimenter pointing to one of them (x in Figure 8.1). Two words *(zup* and *grod)* were taught as follows: The experimenter

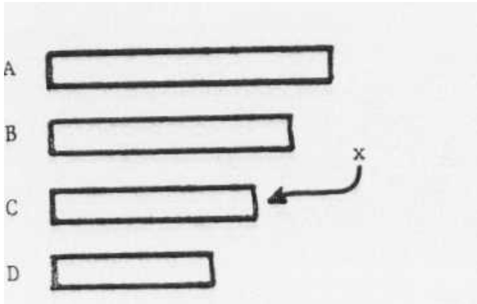

Figure 8.1 Stimuli from Klatzky, Clark, and Macken study

said "Show me one that is zup," giving the child feedback that the correct answer was D, and "Show me one that is grod," giving the child feedback that the correct answer was A or *B. Zup,* then, meant "longer than this one," and *grod,* "shorter than this one." But in English there is no single lexical item that packages comparative information in that way-such a word violates English constraints on possible lexical items. If the child has knowledge of those constraints, it is not surprising that these words were difficult to learn.'

None of the studies reviewed above, neither those in which the child picked up a new word effortlessly, nor those in which he had great difficulty, was designed to study the word-learning process, extended over time. Nor did any of the researchers probe for partial meanings (it may be that the children had learned something of the meanings of *wamplum, zup,* or *snorkel-had* at least entered these forms into their lexicons-after single exposures). None of the studies tried to find the limits of the child's proficiency: the words were explicitly taught; the child was not required to pick them up from their use in normal contexts. Moreover, none of the words filled a gap in a well-structured lexical domain, so they could not be studied relative to the child's knowledge of that domain. Finally, none of the studies tested the permanence of the acquisition of *sib, dax, piss* or *seb.*

Anecdotal evidence suggests that the child's proficiency at acquiring a new word could be put to a much stronger test. In an experiment designed to explore the claim that there is a point in development when the word *less is* incompletely represented as a synonym of *more* or *some* (Carey 1977), children were asked to

2 The study by Klatsky et al. (1973) was not intended to model normal word-learning. I am merely speculating why children found learning these new words so difficult. The points I raise are irrelevant to the goals of the original study.

"Make it so there is tiv to drink in this glass." *Tiv* was not emphasized, and more than half of the children simply added water to the glass or poured some out. Very few asked what *tiv* meant. The instruction was repeated three times, once using the word *more* instead of *liv,* once using *less,* and once again using *tiv.* Thus, the children heard *tiv* only twice. Two to six weeks later, fourteen of the three- to four-year-olds who had been in this study were in another study requiring that they generate opposites to whatever a puppet said; for example, if the puppet said *up,* the child said *down.* The research assistant collecting these Opposites data tacked on *tiv* to the end of the list. Most children treated it as a nonsense syllable, either giving no response or producing a rhyme. But three of the fourteen responded *more* (or *less)* and of these only one had asked what liv meant (and had been told *less).* For one of these children, six weeks had intervened between the *tiv* study and the Opposites task. We are close here to the limits of the child's proficiency-eleven children did not remember *tiv.* But the three who did clearly illustrate the way in which preschool children are magnets for new words.

Building on this observation, Elsa Bartlett and I decided to study the process of word learning when the word was part of a structured lexical domain.

Mapping between lexical and conceptual domains Corresponding to the distinction between word and concept is the distinction between lexical domain and conceptual domain. The lexical domain is simply the structured set of words that "covers" the conceptual domain. The conceptual domain is the relevant part of the internal representational system in terms of which the person or animal describes and understands the world and his own actions in it (see Fodor 1975 for an extended discussion of the internal representation system). Take, for example, the lexical domain of color words and the conceptual domain of mental representation of colors.

The structure of the conceptual domain can largely be captured by the three dimensions of the color solid: hue, saturation, and brightness. But other facts also are represented in the conceptual domain of color, facts such as the characteristic colors of particular objects; sometimes such facts may play a role in the recognition of those objects. Psychologists can learn about the conceptual domain without recourse to language; for example, perceptual matching and discrimination abilities and the phenomena of color contrast, color mixture, and color blindness suffice to reveal the properties of the color solid.

Some sample facts about the lexical domain are that the words *red, green, blue, scarlet,* and *chartreuse* are hyponyms of the word *color* and furthermore that *scarlet* is a hyponym of *red* and *chartreuse* is a hyponym of *green.* Ten or eleven primary hue words can be combined productively to form terms like *yellowish green* or *yellow green,* for example. Properties of the color solid other than hue are handled lexically by adjectives such as *dark, light,* and vivid.

This characterization does not begin to do justice to either the lexical domain or the conceptual domain. (See Miller and Johnson-Laird 1976 for complete analyses.) Nevertheless, this sketchy characterization will suffice for a few crucial observations. First, there are many things we know about colors that are not captured in the structure of the color lexicon. Closely related to this point, the lexical domain is highly structured as to possible words. There could never be, in any language with a color lexicon, a word whose meaning was the concept "the color of the flowers of lilac bushes," the reason being that the colors picked out by that concept (white, lilac, blue, and purple) are not adjacent on the color solid, and hue words can name only continuous regions of the rainbow. On the other hand, a color word meaning, roughly, "green or blue," would be allowed by the rules of lexical formation of color words; indeed, some languages have such a term. Finally, it should be emphasized that each domain (the conceptual and the lexical) has its own identity and structure.

For the child learning language, there can be development within each domain separately before any mapping of one onto the other is begun. Clever psychophysicists have shown that the color solid captures the structure of the conceptual/perceptual domain of color for prever-

color is a supernym of at least some hue names. That the

mapping between color words and perceived colors had not yet begun was shown by a child's failure to consistently pick a color chip that matched the name of any of the colors being probed and his failure to consistently produce the correct color name for any of the color chips.

For our initial pilot study on the process of lexical acquisition, Bartlett and I decided to use the domain of color words. Our goal was to explore the limits of the preschool child's capacity as a word learner, There would be no explicit teaching; the new word would he introduced in an entirely neutral context. Initially, each child would have only one exposure to it and at least a week would intervene before any probing to determine whether it had been taken up into any of the children's lexicons. The color lexicon was chosen mainly because the subjects were currently being studied by Bartlett in her work on ordinary color words, allowing us to see how learning a new color word would restructure the child's existing lexical and conceptual color domains. The subjects were fourteen three- to four-year-olds, all of whom had begun the mapping from color words onto colors. One knew only one word-color correspondence (green); the rest reliably comprehended and produced from six to eleven color words.

First we had to choose the color to be named. Since a case can be made that the color olive is a hue on the same level in the color hierarchy as blue, red, yellow, brown, and white (one can say *greenish olive* or *brownish olive* but not *reddish scarlet* or *orangish scarlet),* and since the name of the color was unknown to all fourteen of our subjects, olive was chosen as the color. Since some of the children might know what olives are, an unknown word (an actual color name), *chromium,* was chosen as the name for olive.

By far the most important choice in this pilot was the introducing event. The subjects attended a nursery school run in George Miller's laboratory at The Rockefeller University, and the teacher in the school agreed to do the introducing. One tray and one cup in the classroom were painted olive. There was one identically shaped cup (red) and one identically shaped tray (blue). In a natural context such as setting up for a snack, each child, individually, was told, "Bring me the chromium tray, not the blue one, the chromium one" or "Bring me the chromium cup, not the red one, the chromium one." Since we wanted a strong test of the child's efficiency as a word learner, we wanted to avoid explicitly calling the child's attention to the new word. To carry out the task (bringing the correct cup or tray) the child need not focus on the word *chromium* at all, for "not the red (or blue) one" was sufficient to determine the response. To further ensure a strong test of the child, subsequent production and comprehension tests were administered in a different room by a different person from the teacher who had introduced the new word. Although great care was taken to minimize the child's sense that he was being taught a new word, the introducing event provided syntactic and lexical cues that *chromium* was a color word as well as contextual cues to which color it named. That is, the information was adequate to the full mapping.

Prior to the introducing event, each child was given a production test, which included an olive color chip. Most called it *green,* a few *brown.* These responses served as a base line against which to assess a later production test.

All of the nursery school sessions were being video-taped, allowing full analysis of each child's introduction to the word *chromium* and the monitoring of any spontaneous use of the new word in the classroom as the pilot study progressed. What happened?

At first exposure, only one child had trouble picking the correct tray or cup. Four spontaneously repeated some approximation to *chromium,* and most asked for some confirmation: "You mean this one?" These children clearly flagged "new word!" upon hearing so phonological sequence with no current lexical entry. And they even though they did not need to focus on that word to carry out their

i

[illegible]

Six weeks after the teaching (five weeks after the comprehension test) the children were given a production test. What happened that test surprised us. Eight of the fourteen children changed their

response from their base-line response. Two said they did not know what to call olive, indicating that they had learned that olive had a different name from *green* but could not remember what. The rest used another color name for it *gray, blue,* or *brown*-and in each case the child chose a word he had not yet stably mapped onto a color. The child had learned and retained for over a month that olive is not called *green;* in searching his lexicon for a name to call it, he found another color word with no stable referent which was more highly accessible than the new word *chromium.* Thus, for these eight children at least, the process of restructuring the conceptual and lexical domains had already begun.

The pilot was continued, with a few further teaching experiences and comprehension and production tests. Two different routes to full acquisition were identified. Some children adopted a false hypothesis about the structure of the lexicon-that *chromium* was a synonym of *green.* In comprehension tests they often picked green when asked to choose the chromium chip. In production tests they usually called both olive and green *green,* but sometimes called focal green *chromium;* one child called focal green light *chromium.* For these children, working out the correct lexical relation between *chromium* and *green* was an extended process, not completed in some cases after 18 weeks of testing. Let us call this group the "false synonym" group. The other group, in contrast, knew from the beginning that olive needed its own name. In comprehension tests they always picked olive for the word *chromium. In* production tests they said they could not remember the name or chose some color name with no stable referent. Some of the children who adopted this "odd color, odd name" strategy also did not achieve full mastery by the end of testing.

When two children on different paths to full acquisition (one who had adopted the false-synonym (FS) strategy, and the other the odd--color-odd-name (OCON) strategy) tried to communicate in a natural exchange in the classroom, predictable confusion resulted. About halfway through the study the children were making Easter baskets. Material included pieces of yarn and paper that were colored fuchsia, light green, and dark forest green *(not* olive). The FS child, Albert, said, "I want chromium, ah, a green, a green, I mean chromium and that red." The teacher asked the OCON child, Ellen, which one she wanted; Ellen indicated the forest green and said "The dark one." Albert (FS): "No, that's chromium." A third child picked up fuchsia and asked "Chromium?" Albert (FS) took the fuchsia from him, gave him a forest green and said, "No, chromium is a green one." Ellen (OCON), looking puzzled through all this, picked up a forest green piece of paper, and asked the teacher, "Is THIS the chromium tray?" Teacher: "What do you mean?" Ellen (OCON): "It doesn't look the same." Teacher: "The same as what?" Ellen, confused, did not respond.

The answer to the teacher's question is, of course, olive green, the color of the tray and cup. Ellen, an odd-color-odd-name child, had focused on the color olive and always picked olive in comprehension tasks probing *chromium*. What she had learned was that olive was not called *green* and that *chromium* did not name green, so Albert's (FS) calling forest green both *green* and *chromium* was simply not acceptable.

This tentative pilot study, currently being replicated and extended, shows that one can study the process of lexical acquisition. Several lessons emerge. First, lexical acquisition is indeed a very efficient and rapid process. (Only one child appeared never to have learned anything about the word *chromium* or the color olive.) One experience with the word, or at most a very few, sufficed for the remaining children to adopt one of the two partial mappings between the color lexicon and the conceptual domain. We have dubbed this initial, speedy process "fast mapping." Second, after this initial fast mapping, protracted further experience was required before learning was complete, and many children had not progressed beyond the initial mapping by the end of testing. Finally, I would like to stress that there were two parts to the process: (1) the restructuring of the lexicon by finding the right place for the word *chromium*, and (2) the restructuring of the conceptual domain by learning that olive was not included in the category of green or brown, but was a color that had its own name. Full coordination of both of these developments had been achieved by only half of our subjects after several months of weekly experiences with the word.

Given the speed and efficiency of the fast mapping, the slowness of the process of attaining the full mapping is perhaps surprising. After all, the color lexicon is relatively simple, and information adequate to the full mapping was repeatedly presented. Suppose that, on the average, six months is required for the full acquisition of a new word (surely an underestimate, as we will see). If the child is learning nine new words a day, then he is working out the meanings of over 1,600 words at a time. This fact is a clue to the real significance of the fast mapping. What is included in that initial mapping-that the new word is a word, along with some of its syntactic and semantic properties-must allow the child to hold onto that fragile new entry in his lexicon and keep it separate from hundreds of other fragile new entries, and it must guide his further hypotheses about the word's meaning. In the next section we will turn our attention to the nature of the early representations and to the process of moving from such beginnings to full meanings.

Immature Lexical Entries: Missing Features

On one widely held view of lexical development, the missing-feature theory, the lexical entries of children differ from those of adults in being incomplete. Only some of the semantic features that characterize the full meaning are initially represented by the child. According to the missing-feature theory, the process of lexical development consists of adding features until the full entry is achieved. Many recent controversies about semantic development have presupposed this theory-the controversy whether the addition of features proceeds from general to specific or vice versa, for example (Clark 1973; Anglin 1970). Furthermore, the missing-feature theory has provided a framework within which to pose many specific questions about the incomplete entries within particular lexical domains.

In this section two competing hypotheses about what features are initially missed during the acquisition of spatial adjectives will be reviewed. Following this review, the missing-feature theory itself will be put to a test and will be found wanting. In the next section a revision of the missing-feature theory will be developed.

Spatial adjectives The spatial adjectives include: *big, little; long, short; tall, short; wide, narrow; thick, thin; deep, shallow;* and *high, low.* Two aspects of the linguistic structure of this lexical domain are relevant to our discussion. The first is the core comparative structure, which the words in this domain share with all other relative adjectives (*fast/slow* and *heavy/light,* for example). The second is the feature system in terms of which spatial adjectives differ from each other.

The core comparative structure underlies both contrastive uses (*He is a tall man* or *He is a short man)* and comparative uses *(She is taller/shorter than Harry)* of relative adjectives. Both uses lire a dimension of comparison (height in these examples), a standard of comparison (average height of men and the height of Harry, respectively), and a direction from that standard (greater than or less than). The particular adjective specifies both the dimension of comparison and the direction from the reference point. In comparative uses, the reference point is explicitly given; in contrastive uses it is inferred, and is usually some average value. Relative adjectives come in pairs, which differ in their polarity-that is, in their direction from the reference standard. In the pairs listed, the positive-pole words (values greater than the reference point along the dimension of comparison) are given first.

The domain of spatial adjectives that is of interest here is defined by the feature [spatial extent]. I mean this feature to capture the fact that size, height (both altitude and tallness), length, width, thickness, and depth can all be measured in inches (linear inches, square inches, or cubic inches). *Fat* and *skinny,* while relative adjectives are not part of this domain, because the dimension of comparison underlying these two words is not one of spatial extent. An answer to the question "How fat is John?" cannot be "Two feet" nor even "Two hundred pounds." Rather, it must be something like "Very." For *fat* and *skinny* the relevant dimension of comparison is a ratio among spatial extents-relative width or thickness, given some height. For the purposes of this chapter I will call such a dimension "relative shape," which is in contrast to simple spatial extent.

Several linguists have proposed feature systems to capture the differences among spatial adjectives (for example, Bierwisch 1967; see also Miller and Johnson-Laird 1976). Of the words in this domain, *big* and *little* are in some sense superordinate to the rest; each of the other pairs picks out a particular way of being big or little. The dimension of comparison underlying *big* and *little* is spatial extent alone; all other spatial relatives require further features for their specification, such as [vertical]. Bierwisch's system is complex, and it is not obvious that the dimensions of comparison can be totally captured by such a feature system. For example, take the pair *thick* and *thin.* The dimension, thickness, is characterized as the third, or tertiary, dimension whenever length (or height) and width are also specifiable. Doors have height and width, pavement and ribbons have length and width; thickness of each is the third dimension. This analysis applies straightforwardly to many common uses of *thick* and *thin,* but for other uses it applies only abstractly. Consider the

nonmetaphorical senses of *thick skin, thick skull, thick crust.* It is not obvious that thickness is the tertiary dimension of skin, for example, until one imagines the skin (of an orange or an animal) laid out flat as a two-dimensional peel or hide. Then, indeed, thickness of skin is analogous to thickness of doors or ribbons. In other common and systematic uses the characterization of thickness as the tertiary dimension does not apply at *all-thick rope, thick cigarette, thick tree trunk.* There may also be nonsystematic uses, perhaps as in *thick lips.* At any rate, the feature system underlying the adjectives of spatial extent is complex; the child does not face an easy task in mastering it.

Both aspects of the lexical structure of the domain of spatial adjectives-the comparative core and the feature system that specifies the differences among the dimensions of spatial extent-reflect the conceptual system. Underlying the comparative core are the abstract concepts of reference point and polarity, which presuppose concepts of dimension of comparison and zero point. These concepts can be probed nonlinguistically. For example, animals can be taught to choose the smaller, or larger, of two stimuli. It is likely that the standard sizes of objects are represented conceptually for the purpose of object recognition. Presented with a box the size of my desk, I am not likely to entertain the hypothesis that it is a box of Kleenex. Underlying the feature system characterizing the dimensions are many aspects of man's representation of space; concepts like vertical and horizontal, cross-section, and spatial extent itself are reflected in many nonlinguistic sensorimotor routines (see Miller and Johnson-Laird 1976; H. H. Clark 1973).

As the child learns a new spatial adjective, what aspects of its conceptual underpinnings are mapped onto it early and what aspects, if any, take years to work out? Two positions within the framework of the missing-feature theory have been held.

The first position is that the child's initial mapping is between the word and the features specifying the relevant dimension of comparison. The missing feature is polarity, the direction from zero. On this view, both *narrow* and *wide* would have the incomplete lexical entry: [Adj] [comparative] [spatial extent] [-primary] [-vertical], making the two words synonyms (Donaldson and Wales 1970; H. H. Clark 1970; E. V. Clark 1973; Klatzky et al. 1973). The two words need not have identical incomplete lexical entries simultaneously. For example, [+pole] might be added to the representation of *wide* before [-pole] is to *narrow.* In this case *narrow* means what *wide* did before

[+pole] was added to the entry of the latter: *narrow* is a synonym of an earlier representation of *wide.* Suppose the child had such incomplete lexical entries for *narrow* or *wide.* How would he decide which of two boards was wider (or narrower)? Lacking information about polarity, he might pick the board that best exemplifies width in both cases-namely, the wider board (H. H. Clark 1970).

The second position is that the child first maps the word onto the comparative core (including polarity), but the correct dimensional underpinning takes years to work out. Early on, each spatial adjective will be marked only [Adj] [comparative] [±pole] [spatial extent]; gradually the other features-such as [vertical] and [primary]-will be added. On this view, *tall, snide, high,* and so on, are at first all synonyms of *big;* later, *tall* and *high* might be synonymous, [vertical] having been added to each entry. Similarly, at first *short, narrow, low,* and so on, are all synonyms of *little;* later, *thin* and *narrow* might be synonyms (Wales and Campbell 1970; E. V. Clark 1972). All that is required is that each word be synonymous with *big* or *little* at some point during its acquisition. For example, *high* may be marked [vertical] while *tall* and *wide* are marked only [spatial extent]. Of the three, only *tall* and *wide* are synonymous with *big;* before the addition of [vertical], so is *high.*

The dust from a great deal of experimentation has settled, and it is now clear that of the two positions, the weight of evidence strongly favors the second. That is, when a new spatial adjective is learned, its polarity is represented early; it is its underlying dimensionality that takes years to work out. I will review that evidence and then present new data that suggest that although the process of working out the dimensions of comparison is indeed an extended one, the missing-feature theory provides an inadequate account of that process.

The evidence that polarity is present in early lexical entries while information about dimension of comparison is sometimes missing will be presented in three steps: (1) Since *big* and *little* are the first spatial adjectives learned, it must be shown that each is mapped onto the core comparative structure (including correct polarity) very early on. (2) It must be shown that the other spatial adjectives are also mapped onto this core comparative structure (including correct polarity) as soon as each is learned. And (3) It must be shown that this mapping precedes full dimensional specification in the cases of spatial adjectives other than *big* and *little,*

(1) Big and *little* become part of the vocabularies of most children

before they are three years old. Is there a point where the negative polarity of *little* or the positive polarity of *big* i

big and *little.* Each child was then shown a series of objects, a big shoe (size 13) or a little shoe (toddler's size), a big chair or a small chair, a big paper bag or a small paper bag, and asked of each "Is this a big x or a small x?" The big shoe was smaller than the little chair; judgments, if correct, must have been made relative to what the child had represented as the standard-sized shoe, chair, paper bag, and so on. The eight children who had *big* and *little* in their vocabularies made very few errors (90 percent correct as a group; for seven of the eight children the error score was zero or one on seven items). There were no more errors on *little* than on *big.* Not surprisingly, those children whose vocabularies did not include *big* and *little* performed at chance (Carey and Potter 1976).

The children's ability to use *big* and *little* contrastively (in constructions like "It's a big chair") establishes that as young as age two, children have representations of the standard sizes of common objects. The two words, *big* and *little,* seem to be acquired at the same time and are mapped onto the core comparative structure (including polarity) immediately.

(2) That other relative spatial adjectives also are mapped onto this core comparative structure early was demonstrated analogously: two- to four-year-old children were shown objects such as very long pencils or very short pencils and asked of each "Is this a long pencil?" or "Is this a short pencil?" While mastery of the different antonymous pairs varied *(tall/short* is learned before *wide/narrow)* the children did just as well on negative-pole words as on positive-pole words, showing no evidence of learning the negative-pole words later or of failing to represent their negative polarity (Carey 1976).

These observations indicate that the child has mapped spatial adjectives onto the core comparative structure (including polarity); three additional paradigms establish that this mapping precedes the correct dimensional mapping.

(3) Shown an array of objects varying in two dimensions (see Figure 8.2) and asked to indicate the shortest one, for example, errors respected polarity but not dimension (Brewer and Stone 1975; Carey 1976). That is, the child picked either of the two little ones, A or B (little in height and width respectively). If the child had mapped [vertical] onto *short,* but had not yet represented its polarity, one

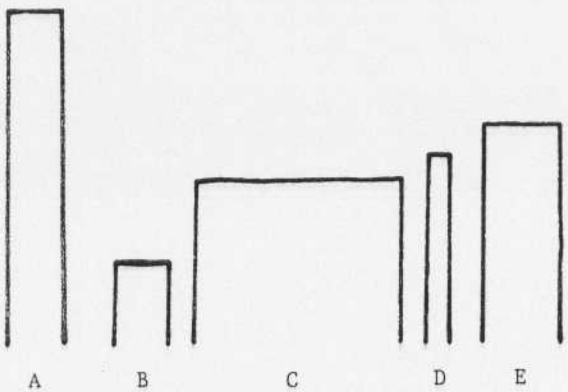

Figure 8.2 Stimuli varying in two dimensions

would expect the erroneous choice of D. This did not occur. Another class of errors in my data suggest that the confusion over dimensionality extended even further. When the children were asked to pick the tallest block, most errors were choices of B-the narrow block-which happened to be the one with the greatest ratio of [illegible] have failed to map simple spatial extent onto *tall*, but to have adopted the false hypothesis that *tall*, like *fat* and *skinny*, specifies relative shape.

A second paradigm supporting the position that the underlying dimensionality of each spatial adjective is learned slowly was adapted from Ervin-Tripp and Foster (1960). Shown an array such as that shown in Figure 8.3a, children were asked, "Is one of these fatter than the rest, or are they all the same in fatness?" (Care was taken that the child understood this locution. Each child was pretrained on arrays like those in Figure 8.3b, where the question was "Is one of these bigger than the rest, or are they all the same in bigness?" or "Is one of these darker than the rest, or are they all the same in darkness?") On over 60 percent of the test trials with inappropriate dimensions, the child indicated the item that respected the polarity of the probed adjective (that is, picked A in Figure 8.3a as fatter than the others). The remaining responses were predominantly "all the same in fatness" (Carey 1976).

Finally, the third paradigm indicating that polarity is firmly represented before the underlying dimensionality is worked out is the elicitation of opposites. If three- and four-year-old children are asked for opposites of spatial adjectives, they often err, but their errors

almost always respect the polarity of the correct response. Most commonly, if probed with a specific dimensional adjective, the child responds *big* or *little*. To *wide* he responds *little;* to *high, little;* to *deep, little*. And to *low, skinny,* and *short,* he responds *big* (E. V. Clark 1972; Carey 1976). Other errors involve responding with an incorrectly specific adjective: to *high* he responds *short;* to *low, tall;* to *thick, skinny;* to *low, fat;* to *tall, long*. Most of these errors, too, respect correct polarity. Some of the errors suggest an intermediate stage in the representation of dimension; for example, errors such as *high :short* and *low :tall,* support the hypothesis that the child knows these words apply to the vertical dimension, but their lexical entries still lack the features that distinguish *high* from *tall* (E. V. Clark 1972).

It seems firmly established, then, that the mapping from a newly learned spatial adjective onto the core comparative structure, including polarity is established early while the exact dimension of comparison takes very much longer-even years-to work out.

Why should the core comparative structure (including polarity) be part of the early mapping, whereas the dimensionality takes much longer to work out? There are probably several answers to this question.

The child already knows the words *big* and *little* before he learns any of the specialized spatial adjectives. Since the core comparative structure (including polarity) is part of his early representations of *big* and *little,* these features are already available as lexical organizers when the child encounters a word like *wide* or *low*. By "available as a lexical organizer," I mean already part of the lexical entry of some word. Although the features underlying the dimensionality of the

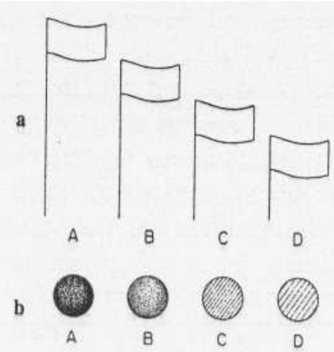

Figure 8.3 Stimuli for inappropriate dimension study

domain of spatial extent are part of the child's conceptual system, their linguistic relevance might not yet be recognized. That is, the *child* might not yet realize that the spatial concepts mark contrasts between words, but because of his knowledge of *big* and *little* (and possibly other relative adjectives like *fast* and *slow)* he does realize that [+pole] and [-pole] mark linguistically relevant contrasts. It is not unreasonable that features available as lexical organizers are mapped onto new words more easily than those that are not yet so available.

But how is the child to know, upon first hearing a new spatial adjective, that the features contained in *big* and *little* are relevant to it'? As pointed out earlier, the child's representations of both linguistic and nonlinguistic contexts provide him with information about a newly encountered word. That the spatial adjectives are mapped onto the comparative core has syntactic consequences, and some linguistic contexts give cues to the adjective's comparative status (for example, comparative and superlative constructions, and continuum-naming constructions such as "five feet tall" or "How tall is he?"). The child must know at least some of the syntax of comparatives-that required for his early uses of *big* and *little.* It is likely that he is able to use these syntactic cues in just the efficient way he uses the absence of an article as an indicator of proper-noun-hood.

This explanation of why polarity might be more learnable than dimensionality depends upon the child's mastery of the syntax of comparative constructions and of the lexical entries for *big* and *little* prior to his acquisition of the specialized spatial adjectives. The child's original, and early, learning of *big* and *little* has not been explained. This accomplishment should be studied further.

Representing polarity first, while leaving dimensionality to be worked out slowly, can be accounted for in terms of corrigibility also. The child minimizes both production and comprehension errors. In production, if he does not know the exact word, he can simply use the more general *big* or *little.* In comprehension, the dimension in question is usually obvious from context-the objects talked about vary in only one dimension or are markedly bigger or smaller than absent standards in only one dimension. Therefore, the child can tolerate vaguely characterized dimensions of contrast. If he had misrepresented the polarity, he could not avoid errors and would be likely to be corrected by his parents (or find himself with the opposite of -hat he wanted on some occasions).

That the dimensionality underlying spatial adjectives is worked out slowly during development has been established, but the nature of the immature lexical entries and the process of lexical growth have not yet been discussed. The results presented are consistent with the missing-feature theory: most errors are as would be expected if specialized spatial adjectives were synonyms of *big* or *little.* Nevertheless, the results suggest an alternative beginning as well: some children may initially map [relative shape] onto *tall,* making it a synonym of *skinny*. In this case, as in the cases where *fat* and *skinny* initially include the feature [spatial extent], the process of lexical growth would involve unlearning as well as adding further features. This alternative requires only a minor revision of the missing-feature theory. Both routes to the acquisition of *tall,* for example, include partial representations such that *tall* initially has a lexical entry identical to that of some other word.

Although the missing-feature theory is consistent with the data, it is greatly undersupported. According to the theory, lexical entries of spatial adjectives initially contain *only* the feature [spatial extent] for the specification of dimension; later on, the lexical entry for *high,* for example, contains only the features [spatial extent] and [vertical]. It is important for the theory that no other information about dimension be represented; otherwise the predictions about synonymy would not obtain. The weakness is that the data supporting the theory in the case of spatial adjectives is almost entirely cross-sectional, each child contributing one response for each adjective. Certainly, a child will give *big* as an opposite to *short* and another child will pick the narrowest block when asked to pick the shortest. But are we justified in concluding that these two children had represented nothing more in their entries about the dimension of comparison underlying *short* than [spatial extent]? That is, is *short* really a synonym of *little?* Further evidence is needed to establish such a strong claim: at the very least, a child should make consistent errors on several different tasks that diagnose, in different ways, the putative partial entries.

To check for such consistency, Tom Considine and I presented sixteen three- and four-year-olds with the five tasks shown in Table 8.1. Each task involved many adjectives, some of them not spatial relatives. *Tall, short, fat,* and *skinny* were common to all five tasks, yielding 64 cases to analyze for consistency. To continue with the example of short, if short is a synonym of little the child should make no errors on the first two tasks of Table 8.1: a short piece of string

Table 8.1. The Five Tasks to Probe Consistency of Errors on Spatial Adjectives

Task	Sample	
1. Appropriate Dimension (contrastive)	Stimulus:	A drawing of a tall man, or a drawing of a short man
	Question:	"Is this a tall man'?"
2. Appropriate Dimension (comparative)	Stimulus:	A drawing of four flagpoles, differing only in height (Figure 8.3a)
	Question:	"Is one taller than the others, or are they all the same in tallness?"
3. Opposites	Stimulus:	"1 say *tall.*"
	Question:	"You say ?" (*short* is expected)
4. Two Dimensions	Stimuli:	Five or six blocks, varying in both height and width (Figure 8.2)
	Question:	"Which one is the tallest one?"
5. Inappropriate Dimension	Stimulus:	A drawing of four ladies, differing only in girth
	Question:	"Is one taller than the others, or are they all the same in tallness?"

will be judged *short* (it is indeed little), and the shortest flagpole will be chosen when the child is asked "Is one of these shorter than all the rest, or are they all the same in shortness?" (it is the littlest). In contrast, errors would be expected on the next three tasks of Table 8.1: on the Opposites task, the child might give to *short* a response such as *big* or *high* or *fat,* respecting polarity but violating specific dimensionality; he might also pick the narrow block when asked for the short one on task 4; and he might also judge that the thinnest lady is shorter than all the rest on task 5. Thus, the strongest evidence for the missing-feature theory would be the following pattern: no errors on the first two tasks of Table 8.1 and errors on the three remaining tasks, all consistent with the same incomplete lexical entry.

In 25 of the 64 cases analyzed for consistency the child made no errors, indicating rather complete lexical entries for those adjectives. Thus there were 39 cases in which errors were made and where the predicted pattern of errors might possibly be observed. Only one

yielded the full pattern, and that was on *short.* No child gave the analogous pattern on *tall, fat,* or *skinny.*

It is possible that even when a child's lexical entry includes only the feature [spatial extent] as dimensional specification, he still might be correct on one or more of the three tasks where errors are predicted. For example, on the Opposites task (task 3) he might produce the correct opposite because of learned associations between *tall* and *short* or between *fat* and *skinny.* Or on the Two Dimensions task (task 4) he might pick the correct block because it is one of two exemplifying the polarity of the word in question; this task was presented twice with different stimuli, however, so being correct required two correct choices. A child with the hypothesized partial entry should never be correct on the Inappropriate Dimension task. Thus, support for the missing-feature theory need not require the full pattern of predicted responses; two errors consistent with the same partial meaning would be strong support.

In 23 of the 39 cases where errors were made, there was only a single error. Even when the error was consistent with the missing-feature theory (for example, responding *big* as an opposite to *short),* these single-error patterns suggest that the child represents more about the dimension of comparison than [spatial extent]; he apparently calls on additional dimensional knowledge on other tasks involving the same word. Furthermore, some of the single-error patterns were actually inconsistent with the theory: the child was correct on the Inappropriate Dimension task (task 5), and in a few cases, the child responded correctly on all three of the tasks that are diagnostic of incomplete dimensional representations but made an error on task I or 2.

If the single-error cases are equivocal, the sixteen remaining cases where the child made two or more errors are not. In only four of these cases, including the one full predicted pattern on *short,* were the errors consistent with each other as reflecting the same putative incomplete representation. The inconsistencies were of several types. First, on one task a child might make an error that seemingly reflected the relative-shape partial meaning (for example, he might pick the skinny block when asked for the tall one on the Two Dimensions task) and on another task make an error that seemingly reflected the simple spatial-extent partial meaning (picking the fattest lady when asked if one is taller than the others on the Inappropriate

Dimension task). In a second type of inconsistency, the child might err on one of tasks 3, 4, or 5, seemingly because he had one of the putative partial meanings, yet also err on task 1 or 2, where he should have been correct, given that partial meaning.

These data, although scanty, are clear. When each of the five tasks is considered alone, the results support the missing-feature theory. Errors, when they occur, are consistent with the view that the child has represented the meaning of *tall,* for example, as if it were a synonym of *big* or *high* (or of *skinny,* where the child seems to think that *tall is* a relative-shape predicate.) However, the across-task consistency required to support such a view is not found.

Thus, something is amiss with the missing-feature theory. There is evidence against stable incomplete lexical entries: *short* may never mean *little.* The child seems always to represent more about the dimension of comparison underlying each spatial adjective than is captured in the feature [spatial extent]. But whatever more is represented, it does not protect the child against errors that indicate confusion about the dimension of comparison. What might early lexical entries for spatial adjectives be like such that they would have these properties?

Immature Lexical Entries, Missing Features and Haphazard Examples

The immature lexical entries for spatial adjectives might contain information about some particular objects to which each adjective applies. The entry for *tall* might include that it applies to buildings and people; of *short* that it applies to hair, people, and distances; of *high* that it applies to chairs and shelves; and so on. Of course, the child must represent the dimension of variation within each object that is relevant to each adjective's use. Thus, sample lexical entries might be:

tall: [Adj] [comparative] [+pole] [building, ground up; person, head to toe]

short: [Adj] [comparative] [-pole] [person, head to toe; hair, root to end; distance, direction of motion]

On this view, the child learns, object by object and particular part by particular part, what spatial adjective applies to what kinds of variation. (Loaves of bread have overall size and also length, height, and width; slices of bread have thickness, height, and width; crusts of bread have thickness.) At the very beginning of the child's experience with a word, he will have represented only one or a few such aspects of particular objects to which the adjective applies. Those included will be a reflection of the child's haphazard encounters with the word.

It may seem that the specification of the dimensions of variation within particular objects requires the very feature system that characterizes the adult lexicon. If so, why does the child not represent those features directly, rather than list exemplars? The reason is that the relevant spatial predicates characterizing the dimensions of variation may not have been differentiated from irrelevant spatial and nonspatial predicates. Suppose the child first learns *deep* and *shallow as* applying to ends of pools. If he can use the words correctly faced with novel swimming pools, not confusing depth with the length or width of the pool, then certainly he has the concept of depth of swimming pools. But he may not see the similarity between the way that the deep end of a swimming pool is deep and the way that bowls, holes, and puddles are deep. He may not know that *deep* can apply when there is no contrast between two parts of a single object, or that it does not require a liquid medium. Each of these, plus many other irrelevant features, may be part of his unanalyzed conception of the depth of pools.

Many linguists and psychologists have discussed the relation between the perceptual system and the structure of language (Bierwisch 1967; H. H. Clark 1973; Miller and Johnson-Laird 1976). As all of these writers have noted, the features underlying spatial adjectives (and also spatial prepositions, verbs of motion, and other lexical domains) correspond to perceptual predicates. This does not mean that such predicates are already available to the child as lexical organizers. Rather, the lexical features must be worked out relative to all of the words in any particular lexical domain. The representation of specific exemplars could provide the basis for abstraction of common features within the uses of a word as well as for the contrasts with other words in the domain.

The missing-feature-plus-haphazard-example theory differs from the missing-feature theory both in its description of immature lexical entries and in its account of the process of lexical growth. In the revised theory, immature lexical entries include privileges of occurrence that contain dimensional information, albeit in an unanalyzed format. The process of lexical growth has two components. (1) The discovery of semantic features: the child must discover what aspects of the conceptual system are relevant to the structure of the lexical domain. (2) The mapping of semantic features onto words: the child must work out how the semantic features apply to all of the words in the domain. Both of these processes go on together.

The two theories are alike in that before a feature becomes a lexical organizer, it is in fact missing from the lexical entry of any word, so lexical development does consist partially in the addition of features to words. Left out of the missing-feature theory is the process by which new lexical features are discovered.

Several considerations favor the missing-feature-plus-haphazard-example theory over the missing-feature theory. From the beginning, according to the revised theory, more is represented about dimension than simple spatial extent. Therefore, *short* never should mean *little.* Moreover, the inconsistencies Considine and I found are to be expected. Any probe of a word necessarily requires some particular object to which the child must try to apply it. In the early stages of that word's lexical entry, the child's response will be determined by the haphazard examples he has represented. Thus, to return to the *deep* and *shallow* example, the child might respond correctly if he happens to be probed about ends of swimming pools, and he might be able to give the fully appropriate opposite to each word, since they were learned as explicit contrasts. But if shown a picture of a deep puddle-a girl sinking into a mud puddle up to her knees - and asked "Is this a deep puddle?" the child might answer, "No, a big one." The responses of several children in our sample followed such a pattern.

The most compelling consideration in favor of the missing-feature-plus-haphazard-example theory is the very complexity of the feature system underlying the lexical domain of spatial adjectives. It is possible that even some adults do not discover all the regularities in the domain, never fully representing, for example, how *fat, wide,* and *thick* differ, although they know very well some paradigm cases of things that can be each.

The missing-feature-plus-haphazard-example theory relies on the contrast between semantic information that is part of an integrated representation of a particular object and semantic information as features that are lexical organizers. Just this distinction is necessary in the discussion of the acquisition of causative verbs.

Causative verbs In her analysis of her daughter's acquisition of causative verb constructions, Bowerman presents convincing evidence for there being a point in development when the feature [cause] has become a lexical organizer. Earlier uses of causative verbs involve notions of causation, but only as part of unanalyzed conceptual packages corresponding to the whole meaning of verbs (Bowerman .1974). Thus Bowerman's analysis directly parallels that offered above for spatial adjectives.

In English, causal relations between two propositions can be expressed in at least two ways. In the first, both propositions are reflected on the surface, as in *He made her steal the diamond.* In the second, the causal proposition is incorporated lexically into the verb, as in *He warmed the bottle* ("made it warm"), *He sharpened the knife* ("made it sharp"), and *He killed her* ("caused her to die"). Many noncausal transitive and intransitive verbs and many adjectives can be used as causatives. Some causative verbs, like *open* and *warm,* are lexically identical to their noncausative forms. Other causative verbs, like *enrich, sharpen,* and lay, exploit productive morphological rules that relate the two forms. Still others, like *kill* and *give,* are completely unrelated lexically to their noncausative counterparts.

Children use certain causative verbs before they are two years old, while they are still producing predominantly two-word utterances. Examples include *Mommy open, Open box,* and *Give cookie.* The question of concern to Bowerman is whether the child should be credited with the two-clause structure (CAUSE BOX (OPEN sox)) that underlies such an utterance in a grammar of adult English. If not, when is such an analysis appropriate for child speech? And when does the child have command of [cause] as a feature of verb meaning?

When she was just a little over two years old, Bowerman's daughter Christy began making many productive errors in which she used noncausative words in causal constructions. Some examples:

(1) How would you flat it?
(said while trying to flatten a piece of paper)

(2) Down your little knee.
(said while trying to push her sibling's knee down)

(3) She came it over there.
(said of a dog who brought something from one place to another)

(4) Don't eat her yet; she's smelly.
(said as her mother was about to feed the baby, who needed a diaper change)
(5) I' m singing him
(said while pulling a string, trying to make a toy sing)

Some of Christy's errors were purely lexical; she used a word causatively when English demands a different form-as in examples (1)-(4), where the correct forms would *be flatten, lower, brought,* and *feed.* Other errors actually violated English restrictions on the semantics of causatives, as in example (5) (in English there is no lexical item meaning "cause to sing"); in order to be used causatively, verbs and adjectives must describe states or changes of states (including location). These errors were frequent and persistent; by the age of four Christy had produced hundreds and was still making new ones. Bowerman argues that such errors are overgeneralizations of the relation between the causative and noncausative uses of words like *open* and that they demonstrate that Christy had command of [cause] as a feature of verb meaning.

The flowering of these productive errors exactly coincided with Christy's first productions of two-clause coordinate structures expressing causation. Just one week before the first lexical overgeneralization, a large number of sentences such as I *got her wet* and *I made it warm* first appeared. Other constructions appearing for the first time were forms like *Put the hat on,* which also requires the coordination of an action (putting) and a resulting change of state (change of location of the hat). Bowerman argues that Christy's mastery of the syntax of these coordinated structures was a prerequisite to her noticing the relation between forms like I *made it warm* and I *warmed it* and thus to her abstraction of [cause] as a lexical organizer.

Christy had been using causatives like I *warmed it* for many months before she showed any surface realization of coordinated structures or made any productive errors. With these two developments, some of her former causatives like *bring* dropped out in favor of incorrect forms like *Come it over here.* This case provides the classic features of the sudden appearance of rule-governed productive control over what had previously been unanalyzed isolated lexically-bound routines.

If Christy did not represent the [cause] feature in her lexical entries

for *give, bring,* and *open,* even though they were used causatively, what did they mean to her? Bowerman says that there is no doubt that when Christy used *give,* for example, she knew who was giving what to whom and who would end up having what. But, Bowerman argues, the meaning of *give* should be thought of as an unanalyzed whole, corresponding perhaps to the entire act of giving. It is a matter of discovery for the child, Bowerman concludes, what aspects of the act of giving are linguistically relevant to the syntax of causatives and to the lexical organization of causative verbs.

After the abstraction of [cause] as a lexical feature of the meaning of causative verbs, Christy spent a number of years mastering the English restrictions on causative constructions. Thus, here were three stages in her development: the representation of causative verbs as unanalyzed cognitive units, the abstraction of the lexical feature [cause], and finally the long-drawn-out process of working out the details of the semantics of causatives.

There are, of course, large differences between learning spatial adjectives and learning causative verbs. Spatial adjectives form a well-defined, limited lexical domain; causative verbs are an open-ended set. Learning causatives fundamentally involves syntactic development as well as semantic development; some of the syntax of comparative constructions is already mastered before any of the specialized spatial adjectives are learned. Indeed, were it not for the syntax of causatives, Bowerman's pinpointing of the moment of the isolation of [cause] as a lexical organizer would not have been possible. But in spite of the differences between the two cases, they both support a view of lexical growth in which an important part of the process is discovering which aspects of conceptual structure fill the role of semantic features.

Summary and Conclusions

The preschool child effortlessly meets the challenge of learning his first language - quite an accomplishment in the case of the lexicon. In the first part of this chapter, the time course of learning a single new word was assessed. A distinction emerged between fast mappings and slow, extended mappings. One, or a very few, experiences with a new word can suffice for the child to enter it into his mental lexicon and to represent some of its syntactic and semantic *features.*

The importance of the fast mapping was seen in terms of the rate of vocabulary growth and the weight of information contained in each lexical entry. Given a fast mapping, an entry can be completed slowly as the child encounters the word again and contrasts it with other words. In the chromium study this slow process was seen to include reorganizing the lexicon and reorganizing the conceptual domain.

Further evidence that some features of a word's meaning are represented early while others are worked out slowly was found in the domain of spatial adjectives. But it is not known whether the early entries for spatial adjectives result from a fast mapping-that is, from just one or a very few encounters with the new word. A study parallel to the chromium study might settle the question. In the case of spatial adjectives it was suggested that the initial lexical entry for each new word contains the abstract comparative core, including polarity, plus specific examples of contexts in which it is appropriate to use the word. The process of lexical growth involves discovering which features of those contexts serve as lexical organizers. Further evidence for this process was provided in Bowerman's analysis of causative verbs.

Consider again the information available to a child when he first hears a new word. All he has is his representations of the linguistic and nonlinguistic contexts. The importance of the linguistic context should be stressed; for example, it was suggested that the child's knowledge of the syntax of comparatives is one possible basis for the inclusion of the comparative core in his initial representation of a new spatial adjective. If the semantic features relevant to the structure of the lexical domain are not already known to the child, then remembering the specific nonlinguistic context in which the word was first acquired provides a basis for future uses of the word. Thus, the missing-feature-plus-haphazard-example theory characterizes the lexical entries that result from fast mappings for new words in relatively unstructured domains-that is, in domains for which the child has not yet discovered the relevant semantic features.

In this chapter I have gone much further than available data license. Nonetheless, several topics for research can be suggested. The process of learning a new word from context should vary as a function of the lexical organizers already available to the child. Learning big and little may be different from learning a specific spatial adjective like tall, because the linguistic relevance of the core

comparative structure may not even be realized before big and little are learned. And learning thick, for example, should involve different initial mappings if some of the semantic features of the domain have already been added to the lexical entries of other spatial adjectives.

Finally, the method used in the chromium study could possibly be adapted to assess when the child is aware of the constraints on possible lexical entries within particular domains. When does he know, for example, that there is no word in English that means "cause to sing" or "bigger than this one" or "the color of lilac flowers"? The child could be taught new words that correspond to gaps in the lexicon-gaps for words that are allowed by the constraints on word formation but which the child does not yet know, as in the chromium study, and gaps left because words that would fill them would violate English semantic constraints. If the child finds "words" of the latter kind much harder to learn, we would have evidence that he is sensitive to the constraints on word formation within that domain. Such research might strengthen the central assumption of this chapter, namely, that consideration of the process of learning a single new word can lead to novel insights about lexical development.