Chapter 4
Object Names and Other Common Nouns

The first words of a child of Harvard graduate students in the 1960s are much the same as those of a child learning French in the Paris of the 1920s or one currently learning Kaluli in the highlands of Papua New Guinea. Many early words refer to middle-sized objects—things that can move and be moved. These include names for specific people (*Mama, Dada*), animals (*dog, cat*), toys (*ball, block*), articles of clothing (*sock, shirt*), and other artifacts (*fork, chair*). There are names for substances (*juice, milk*), names for parts, typically body parts (*nose, foot*), modifiers (*hot, more*), words that refer to actions or changes (*up, allgone*), and routines that are linked to certain social interactions (*bye-bye, peek-a-boo*) (see Clark, 1993, for review). Soon afterward, verbs appear (*go, make*), as do prepositions (*in, on*) and more abstract terms (*kitchen, nap*).

This chapter addresses the question of why children's early vocabularies are the way they are. The initial focus is on object names, but it also discusses nominals in general, including those that refer to substances, actions, parts, and collections. The chapter concludes by shifting attention away from meaning and toward form, asking how children figure out which words in an utterance are the relevant ones to attend to. This problem is largely ignored in discussions of word learning, but it is a serious one, and its solution leads to some rather surprising conclusions about the sort of experience necessary for successful word learning.

Explaining the Words People Know

Learning a word involves mapping a form, such as the sound "dog," onto a meaning or concept, such as the concept of dogs. This perspective leads to three considerations underlying why children and adults know the words they do.

Access to the Form

People cannot learn words unless they are exposed to them. We can explain much of the character of children's vocabularies in terms of this banal fact, without positing any differences between the minds of children and adults. No matter how smart babies are, their first words are more likely to include *milk* and *spoon* than *weed* and *gene.* These are the words they hear.

Note, however, that there is more to figuring the accessibility of a form than simply determining its frequency in speech to children. It is not how often the adult says the word that matters; it is how often the child processes it. (This distinction is sometimes expressed as *input* versus *intake*). For instance, in English-speaking mothers' speech to one-year-olds, names for things are loudest and most likely to be in final position (Goldfield, 1993; Messer, 1981); such factors make it relatively easy for children to extract these words from the speech stream. In contrast, closed-class morphemes like *a* and *the* are very frequent, far more so than any particular object name, but they are harder for children to process.

Access to the Concept

To learn what a word means, one needs to possess the relevant concept. A two-year-old child of parents who are buying a house might often hear the word *mortgage,* but the word will not be learned because two-year-olds don't have the concept of mortgages.

In some cases, the relative ease with which words are acquired can be explained in terms of conceptual access. For instance, children learn *dog* before *animal* and *car* before *vehicle.* Not only are basic-level terms more frequent in the input (e.g., Brown, 1958a), but children also find it more natural to categorize a novel object as an instance of a basic-level kind than as an instance of a superordinate kind. As a result, children (and adults) find it easier to learn novel basic-level names than novel superordinates, even when the words are used equally often (e.g., Horton & Markman, 1980; see chapter 6).

Access to the Mapping

Form X and form Y could be equally accessible to the child, and so could concept X and concept Y, but word X might still be learned earlier than word Y. It might be easier for children to figure out that form X maps onto concept X than to figure out that form Y maps onto concept Y. To learn a word, after all, you not only need to hear the form and possess the relevant concept; you have to put the two together.

Lila Gleitman and Henry Gleitman (1997) suggest that mapping difficulties are one reason why mental-state verbs such as *thinking* take so long to learn. They report a study by Gillette et al. in which they exposed adults to videotaped mother-child interactions with the sound turned off; the adult heard beeps whenever the mothers used a noun or a verb. The subjects were asked to guess which English words the beeps corresponded to. Under these circumstances, adults found it fairly easy to recognize when the mother was using an object name like *chair* but quite difficult to figure out when she was using a mental-state verb like *thinking.* Since the words were presented an equal number of times, and since adults already have the concepts of chairs and thinking, it is likely that the relative difficulty has to do with establishing the mapping. Under the circumstances in which these words are typically used, it is easier to figure out that *chair* refers to chairs than it is to figure out that *thinking* refers to thinking.

With these three factors in mind, we can go back to children's first words. It has long been observed that names for objects have a special place in child language. This point is often overstated. Not all or even most of children's words are object names. (Typically, fewer than half of children's first 50 words are object names.) In fact, not all of children's *nouns* refer to objects; before their second birthday, children produce nouns referring to substances, parts, actions, and locations (e.g., L. Bloom, Tinker & Margulis, 1993; P. Bloom, 1990; Gordon, 1992; Nelson, Hampson & Shaw, 1993).

Nonetheless, object names really are special. They constitute a much larger proportion of children's early vocabularies than they do of the vocabularies of older children and adults (Brown, 1957; Macnamara, 1982; Pinker, 1984). This is true for every language that has been studied, including English, Italian, Japanese, Kaluli, Mandarin Chinese, Navajo, Turkish, and Tzeltal (see Gentner & Boroditsky, in press, for review). Korean has been argued to be an exception to this generalization. It is a verb-final language and allows for nominal ellipsis, both factors that lead to more emphasis on verbs in the input. Choi and Gopnik (1995) present evidence suggesting that the noun bias is not as prevalent in the vocabularies of young Korean children. This is a controversial claim, as other investigators *have* found a strong noun bias in children learning Korean (Au, Dapretto & Song, 1994) as well as in children learning Japanese, which is also verb-final with nominal ellipsis (Fernald & Morikawa, 1993). But in any case, as Gentner and Boroditsky (in press) point out, even in the Choi and Gopnik (1995) study there does exist a noun bias in children's first 50 words (44 percent nouns versus 31 percent verbs)—despite the fact that verbs are more salient than nouns in the speech that these children hear.

Children are biased to interpret new words as object names. John Macnamara (1972, p. 4) was one of the first to note that this is children's default assumption when learning a new word: "It is obvious that an infant has the capacity to distinguish from the rest of the physical environment an object which his mother draws to his attention and names. It seems clear too that in such circumstances he adopts the strategy of taking the word he hears as a name for the object as a whole rather than as a subset of its properties, or for its position, or weight, or worth, or anything else."

I've said before that children almost never make mapping errors, but those that do occur can be captured in terms of this bias. Macnamara (1972) describes a 17-month-old who would refer to the kitchen stove as *hot*. Assuming that she really thought *hot* was the name for the stove (and was not using the word appropriately to comment on its most interesting property, as an adult might), this could be because she heard someone talk about the stove as being hot and mistakenly viewed the term as naming not a property but the object itself.

Considerable experimental evidence supports the existence of an object bias. When shown an object (such as a rabbit) and given a name for it, children will assume that the word refers to that object and not to a part of the object (the tail), a property (white), the action that the object is doing (hopping), or the stuff that the object is made of (rabbit meat) (e.g., Baldwin, 1989; Dockrell & Campbell, 1986; Golinkoff, Mervis & Hirsh-Pasek, 1994; Hall, Waxman & Hurwitz, 1993; Landau, Smith & Jones, 1988; Markman & Hutchinson, 1984; Markman & Wachtel, 1988; Soja, Carey & Spelke, 1991; Waxman & Markow, 1995). This is often described as a finding about children, but it is actually a fact about people in general. When shown a novel object and given a word that refers to it, adults do exactly the same thing: we take the word as an object name.

What is the nature of this whole-object bias? There are many possibilities. It could be a conceptual bias: we might naturally see the world as containing objects, and so, when we hear new words and have to figure out what they refer to, objects are natural candidates. It could be a mapping bias, based on theory of mind: we might believe that other people, when they use words, typically wish to draw our attention to objects. Or it might be an assumption about language, about how words work. Under all of these accounts, the further question arises as to whether the relevant biases or expectations are learned, perhaps through experience with language, or innate, perhaps evolved to facilitate the process of word learning.

Before addressing these alternatives, it is first necessary to consider more precisely what we mean by *object*. This inquiry bears on a certain class of theories about the origin and nature of the object bias, and it also is relevant for the question of how children learn names for entities other than objects.

Objects

What is an object? Put like that, the question is vague; the word *object* is used in many ways in both colloquial English and within psychology and philosophy. A better question is: When we conclude that children are biased to treat words as object names—because, for instance, they assume that a new word that is used in the presence of a rabbit means *rabbit* and not *white, jumping*, or *tail*—exactly which notion of "object" are we appealing to? Which criteria are children using when choosing *rabbit* over all of these alternatives?

We can quickly reject some proposals. When Simone de Beauvoir says that a woman stands before a man "as an object," she is describing a lamentable state of affairs; it degrades a person to be viewed in this manner. This is plainly not the notion of object we are interested in here. Logicians sometimes use the word *object* to describe anything that one can quantify over and, under this reading, literally anything can be a single object, including, say, my shoe and the top half of the Eiffel Tower. This also won't do. Similarly, when Gottlob Frege (1892) says that every proper name has an associated object, this would include not only individual people but also places (*London*), events (*World War II*), corporations (*Burger King*), and groups (*Spice Girls*). Again, this might be a perfectly good way to use the word for other purposes, but it is not relevant for our purposes.

More promising notions come from perceptual psychology, where similar questions arise. The relationship between parts and wholes poses a particularly important problem. David Marr (1982, p. 270) discusses this and presents a skeptical solution: "Is a nose an object? Is a head one? Is it still one if it is attached to a body? What about a man on horseback? . . . There is really no answer to [these questions]—all these things can be an object if you want to think of them that way, or they can be part of a larger object."

What we are looking for is a notion of object that best comports with children's word learning biases, one consistent with the finding that names for such entities are easy to learn. From this perspective, we can answer Marr's questions: a man on horseback is not an object in our sense because no child could ever learn a word that refers to such an entity. A nose and a head are not objects either; they are parts of objects.

In my view, the notion of object that best corresponds to the findings from language acquisition has been elaborated by Elizabeth Spelke on the basis of her infant research (e.g., Spelke, 1994; Spelke, Phillips & Woodward, 1995). We can call these entities *Spelke-objects,* though most of the time, when there is no danger of confusion, I'll just call them *objects.* To anticipate an argument I make below, it is not an accident that a notion of object developed from studies of how babies see the world so elegantly captures the word-learning biases we find in children and adults. Instead, humans are naturally predisposed to see the world as composed of Spelke-objects—and this explains the object bias present in early word learning.

What is a Spelke-object? Such entities follow principles, the most central one being the *principle of cohesion.* To be an object is to be a connected and bounded region of matter that maintains its connectedness and boundaries when it is in motion. With objects of the right size, this suggests a crude test of objecthood: grab some portion of stuff and *pull;* all the stuff that comes with you belongs to the same object; the stuff that remains behind does not (Pinker, 1997). By this criteria, heads are not typically objects; if you tug on a person's head, the rest of their body follows. When a head is severed, however, it is an object. A man on horseback is two objects, not one, because the man can move and be moved independently from the horse and vice-versa.

The principle of cohesion is a claim about our understanding of what it is to be an object, not a claim about the perceptual cues necessary for online object parsing. You don't have to actually see cohesive and bounded movement to realize something is an object; it is enough that you can infer that there *could* be such movement. Hence adults can parse stationary scenes into distinct objects because the gaps between entities imply that they will not move together (as when we see two shoes next to each other but not touching), because we recognize entities we know from prior experience exist as independently moving objects (as when we see a man on a horse), or because Gestalt cues such as good continuity and sameness of color and texture suggest that different entities have the potential for independent movement (as when we see a shiny red sphere resting on a flat green surface). Such inferences from stationary scenes can be mistaken, however. The man on horseback might be a statue, for instance, carved from a single piece of marble, in which case it would be one object, not two.

Babies are sensitive to shared patterns of movement when reasoning about objects. Kellman and Spelke (1983) showed three-month-olds a wide screen with one stick poking out at the top and another stick poking out at the bottom. If the sticks move back and forth in tandem, undergoing common motion, three-month-olds assume that they are part of a single object—and are surprised if the screen is removed and they are not connected. And eight-month-olds, though perhaps not younger babies, show some ability to use Gestalt cues such as good continuation to parse stationary scenes into multiple objects even if the objects are touching (Needham & Baillargeon, 1997).

There is evidence for an early grasp of other object principles proposed by Spelke. Babies expect objects to follow a continuous pathway through space; they do not disappear from one point and reappear at another (the *principle of continuity*), and they know that objects do not pass through each other (the *principle of solidity*). One finding that demonstrates both principles is that when an object is placed immediately behind a screen and the screen rotates backward, four-month-olds expect the screen to hit the object and stop; when it goes through the space that should be occupied by the hidden object, they are surprised (Baillargeon, Spelke & Wasserman, 1985). This suggests that they expect the object to (1) continue to exist behind the screen and (2) be solid and stop the screen from dropping. A final principle—*the principle of contact*—applies only to inanimate Spelke-objects: babies expect inanimate objects to move if and only if they touch (e.g., Leslie, 1982), though they do not have this expectation about people (Spelke, Phillips & Woodward, 1995).

These principles (cohesion, continuity, solidity, and contact) are said by Spelke to constitute *core knowledge* on the part of children and are likely to be innate. Other facts about objects, such as the fact that they will fall if unsupported, do not seem to be present in young babies and are possibly learned through experience.

How are these principles related? Spelke (e.g., 1994; Carey & Spelke, 1994) treats them as similar in nature and equally important. Certain considerations, however, suggest that the principle of cohesion has a special status. Consider how we deal with violations of the other principles. Violations of the contact principle are frequent; they occur when wind blows papers off my desk, for instance. Even when we have no idea why the principle is violated (as would be the case if my computer was to suddenly slide forward without anyone touching it), we are surprised, but the violation doesn't make us doubt that we are dealing with objects. Science fiction involves frequent violations of the continuity and solidity principles, as when people and machines are instantaneously teleported from one place to another or when an alien has the power to walk through walls. But, again, such violations do not lead us to doubt whether these entities really are objects. When Captain Kirk is teleported onto a planet (violating the continuity principle), you still expect him to be cohesive, solid, and so on. If he subsequently walks into a wall, you expect him to stop, not go through it.

But now consider violations of the cohesion principle. Imagine seeing someone reach for what looks to be a solid pyramid resting on a table, grabbing the top, and lifting . . . and then just the top half rises, the base of the pyramid remains on the table. You would be surprised to see this, just as three-month-olds are (Spelke, Breinlinger, Jacobson & Phillips, 1993). But it is a different sort of surprise than you get with violations of the other principles. With the other violations, you are surprised because they violate expectations of what objects should do: objects should not disappear and reappear elsewhere, go through walls, or move spontaneously. But in the pyramid case, the surprise isn't because violating cohesion is a strange thing for an object to do. It is because the action reveals that, contrary to your expectation, it isn't a single object at all: it is two objects. The principles of solidity, continuity, and contact describe our understanding about *how objects should behave;* the principle of cohesion describes our understanding of *what it is to be an object.*

In fact, this distinction between the cohesion principle and the other principles is implicit in the design of the studies described above. A typical study involves showing babies an object and having it act in some unusual way, such as move by itself, go through a barrier, teleport, defy gravity, and so on. If the babies are surprised, as shown by their pattern of looking, this suggests that they have certain expectations about how objects should behave. But such studies make sense only if one assumes that the babies have some antecedent way of knowing that the object in question actually *is* an object. Babies do know this, and this is because they are first shown that the entity obeys the principle of cohesion: it moves as a single unit. This tells them that it is an object and sets the stage for the experimenter to explore how much they know about the properties that objects do and do not have.

Versions of the principle of cohesion have been around for a long time. Consider Aristotle's (330 B.C. / 1941, 1015b36–1016a9) proposal in his *Metaphysics* that a "continuous thing" "has by its own nature one movement and cannot have any other; and the movement is one when it is indivisible, and indivisible in time. Those things are continuous by their own nature which are not one merely by contact; *for if you put pieces of wood touching one another, you will not say that these are one piece of wood or one body or one continuum of any sort.*"

Aristotle raises an important consideration here, which is that our understanding of what an object is rests in part on our intuitions about its "nature" (see also Cohen, 1996). Suppose I firmly hold a block of wood between my thumb and fingers. Under an overly literal interpretation of the Spelke principles, the block and I constitute a single object: we are solid, bounded, and continuous and move as a single unit. But this is not, of course, the sort of interpretation that people would naturally make. This sort of cohesiveness is seen as *accidental.* It is a different sort of cohesiveness that holds between, say, my head and the rest of my body. Or consider a large desk. It might take a lot of force to move such a desk separately from the rug it is on—more force than it would take to break off a small branch from a tree. Nevertheless, we see the desk as an object and the branch as an object part (and we would do so even if we had never before seen a desk or a tree).

The point here, following Aristotle, is that when we parse the world into objects, we do not merely use a simple algorithm about what moves together and what does not; we use a more sophisticated understanding of whether this common motion is by "nature."

It might be that an understanding of cohesion becomes increasingly more abstract with age. A baby might start off with a simple idea: "Something is a single object if and only if it moves as a bounded and cohesive region." By the time babies are 12 months old, however, their understanding of Spelke-objects is sufficiently subtle that they do not have to actually witness this movement. They can infer it—as when they parse a stationary scene into a cup and a spoon—because they've previously seen these objects in independent motion. Further development would involve an appreciation of more subtle cases, as when something that cannot easily move independently, such as a block of wood that someone is holding, is nonetheless treated as an object.

The importance of cohesion is obvious when we think of what object knowledge is most likely for, which is to parse the world into ecologically relevant units. The other principles don't segment the world in the right way. The principles of solidity and continuity are too broad: they apply to all portions of solid substance. A patch of ground is solid and continuous, and so is the top half of a rabbit. And the principle of contact is too narrow; it applies only to middle-sized inanimate entities, such as sticks and stones, and not to people, birds, or rabbits. But the principle of cohesion is just right. Armed only with this principle, an animal would succeed in identifying entities that *move,* and this includes the most important entities of all—other animals. Initially parsing the world on the basis of cohesion establishes a foundation that underlies the application of the other object principles.

The Bias toward Object Names

The analysis so far motivates an explanation for the object bias in word learning, one originally proposed by Dedre Gentner (1982). It is that we are predisposed to view words as describing whole objects because we are predisposed to think about the world as containing whole objects. Once children know that words are used to refer, objects are the natural candidates for what they are referring to (see also Macnamara, 1982; Maratsos, 1991).

This is not to say that babies are insensitive to properties and entities other than objects. Babies are sensitive to size and color, to movement, and to numerosity. But the objects are seen as the important entities; everything else is secondary. In other words, babies see the world as adults do.

Some of the merits of this approach become clear when we look at the alternatives. One theory is that the whole object bias is primarily an assumption about *words.* Such an assumption might emerge from innate properties of language or thought (e.g., Markman, 1989; Waxman, 1994; Waxman & Markow, 1995; Woodward & Markman, 1997), in which case it would apply to children's first words and should be universal. Or it could be learned through experience with words (Choi & Gopnik, 1995; Golinkoff, Mervis & Hirsh-Pasek, 1994; Gopnik & Choi, 1995; Nelson, 1988), in which case it should apply somewhat later in development, after children have had sufficient experience with language, and it might not be universal.

Another possibility is that the whole-object bias has a syntactic basis. For instance, children know that count nouns correspond to kinds of individuals, and objects are salient individuals; so when children hear a count noun, they are prone to assume that it refers to an object kind (P. Bloom, 1994b). Under this view, the bias should apply only after children are able to identify count nouns and mass nouns in English utterances, which is long after they have started to learn and use words (e.g., Gordon, 1988; Soja, Carey & Spelke, 1991).

A final proposal is based on theory of mind. Children's object bias might emerge from a sensitivity to the referential intentions of others. Children might be biased to assume that when adults use words in certain contexts, they are intending to refer to objects (e.g., Tomasello & Akhtar, 1995). Under this view, the whole-object bias does not reflect how children think about the world, about words, or about nouns: it reflects how they think about the minds of other people.

When you look at an older child or adult, all these factors might apply. It might be that when four-year-olds see someone point to a strange animal and say, "This is a lemur," they see the object as a distinct and salient entity in the world *and* they know that a novel word is likely to be an object name *and* they know that count noun syntax is indicative of reference to an object category *and* they can figure out that the situation is typical of an adult intending to refer to an object.

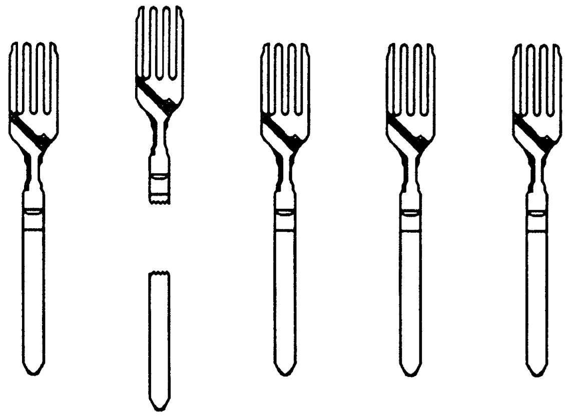

Figure 4.1
"Can you count the forks?" (from Shipley & Shepperson, 1990)

For four-year-olds, the whole object bias could be overdetermined. The debate is over which of these factors is primary.

Two facts about the whole-object bias strongly constrain how we can explain it. The first is that the whole-object bias is precocious. As noted above, object terms—either common nouns that refer to kinds of objects or proper names that refer to individual objects—are frequent among children's first words. Hearing a new word draws even 12-month-olds' attention to an object category (Waxman & Markow, 1995). This precludes certain accounts of where the bias comes from. For instance, it is unlikely to be the product of experience with words. It is also unlikely to exist through a sensitivity to the semantics of words belonging to the syntactic category of count nouns or lexical NPs, since 12-month-olds cannot yet distinguish members of these categories from other parts of speech.

The second fact is that the bias in favor of objects is not limited to word learning. I have already discussed a body of research showing that babies are biased to think of the world in terms of whole objects. Other research has found the same sort of bias in another domain, that of number. A particularly clever study was done by Elizabeth Shipley and Barbara Shepperson (1990). They showed preschool children pictures of objects, such as the picture in figure 4.1.

If you were shown this picture and asked to "Count the forks," you would say either "four" or "five" (depending on how open-minded you are about the broken fork). Shipley and Shepperson find that preschool children tend to answer "six." That is, even when told to count the *forks,* they count each of the *objects.* Similarly, when shown two red apples and three yellow bananas and asked to count the colors or the kinds of fruit, the preschooler's answer is dictated by the number of objects that are present, and so they tend to answer "five." It is not that these children don't know what forks, colors, or kinds are. It is that, in these circumstances, they focus on the objects.

What's the best way to make sense of this result? Stanislaus Dehaene (1997, p. 61) reviews this research, along with several other studies showing that babies can enumerate small sets of objects, and he draws the following conclusion: "the maxim 'Number is a property of sets of discrete physical objects' is deeply embedded in their brains."

This might well be true, but babies' numerical understanding extends beyond objects. After all, they can also enumerate sounds and actions (Starkey, Spelke & Gelman, 1990; Wynn, 1996). And young children have no problem counting entities such as sounds (Wynn, 1990). This suggests that the counting errors do not arise because children can count only objects. A more likely explanation is provided by Shipley and Shepperson (1990), who suggest that children have a strong bias to parse the world into discrete physical objects. As a result, when objects are present in the scene, children are strongly biased to count them. But in the absence of objects in the scene, children have no problem counting other entities.

It would be missing an important generalization to posit three independent object biases—one that underlies how babies track and individuate entities in the world (as found by Spelke), another guiding how they interpret the meanings of words (as found by Macnamara), and a third underlying their counting preferences (as found by Shipley and Shepperson). It makes more sense to use the first fact to explain the other two. Children think about the world as containing Spelke-objects. Hence, when figuring out what people are referring to when they use a new word, and when figuring out which entities to count, objects are natural candidates.

As a final example of the object bias at work, Geoff Hall (1996a) showed four-year-olds and adults entities that were presented in familiar geometrical forms. Some were Spelke-objects, such as a square made of wood; others were substances, such as a square made out of peanut butter. When simply asked "What is that?," subjects would talk about the objects with words that referred to the entire individual such as *square* (even though they knew the word *wood*), and the substances with substance names such as *peanut butter* (even though they knew the word *square*). That is, when shown a Spelke-object, children and adults are drawn to focus on it (and not the material that it is made from) even in a situation that does not involve the learning of words.

Simple Objects and Complex Objects

Gentner (1982) was one of the first proponents of the view that the whole-object bias has its origin in how children experience the world. More recently, in collaboration with Mutsumi Imai, she has proposed an intriguing modification to this view. Imai and Gentner (1997) suggest that all Spelke-objects are not created equal, at least not for the purpose of word learning. Instead, across different languages, complex objects are always perceived as distinct individuals, nonsolid substances are never thought of in this way, and simple objects fall in between. Depending on the language they are exposed to, children can come to think of simple objects as distinct nameable entities or as undifferentiated stuff.

The empirical basis for this claim is an experiment with American and Japanese children and adults, one that was a modified replication of Soja, Carey, and Spelke (1991). The subjects were shown substances (such as sand in an S-shape), simple objects (such as a kidney-shaped piece of paraffin), and complex objects (such as a wood whisk). These entities were named with neutral syntax, as in "Look at the dax" (and the equivalent neutral form in Japanese). Subjects were given a forced-choice task to see how they would generalize the words. They were shown another entity of the same shape made from a different material (consistent with an object interpretation) and another entity of a different shape but made from the same material (consistent with a substance interpretation) and were asked to "Point to the tray that also has the dax on it."

The subjects were young two-year-olds (mean age: two years, one month), older two-year-olds (mean age: two years, eight months), four-year-olds, and adults. The responses of the youngest children are the basis for Imai and Gentner's proposal. Replicating Soja, Carey, and Spelke (1991), they found that children learning English tended to extend the names for both simple objects and complex objects on the basis of shape but did not tend to do so for the names for substances. This is consistent with the view that the Spelke-principles underlie the whole-object bias. But the Japanese children behaved differently. Although they tended to generalize names for complex objects on the basis of shape and names for substances on the basis of material, they responded differently for the simple objects. Here they showed no generalization on the basis of shape; they responded at chance.

Imai and Gentner explain this difference between American and Japanese children as the consequence of the syntax of these languages. English has a grammatical count-mass distinction, in which count nouns (nouns that occur with determiners such as *another* and *many*) differ in their semantics from mass nouns (nouns that occur with determiners such as *much*) (P. Bloom, 1994b; Jackendoff, 1990). Spelke-objects are named by count nouns, as in "a dog" and "many bricks," while material entities that are not Spelke-objects are named with mass nouns, as in "some water" and "much sand." But Japanese has no grammatical count-mass distinction; the words *dog* and *water* fall into the same syntactic class. Imai and Gentner (1997, p. 193) propose that the syntax that children learn affects how they construe simple objects and suggest that the original object proposal of Gentner should be modified accordingly: "Gentner's (1982) natural-partitions hypothesis asserts that object names are learned earlier than relational terms because objects are . . . more easily individuated and parsed out from the perceptual context than other kinds of referents. Our results suggest adding the assumption of graded individuability: for example, that complex objects are more readily individuated (and thus mapped onto language) than simple objects."[1]

Sandeep Prasada (1999) presents a related proposal. He suggests that the psychologically relevant notion of object is strongly linked to intuitions as to how an entity is created (what Aristotle called its *formal cause*). Most important here is whether the entity is seen as having non-accidental structure. Under this view, a wood whisk is an object, but an irregularly shaped chunk of wood is not and is merely a portion of solid stuff. This raises an interesting prediction (1999, pp. 124–125): "A spatiotemporally contiguous amount of matter is distinguished from an object by the presence of the formal cause in the object. Given this understanding of the nature of things and stuff, we predict that subjects should be more likely to construe a solid entity as an object if they are presented with evidence that the entity has a definite form that is the product of a process directed at creating that structure and that it possesses functional properties that depend on that structure."

To test this, Prasada showed different entities to English-speaking adults and asked them if they would prefer to describe them as objects, with a count noun ("There is a blicket in the tray") or as substances, with a mass noun ("There is blicket in the tray"). He found that the object label was preferred when people were shown regularly shaped entities or entities that perform a structure-dependent function, and a substance label was preferred for jagged entities or entities that do not perform a structure-dependent function.

Such findings make a convincing case that a psychological difference exists between simple objects and complex objects. But I doubt that the reason for this difference is that children conceive of simple objects as clumps of solid substance and not as objects at all. After all, abundant evidence from infant research shows that babies individuate, count, and track simple objects, such as ping-pong balls, just as well as they do complex objects, such as Mickey Mouse dolls. There is every reason to believe that babies, like adults, think that a chunk of wood is just as much of an individual object as is a wooden whisk.

Consider a different way to capture the essence of the simple-object versus complex-object distinction, one suggested to me by Susan Carey. Count nouns, words such as *dog* and *whisk,* refer to psychologically interesting kinds of individuals. Their members share relevant and distinctive properties: once you know that an object is a whisk, for instance, you know that it was designed to fulfill a particular function, and this distinguishes it in a useful way from other objects. Chunks of wood, on the other hand, were not created with a specific intent, do not have a common function, and do not share internal properties. They are neither artifacts nor natural kinds, and so we are not predisposed to learn a count noun that refers to chunks of wood as members of a distinct kind.

Simple objects and complex objects, then, are all individuals. But complex objects are more likely to be thought of as members of distinct kinds, and therefore adults prefer to use a novel count noun to describe a complex object than to describe a simple one. All of the Aristotelian considerations raised by Prasada (involving form, structure, and nonrandom causation) still apply—but to the question of what makes a psychologically natural *kind* and not to what makes a psychologically natural individual (see chapter 6 for discussion).

What about the cross-linguistic differences? As discussed above, Imai and Gentner (1997) found that their youngest group of American and Japanese children differed only with regard to simple objects: the Americans generalized them on the basis of shape, and the Japanese performed randomly. But when you look at the other three age groups, you see a more general difference between the Americans and the Japanese. The results are shown in figure 4.2.

The American responses in the substance condition might seem crazy. Why would they generalize the substance name on the basis of shape about half of the time? The answer probably lies in the distinctive shapes in which these substances were originally presented. If you see an S-shaped array of sand, hear it given a name, and then see an S-shaped array of glass pieces, the temptation is to assume that the two go together. It couldn't be an accident, after all, that they have the same shape, so this might be the "right answer." (An even more killjoy argument is that the Americans thought that the new word actually meant "the letter S"—but the other substance stimuli did not correspond to English letters.) The Japanese subjects were less prone to do this.

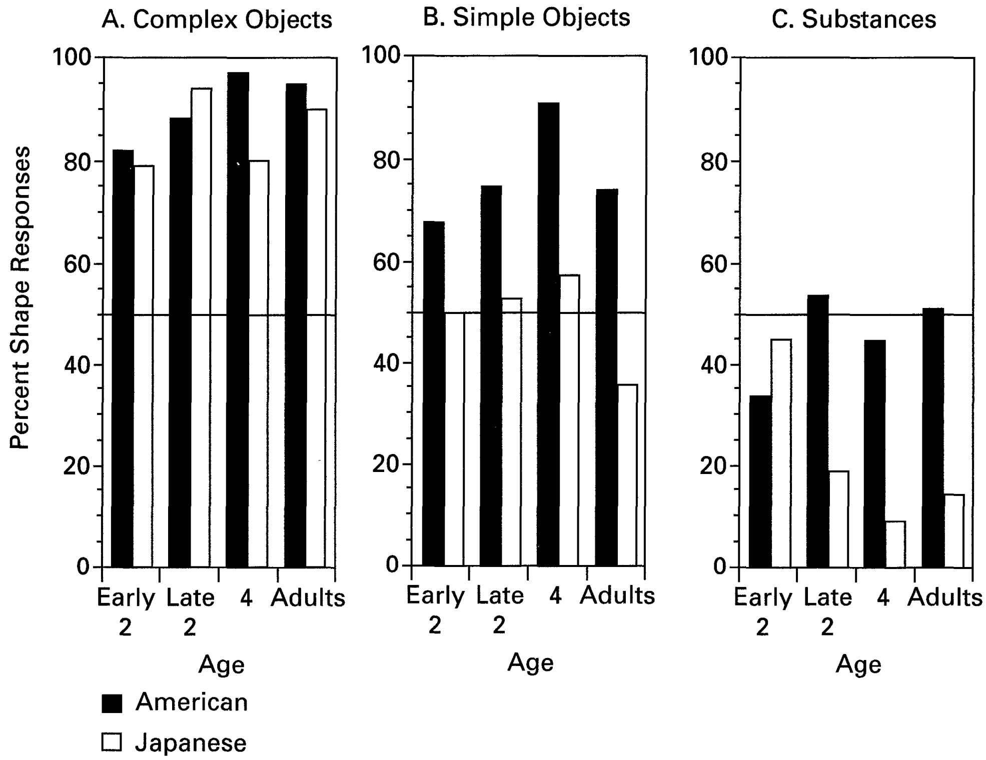

Figure 4.2
Generalization of novel words on the basis of shape by American and Japanese subjects (from Imai & Gentner, 1997)

They almost never generalized on the basis of shape, tending instead to extend the name for S-shaped sand array to the sand piles.

Whatever the precise explanation for this difference in shape generalization, it is clear that for the older two-year-olds, the four-year-olds, and the adults, an effect of equal magnitude exists in both the simple object condition *and the substance condition:* in both conditions, Americans focus more on shape than the Japanese. Such results are inconsistent with the claim that the two populations differ solely in how they think about simple objects. Instead, it may be that the Americans are more eager than the Japanese to construe words as referring to kinds of individuals, even if the words refer to nonsolid substances that are arrayed in a certain nonaccidental form, such as a S-shaped pile of sand.

In sum, the difference between simple objects and complex objects is an intriguing one, but there are reasons to doubt that it exists because only complex objects are thought of as distinct individuals. Instead, complex objects are more readily construed as members of distinct kinds than simple objects and hence are more naturally described with count nouns, explaining the findings of Prasada. And the American and Japanese difference found by Imai and Gentner is consistent with the theory that, for whatever reason, American subjects are more willing than Japanese subjects to assume that a new word (for a simple object or for a substance in a nonrandom shape) refers to a kind of individual.

Overcoming the Object Bias

Consider a child who sees a rabbit, hears the word *rabbit*, and has to figure out what the word refers to. Children who assume that words refer to whole objects can solve this problem: they infer that the word refers to the rabbit and not to its tail, its top half, and so on. The original motivation for constraints on word learning is to solve such problems of induction, and this approach works well for this example.

But the limits of the whole-object bias as a solution to Quine's problem are embarrassingly clear. A whole-object bias is wonderful for explaining how children learn words that refer to whole objects. It is less wonderful for explaining how they learn words that don't. Children learn substance names, part names, verbs, and adjectives, words such as *water*, *tail*, *hopping*, and *white*. What tells children that a new word is not an object name?

One consideration has to do with the nonlinguistic situation in which the word is presented. Children are predisposed to parse the world into objects, but if they have no available objects, other candidates rise to the top. This can be clearly seen in the domain of counting. As pointed out above, it is not particularly hard to get children to count entities that are not objects. The trick is to ask them to do so when there are no salient objects in sight. Under this circumstance, young children have no problem counting nonobject entities, such as sounds.

Similarly, it is not surprising that some of the earliest nonobject words learned are names for substances, such as *water* and *milk*. Such words are learned easily. In an experiment by Soja, Carey, and Spelke (1991), two-year-olds were presented with novel words either in neutral syntax ("This is my blicket"), count syntax ("This is a/another blicket"), or mass syntax ("This is some/some more blicket"). The words were used to refer to objects (e.g., a T-shaped plumbing fixture) or substances (e.g., Nivea cream). Soja et al. found that children extended words referring to solid objects to objects of the same shape, ignoring substance, but extended words referring to nonsolid substances to portions of the same substance, ignoring shape. For the youngest children, syntax had no effect on children's responses. Even before they have learned the syntax of count and mass nouns, children find it just as easy to learn a word like *water* as a word like *spoon.*

The ease with which children learn nonsolid substance names is in sharp contrast to the difficulty they have with solid-substance names such as *wood* and *metal* (Dickinson, 1988; Prasada, 1993). Such names are hard to learn for two related reasons. First, it is considerably easier to construe water as a substance category than it is to think of a solid entity such as wood that way (Bloom, 1994b). And second, children can learn *water* without the distraction of a salient object, but learning *wood* requires that children actively focus on a bounded object and think of it not just as an object but as a portion of solid stuff. This is analogous to a counting task in which they must attend to a set of objects and think of them not as objects but as forks, colors, or kinds (Shipley & Shepperson, 1990).

A second consideration relevant to learning words that are not object names is the pragmatics of the situation. The object bias will be suppressed if children are given good reason to believe that an adult is not intending to refer to an object. This will occur in cases of lexical contrast (discussed in detail in the last chapter). If an object already has a name, then another word used to label it is taken by children as likely to have some other meaning, such as referring to a part or property (e.g., Markman & Wachtel, 1988).

In some cases, speakers explicitly make clear their intent to refer to something other than an object. Consider the naming of body parts. Young children learn words such as *nose* and *eye* and typically do so even before they have a common noun such as *body* or *person,* but they never misinterpret these part names as referring to their whole body. Why not? The answer is that no parent has ever pointed at their child and said "Nose!" Instead, in at least some cultures, part names are presented with linguistic support ("This is *your nose*"), with the addition of other nonlinguistic cues (such as outlining the boundaries of the child's nose with a finger). When naming parts, then, adults take great pains to make it clear to children that they are not referring to the whole object, and apparently even one-year-olds are savvy enough to appreciate these pragmatic cues (see chapter 3).

Cues from discourse context can apply in more subtle ways as well. Tomasello and Akhtar (1995) presented two-year-olds with a new word in isolation—"Modi!"—that described a scene in which a novel object was engaged in a novel action. What children thought the word meant depended on which aspect of the scene, the object or the action, was new to the discourse situation. If, prior to hearing the word, children observed several different actions done on the same object, they tended to interpret *modi* as an action name, but if they had instead previously observed the same action done to several other objects, they tended to interpret it as an object name.

A third consideration is syntax. For instance, if children hear a word used as a verb, as in "he's glipping the table," they can readily infer that it does not refer to an object but instead to an action, and if they hear it used an adjective, as in "that is a glippy thing," they can infer that it refers to a property. Once children have some command of the syntax of their language, syntax can play an important role in leading them away from inappropriate object interpretations.

It is sometimes said that syntax is a weak cue to word meaning, at least early in development. For young children, the argument runs, the bias to treat words as object labels is stronger than their sensitivity to syntax, and hence children will make the object interpretation even in the face of conflicting grammatical cues (e.g., Waxman & Markow, 1995; Woodward & Markman, 1997). One study that is said to support this was done by Hall, Waxman, and Hurwitz (1993). They showed two-year-olds an unfamiliar object, such as glass tongs, and presented the children with a novel adjective, as in "That's a fep one." Children were then shown an object that was different in kind but shared a salient property with the target object, such as a glass napkin ring, and an object that belonged to the same kind but lacked the property, such as red plastic tongs, and they were asked, "Can you find another one that is fep?" Children tended to choose the object of the same kind, treating *fep* as an object name despite the fact that it was used as an adjective. Other studies find that when children are presented with mass nouns that are used to refer to whole objects, as in "This is some blicket," they interpret them as object names, not solid substance names (Dickinson, 1988; Markman & Wachtel, 1988).

These results are usually taken as showing that the object bias is stronger than syntax. But there is a another interpretation. All of the studies that show limits of syntax are based on the failures of younger children to use syntactic cues to learn a word as either a solid substance name or as an adjective denoting the substance an object is made of. Maybe their problem doesn't lie with syntax at all. It is instead that words that refer to solid substances are hard to learn, for reasons discussed earlier.

It is easy enough to test which interpretation is right. Just consider what children do in a situation where they have to use syntax to figure out that a word is not an object name but where the word also doesn't refer to a solid substance. If their problem really does have to do with syntax, then they should find it equally hard to learn adjectives such as *big* or verbs such as *hit.* But if their problem has to do with solid substance names, such words should pose no special problem. In fact, young children do learn adjectives and verbs, and plenty of experimental evidence (reviewed in chapter 8) shows that syntax helps them do so.

Individuals That Are Not Objects

Proper names and count nouns correspond to entities we think of as individuals—entities that can be individuated, counted, and tracked over space and time. Proper names such as *Fido* correspond to specific individuals, while count nouns such as *dog* correspond to kinds of individuals. It makes sense to talk of two dogs or 10 dogs, to say that a certain dog is the same one that I saw yesterday, or to ask what happened to Fido. Not all parts of speech refer to individuals; adjectives like *big,* verbs like *give,* and mass nouns like *water* do not correspond to entities that are thought of in this way.

What sorts of entities are naturally thought of as individuals? We individuate, count, and track dogs: they are psychologically natural individuals. As a result, words that refer to this kind (*dog*) and words that refer to individual members of this kind (*Fido*) are learned by children. But not every possible individual is acceptable from the standpoint of human psychology. For instance, construing the spatially discontinuous entity composed of my dog and his favorite bone as a single individual is conceptually unnatural. We could not easily learn a proper name (*Fidobone,* say) for this entity, nor can we easily track it over space and time. Why is it that Fido is a psychologically possible individual, but Fidobone is not?

It might seem that we have already solved this problem. Fido is a possible object because Fido is a Spelke-object; it satisfies the object principles. Fidobone does not; it does not satisfy the principle of cohesion. Because of this, children see the world as containing Fido and not Fidobone.

This is fine as far as it goes, but children also learn many names that refer to individuals that are not Spelke-objects. Two-year-olds know words for parts such as *finger* and *eye.* They know names for actions such as *sneeze, cough, laugh, kiss, smile, hug,* and *bite.* (Max understood *somersault* at about 18 months; when he was asked to "do a somersault," he would attempt the appropriate motion, and he would appropriately identify the somersaults of others.) They know words for periods of time, such as *minute* and *hour;* for "negative spaces" such as *hole;* for sounds, such as *sound* and *noise;* and for collections, such as *family, forest,* and *bunch.* So while being a Spelke-object may be a sufficient condition for being a nameable individual, it is plainly not a necessary one.

What about these nonobject individuals makes them acceptable candidates for being named? How is it that children parse the world into certain parts, actions, collections, and the like, and hence have these notions available as possible word meanings?

The Generalization Hypothesis

One intriguing possibility is that the principles underlying these individuals are related to those that apply to objects. It has often been argued that abstract language and thought involves the metaphorical extension of spatial notions (e.g., Jackendoff, 1990; Lakoff, 1987; Pinker, 1989, 1997). For instance, we talk about a shoe *in* a box (spatial), as well as a week *in* a semester (temporal), and a character *in* a play (more abstract). Perhaps humans have evolved patterns of thought for physical objects, including the Spelke-principles. These constrain our reasoning about nonmaterial entities, such as sounds, events, and collections. And certain other entities, such as parts, shadows, and negative spaces, may be construed as individuals because they satisfy a subset of the principles; they aren't objects, but they are close enough.

Consider object parts. Parts do not obey the principle of cohesion and do not exhibit independent motion; this is why they are parts and not whole objects. But cohesion may nevertheless be relevant to our understanding of what a natural part is because the cohesion principle involves two conditions—boundedness and connectedness (Spelke, Phillips & Woodward, 1995). A psychologically natural part, while not bounded, will nonetheless move as an internally connected region. Hence fingers are natural parts and so are toes, but it is profoundly unnatural to think of the ring finger and the kneecap as a single body part (a fingerknee) because fingers and knees are unconnected.

But connectedness isn't enough. A one-inch wide ribbon of skin running from the left hand, up the arm, over the shoulder, and ending up at the middle of the lower back is connected (and also conforms to the principles of solidity and continuity), but it is not naturally seen as a body part. Something more is required. It might be that a psychologically natural part is a part that could readily be turned into an object. This is related to the observation by perception psychologists that objects are parsed into natural parts through a sensitivity to discontinuities in surface contour (Hoffman & Richards, 1984; Leyton, 1992). A finger, for instance, is an excellent part because—unpleasant as it is to think about—it is seen as having a potential separateness from the rest of the body, in that it can be cleanly severed. The possibility of a certain amount of independent movement is also relevant; you can wiggle your fingers while the rest of your body stays still. (Though not all parts move independently; consider teeth in a mouth or the handle of a cup.) Parts are likely to fall into a continuum in terms of how natural they are. All languages have a word corresponding to *finger* because it is very natural to parse the body into fingers; other English body part names, such as *chest* and *shin*, correspond to parts that are less intuitively natural and so are not as frequent in other languages (see Andersen, 1978).

Negative spaces are a particularly fascinating domain of study (Casati & Varzi, 1994). We naturally parse the world into negative spaces (holes, cavities, gaps, tunnels, caves, and so on), and even three-year-olds have no problem identifying and counting holes (Giralt & Bloom, in press). A psychologically natural negative space is a mirror image of a Spelke-object: instead of being a connected portion of matter moving through empty space, it is a connected portion of empty space nestled into matter. And some negative spaces even move, as when an air bubble rises to the top of a swimming pool or a whirlpool makes its way across an ocean.

Karen Wynn has suggested to me that a temporal analog to the principle of cohesion is relevant for some sounds and actions. To be counted as a single sound, a noise must be bounded and connected in space and time; if one hears a beep from the right side and a simultaneous beep from the left side, this is two sounds, not one. And if one hears, from a single location, a beep, a pause, then another beep, this is again two sounds (Bregman, 1990). By the same token, motion is parsed into distinct action if separated either in space (two people each jumping at the same time) or time (one person jumping twice with a pause in between). When motion is connected in time and space, it is naturally thought of as a single action.

Consider finally collections, such as families, flocks, armies, bikinis, and bunches. A noun such as *flock* is not itself an object name; instead it refers to a collection that is composed of objects. Hence learning such a word requires that children somehow override the bias to focus on whole objects and instead focus on the group. Two-year-olds typically know some collective nouns, such as *family*, and evidence suggests that such nouns are easier to learn than superordinates such as *animal* and *furniture* (Callanan & Markman, 1982; Markman, Horton & McLanahan, 1980). My students and I have been interested in what makes some groups of objects, and not others, natural individuals (e.g., P. Bloom, 1994a, 1996b; Bloom & Kelemen, 1995; Bloom, Kelemen, Fountain & Courtney, 1995).

It is not trivial to teach someone a new collective noun. If you simply show children or adults a display of three groups of stationary objects (as in figure 4.3A) and describe the display with a plural count noun, as in "These are fendles," the tendency is to view *fendle* as an object

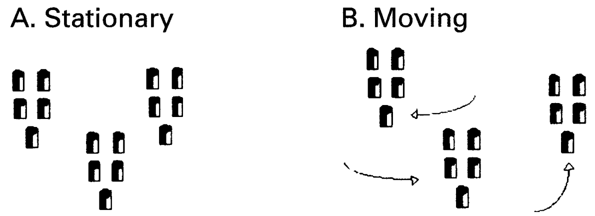

Figure 4.3
Stationary (A) and moving (B) collections

name, not as a collective noun. The display is seen as containing 15 fendles, not three fendles.

What can drive people to view such groups as individuals? One factor is motion. In one study, adults were shown three groups of five objects each on a computer screen, each object moving independently within its collection while also following the trajectory traced by the collection as a whole (see figure 4.3B). The groups behaved in an animate manner, like three swarms of bees, tracing paths along the screen and moving past each other. Under these circumstances, the groups are seen as individuals, and *fendle* is interpreted as a collective noun: subjects view the display as containing three fendles, not 15.

To test whether this effect is due to experience with real-world collections like flocks of birds or schools of fish, Karen Wynn, Wen-Chi Chiang, and I conducted a modified replication of this study with five-month-olds. We showed half the babies two collections of three objects each and the other half four collections of three objects each. Each collection traced a vertical path up and down on a computer screen. Once babies were habituated to this display, they were presented alternately with two collections of four objects and four collections of two objects, each moving horizontally back and forth on the screen. Babies looked reliably longer at the new number of collections, showing that they treated each of the collections as an individual for the purposes of enumeration (Wynn, Bloom & Chiang, under review).

Why does movement have this effect? It might make the collections objectlike, as if each collection were a surrogate object with unattached parts. This would be the inverse case of parts: parts satisfy connectedness but not boundedness; moving collections satisfy boundedness but not connectedness.

Limits of the Generalization Hypothesis
Unfortunately, the generalization hypothesis fails to account for many of the names we learn. Many refer to individuals that are *not* plausibly seen as objects, potential objects, or surrogate objects. They instead

refer to entities that emerge through our intuitions about the goals, intentions, and desires of others.

For instance, event nouns such as *conference, fight,* and *party* refer to individuals that are bounded on the basis of subtle intentional and social factors, not psychophysical ones. An ability to parse motionful scenes on the basis of such factors emerges quite early. Eighteen-month-olds who watch an adult attempt to perform an action and fail will often imitate the entire action that was intended, even though they never witnessed it (Meltzoff, 1995). Much younger babies—six-month-olds—can count the jumps of a continually moving puppet, one that wags back and forth between jumps (Wynn, 1996). This shows, at minimum, that they are not limited to individuating actions that are bounded by stillness. This ability to parse the scene into jumps might be the result of their sensitivity to systematic patterns of motion, or it might be the result of their construing of jumping, and not wagging, as an intentional act, which motivates them to extract it from the motionful scene.

Other nonobject individuals exist because they are seen as the objects of intentional regard by other people. Chapters, stories, and jokes are individuals just because they are created and thought of by others as singular entities. There is no independent motivation for treating them as such. A bikini can be viewed as a single individual because it is created and used for a singular purpose; and the world is divided into distinct countries (many of them, such as the United States, physically discontinuous) through social and historical factors, not solely physical ones. And while it is true that many parts can be identified on the basis of discontinuity of contour and the potential for independent movement, this might be because such factors are cues to the presence of deeper causal processes. In the case of body parts such as fingers, such processes have to do with growth and adaptation; in the case of artifact parts such as pedals, they have to do with design and function.

With this in mind, let's return to the question of what people see as a natural collection. As discussed above, if subjects are simply shown three groups of objects and given a name for them, the tendency is to treat the word as an object name (see figure 4.4A). But a collective interpretation can be induced if the subjects are convinced that each of the collections is thought of as a single individual in the mind of the experimenter. One way to do this is through syntax and discourse cues. If you point to each of the groups in turn, and say "This is a fendle, and this is a fendle, and this is a fendle," adults and five-year-olds tend to interpret *fendle* as a collective noun (Bloom & Kelemen,

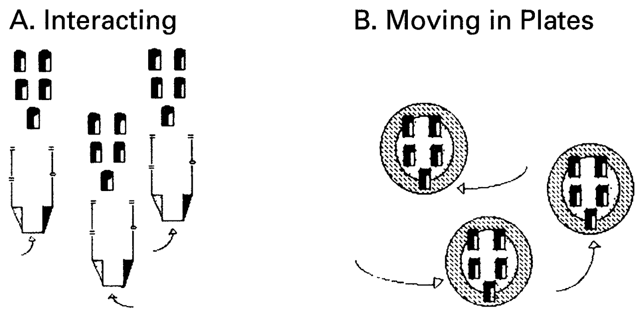

Figure 4.4
Machines interacting with collections (A) and moving collections on plates (B)

1995). But one can get the same effect in a more subtle way. In another study, an experimenter carefully arranged three groups of objects in front of the subjects, placing a picture frame around each group, giving the impression that each group was an independent artistic creation. The frames were then removed, and the entire display was described as "These are fendles." Under these circumstances, even four-year-olds tended to interpret the word as a collective noun (P. Bloom, 1996b).

In another study, subjects were shown a display in which three "machines" came into view, each surrounding one of the groups (figure 4.4A), with colored balloons popping out of each machine when it made contact. The machines then gradually retreated from the screen, leaving the screen empty except for the original, still stationary, groups. Subjects again tended to interpret *fendle* as a collective noun, either referring to each group (the target of an action) or to each group and its corresponding machine (all of the participants in an action). This supports the above finding that something can be seen as a collection if it is treated as a single individual by someone or something else (Bloom, Kelemen, Fountain & Courtney, 1995).

This brings us back to the question of why movement induces a collective interpretation. In another study with adults, each moving group was surrounded by a circle, and subjects were told "These are fendles on plates" to make it clear that we were defining the circles as plates (figure 4.4B). Placing the groups on "plates" explains the common movement of the objects within a group without appeal to any deeper shared properties of the objects. Without such a motivation to treat the groups as distinct causal entities, subjects abandoned the collective interpretation and treated *fendle* as an object name (15 fendles), despite the fact that the groups were moving. This suggests that, at least for adults, common movement is a cue to individuation only to the extent that it indicates some underlying causal property of the moving group. It is an open question whether this is also the case for children.

A different domain, studied by Nancy Soja and Tracy Burns (Soja, 1994; Burns & Soja, 1995) is that of *NP-type nouns.* Words such as *church* and *school* can refer to physical entities, as in "There is a church on the corner." But they can also be abstract proper names that do not refer to Spelke-objects, as when one says "John goes to church regularly," where the lack of a determiner indicates that the word is a noun phrase, or NP. Not all common nouns work that way; one can say "There is a bookstore on the corner," but it is unacceptable to say "John goes to bookstore regularly." Preschoolers use NP-type nouns correctly in their own speech and are sensitive to the subtle semantic criteria that determine whether a word falls into this category. If they are given a novel word that refers to an enduring cultural institution, something that people participate in at predictable times, children will treat the word as an NP-type noun, as a proper name for an abstract individual, but will not do so if these semantic criteria are not met. Such criteria are rooted in social considerations and have nothing to do with the Spelke-principles.

I am not denying here that once we think of something as a distinct individual, we tend to view it as objectlike in certain respects (e.g., Lakoff, 1987; Langacker, 1987). After all, we talk about individuals like jokes and chapters as if they are moveable sizable entities, as in "He shortened the joke and moved it out of the first chapter and into the second." But there is a big difference between claiming that once we think of an abstract entity as an individual, we then see it as having objectlike properties (something that is probably true) and the much stronger claim that we think of an abstract entity as an individual just *because* we see it as objectlike (something that is probably false in many instances).

The considerations reviewed above are consistent with the view that many of the candidate referents for common nouns and proper names emerge through one of two distinct cognitive systems. The first is an object system with an eye toward portions of matter that satisfy the Spelke-principles, particularly cohesion. This gives us dogs and bricks, whisks and chunks of wood. It might also provide us with—as a result of the extension of these principles—individuals such as fingers, toes, holes, shadows, and jumps. The second system is theory of mind, which parses motion and matter through an understanding of goal, function, and intent, giving rise to individuals such as games, parties, chapters, families, and church.

Finding the Right Words

Up to now we have viewed the problem of word learning as that of hearing a word and trying to discover its meaning. But another problem, raised by Landau and Gleitman (1985), is how children find the relevant words in the sentences they hear. Imagine a child who sees a dog, notices that the adult is also looking at the dog, and hears "Look at the dog." Suppose that the child can parse such an utterance into distinct words (see Jusczyk, 1997). How does she figure out which of the words, if any, correspond to the dog?

It is true that some Western parents are considerate enough to simply point and say "Dog!" But this is not a universal human behavior. Furthermore, some words are never used in isolation. Nobody ever points and says "The!" or "Of!" The problem of finding the right words has to be solvable by all children, for at least some of the words they learn.

As a starting point, we have to rethink certain assumptions about words and reference. The problem of word learning can be framed, as I have done up to now, in terms of how children figure out what words refer to, how they come to know that *dog* refers to dogs, *cup* refers to cups, *eating* refers to eating, and so on. This suffices for most circumstances, but it is not precisely right. Nouns do *not* refer to objects, and verbs do *not* refer to actions. Noun phrases refer to objects, and verb phrases refer to actions. As children attempt to make sense of what they hear, these NPs and VPs stand out as the referential elements and are used to refer to objects and actions in the world.

Consider how someone might draw your attention to a dog. She might say "Look at the dog" or "That's a dog" or "There's that big dog again." The part of the sentence that refers is not the noun *dog;* it is the NP *the dog* or *a dog* or *that big dog.* The noun contributes to the meaning of these phrases, but it does not, by itself, refer to any particular thing in the world. If common nouns refer at all, they refer to the kind or category of dogs (e.g., Macnamara, 1986). Pronouns and proper names are different; they stand alone to refer to specific individuals ("Look at him. Look at Fred") and hence are lexical NPs, not nouns.

In theories of formal semantics, nouns are viewed as predicates, which must combine with determiners to establish reference (e.g., Barwise & Cooper, 1981). More generally, it is often said that nouns are semantically incomplete (or "unsaturated"; see Higginbotham, 1983); only NPs are semantically complete. As such, only NPs can participate in certain forms of semantic interaction, such as being able to refer to entities in the world (as in "Look at the big dog"), having thematic roles (in the sentence "The big dog bit the ugly cat," the NPs, not the nouns *dog* and *cat,* are the agents and patients), and participating in coreference relations (in the sentence "The big dog likes himself," the NP "the big dog," not the noun *dog,* corefers with the reflexive *himself*).

The way children learn what nouns and verbs mean, then, is through their contribution to the meaning of the phrases in which they are used. It is easy to ignore this when discussing English because English noun and verb stems often appear in isolation. In other languages, such as Quechua, verb roots cannot stand as independent morphemes, and hence their meanings can *only* be determined through their contribution to the larger syntactic constituent that they belong to (Courtney, 1994; Lefebvre & Muysken, 1988).

Because of this, in an important sense the learning of nouns and verbs is similar to the learning of determiners. Children obviously don't learn the meaning of *the* by hearing this word used in isolation and figuring out what it refers to. They instead attend to NPs such as "the dog" and "the cups" and figure out the semantic role of the determiner across all of these phrases (Pinker, 1984, 1989). The suggestion here is that much the same holds for nouns and verbs.

This might seem somewhat baroque. Is it really true that to learn names for things (words like *dog* and *cup*) children have to figure out the contribution that these words make toward the meaning of the phrase, in the same way that they figure out how words such as *the* and *another* contribute? If so, why do children find it so much easier to learn the meanings of nouns than the meanings of determiners?

The thing to keep in mind here is that it is not difficult to find the noun inside the NP. In some cultures, parents might solve the problem for children by omitting the determiner: they might point to a dog and simply say "dog." Even if they are uttered, determiners are not phonologically salient, and so young children may not perceive them (Gleitman & Wanner, 1982). Parents might say "the dog," but children might hear just "Dog." Finally, children might have certain expectations about phonological and semantic differences between closed-class words and open-class words, allowing them to categorize these parts of speech (Gerken, Landau & Remez, 1989) and guiding them to expect that it is the open-class term that refers to the kind of entity, not the closed-class one (Pinker, 1984, 1989).

To see how this would work, imagine learning a new language. You are capable of parsing phrases into words and of figuring out what the phrases refer to, but you don't yet know what any words mean. You hear the following phrases in the following contexts:

Phrase	*Used to Refer to . . .*
za loob	a dog
te murpet	a chair
he murpet	two chairs
he loob	several dogs
wo murpet	a chair

If you don't hear the determiners *za, te, he,* and *wo,* there is no problem at all in finding the word that corresponds to the relevant entity or entities. Based on what you *perceive,* the mappings are straightforward:

Phrase	*Used to Refer to . . .*
loob	a dog
murpet	a chair
murpet	two chairs
loob	several dogs
murpet	a chair

But even if you do hear the determiners, it is not difficult to extract the nouns and infer that *loob* means dog and *murpet* means chair. This is because the most phonologically salient part of the phrase is the noun and the most semantically salient part of the context is the object.

These factors explain why it is harder to learn determiners than nouns. Not only are determiners less phonologically salient, but they also correspond to contrasts that are less semantically salient. The difference between dogs and chairs is more striking than the differences between one chair versus multiple chairs. Some of the semantic notions encoded by determiners are subtle indeed, requiring a sophisticated analysis of the scene and multiple exposures to narrow down the hypothesis space. What does *wo* mean? Based on the data above, it is impossible to tell. It could have a meaning akin to *the, this, that, a,* or *my,* to give only a few examples. To make matters worse, determiners often encode nonsemantic contrasts, as in the French contrast between *le* and *la.* These require a different sort of learning procedure altogether, one that is again based on multiple trials (see Levy, 1988).

Not all phrasal decompositions are equally easy. Finding the verb in a VP might be considerably harder than finding the noun in an NP. For one thing, a VP typically contains at least one NP, and hence many content words must be sifted through. Finding the verb might have to wait until the child either already knows the meanings of the nouns or has enough linguistic knowledge to use morphological and syntactic cues to parse the sentence. Furthermore, the referent of the verb might not "jump out" of the scene in the same way as the referent of a noun. For these reasons, among others, it should be harder to learn verbs than to learn nouns—and it is. We return to this issue in chapter 8.

Even if children have a procedure for extracting nouns from NPs, this might not entirely solve the problem posed by Landau and Gleitman (1985). Children still have to find the NPs. To see the problem this raises, imagine seeing a bird in the sky and hearing, in a language you wish to learn, "Zav bo goop wicket mep!"

By hypothesis, you know no words at this point, no inflections, and nothing about the syntax. (And we might be simplifying here by assuming that the sentence is parsed into strings of words; if not, the input is "Zavbogoopwicketmep!") So how can you determine which word or strings of words (if any) refer to the bird?

There are two possible solutions. The first is that learning a language requires a constrained distributional analysis; children store sentence-situation pairs in memory and then look for correlations between specific words or phrases and discrete aspects of meaning. For instance, if children later hear "blub mendle wicket mep" to refer to another bird, they might infer that "wicket mep" (which was present in both sentences) is used to refer to birds and is an NP. The analysis might be simplified if children store only strings of stressed phrases (instead of whole sentences) and representations of relevant individuals (instead of entire situations).

Such a distributional analysis might conceivably lead to problems. Imagine a child who sees one bird, hears the equivalent of "Isn't that pretty?," sees another bird, and hears the equivalent of "That's also pretty." Such a child, using the procedure above, would infer that "pretty" is an NP referring to the bird. Then again, some such confusions, such as using "hot" to refer to a stove, might actually occur in child language.

The alternative is that the input to children might be more congenial to word learning than we first assumed. The problem above might never arise. Perhaps all children are exposed to *some* NPs in isolation, and this can serve as a starting point for word learning. For instance, proper names have a special role in language acquisition. They appear among the very first words of children learning a range of different languages (Gentner, 1982), and they might be the one class of words that all children are guaranteed to have been exposed to in isolation or at least in some special stressed context.

Even in cultures in which adults do not usually label objects for children, proper names might be taught to children. Looking at Kaluli children, Bambi Schieffelin (1985, p. 534) notes:

> There are no labeling games to facilitate or encourage the learning of object names. This is primarily due to the linguistic ideology of the culture. It is only in families who are acquiring literacy that

> one sees any attention paid to saying the names of objects, and this activity is initiated by the child when the mother is looking at books. When extended by the child to other contexts, the mother's response is disinterest.
>
> In contrast, because of the cultural importance placed on learning the proper names and kinterms of the individuals with whom they interact, Kaluli children are consistently encouraged to master a large number of proper names, kinterms, and other relationship terms.

Perhaps every culture has some class of nominals that are special with regard to interaction with children. In Western societies, this is a very broad class, while for the Kaluli it is much more narrow, restricted to proper names and relationship terms. Another candidate for privileged nominals is the class of deictic pronouns, like *this* and *that*. These are also universal, show up early in child language, and draw children's attention to objects and other individuals in the environment.

In sum, it might be that all cultures will use some NPs that refer to individuals in isolation, allowing children to learn their first object names. If not, then children must be capable of somehow learning the meanings of words that are embedded in sentences, by extracting the referential NPs from such sentences through a constrained distributional analysis. An adequate theory of word learning must assume either strong extrinsic constraints (all cultures use some nominals in isolation) or a powerful learning mechanism (one that can learn words not presented in isolation).

Going beyond the Mapping Problem

Much of the previous chapter addressed the mapping problem: Given that children have access to both words and the sorts of things that words refer to, how do they bring them together? It was proposed that they use their theory of mind and figure out the referential intentions of other people. The present chapter asked where this prior access to meaning and form comes from. How do children parse the world into the right sorts of entities, and how do they parse the language to find the names for these entities? Candidate proposals again make reference to theory of mind, but the emphasis was on other cognitive mechanisms, such as a grasp of the object principles and the capacity to perform a distributional analysis on the linguistic input.

Solving the mapping problem is just the first step. It is one thing for a child to learn that a word is used to talk about a particular dog and quite another to know what the word means. The next two chapters address the question of how children figure out whether a word refers to an individual (as with *that* or *Fido*) or to a kind (as with *dog*) and how such words are extended to novel instances.

Note

1. It is not entirely clear what Imai and Gentner (1997) mean by *simple* and *complex* here. Elsewhere, Gentner and Boroditsky (in press) suggest that complex objects have "perceptual coherence," which entails a large number of internal links between the components (such as geons or parts) of an object. Complex objects are also said to possess "well-formed structure," which involves symmetry and regularity. But this gets confusing because Imai and Gentner (1997, p. 193) give *spheres* as an example of simple objects, even though spheres are perfectly symmetrical and regular and so should have well-formed structure and hence be complex. This is not a criticism of their study, since the contrast between simple objects versus complex objects is intuitively clear for the materials that they used (and this was confirmed by the ratings of naive adult subjects), but it is an area that would benefit from some clarification.

Chapter 5
Pronouns and Proper Names

A central question in cognitive psychology is how humans and other animals determine the category or kind a novel entity belongs to—how we categorize something as an apple or a table, a face or some water. And most research in word learning addresses how we learn names for these kinds—*apple, table, face,* and *water.*

But we also think about and name individuals. If someone tosses you an apple, it is not enough to know the kind it belongs to; you need to follow that specific apple, tracking its movement through space. Our emotions are tied to specific people and things. Original artwork and autographs can be worth fortunes, while perfect duplicates might be worthless. You might love your own newborn baby and be indifferent toward somebody else's—even if you are unable to tell them apart. In fact, without the ability to individuate, you couldn't tell the difference between one baby and two, except that two usually make more noise and take up more space. Although the understanding of individuals is much less studied than the understanding of kinds, it is every bit as central to our mental life.

The following three sections address how children learn names for individuals—pronouns and proper names—distinguishing them from common nouns that refer to kinds. The rest of the chapter addresses the broader question of the relationship between our understanding of individuals and our understanding of kinds.

Pronouns

Preliminaries

Pronouns belong to a class of linguistic expressions known as deictics or indexicals. These are words whose interpretation changes radically as a result of the contexts in which they are used. If you hear "Dogs like to chase cats," you can safely assume that the person is talking about the same types of entities (dogs and cats) and the same activities (liking and chasing) that anybody else would be talking about when