ch. 6

read 185-205
show down
187 - stages of a.
195-7 PS rules
Bot 6.2
197 - rules for st

Language in Mind

An Introduction to Psycholinguistics

JULIE SEDIVY *University of Calgary*

SINAUER ASSOCIATES, INC. • Publishers • Sunderland, Massachusetts • USA

6 Learning the Structure of Sentences

Achieving a vocabulary of 60,000 words or more is an impressive learning feat. But it's not nearly as impressive as the fact that you readily *combine* these words deftly and creatively. To get a quick feel for the scale of combinatorial possibilities that language offers, consider chemistry: with a measly 118 elements in the periodic table, there are *trillions* of known molecules that combine these elements. Just think what you can do with 60,000 units!

The combinatorial power of language allows you to convey entirely novel ideas that have never been expressed in language before. For instance, I'm guessing that you've never heard the following sentence:

> *It was all because of the lucrative but internationally reviled pink hoodie industry that the president came to abandon his campaign promise to ensure that every household parrot had recourse to free legal counsel.*

This sentence may be a touch on the enigmatic side, but chances are you had no trouble understanding it (though perhaps not all of its implications). On the other hand, you'd have no hope of understanding that sentence if its words were served up to you in this order:

> *Industry ensure because that internationally reviled had legal household parrot was it abandon all pink president every of campaign promise the but lucrative hoodie the came to his to that counsel recourse to free.*

Clearly, being able to string multiple linguistic units together is not enough. In order for us to be able to understand sentences made by combining words, they obviously can't just be tossed together in a bag. There has to be some underlying order or structure. That is, language has a **syntax**, a set of rules or constraints for how the units can be put together. The syntactic structure

syntax The structure of a sentence, specifying how the words are put together. Also refers to a set of rules or constraints for how linguistic elements can be put together.

semantics The meaning of a sentence; the system of rules for interpreting the meaning of a sentence based on its structure.

telegraphic speech Speech that preserves the correct order of words in sentences, but drops many of the small function words such as *the*, *did*, or *to*.

compositionality The concept that there are fixed rules for combining units of language in terms of their form that result in fixed meaning relationships between the words that are joined together.

of a sentence is obviously intimately tied to its linguistic meaning, or its **semantics**. You can combine the same assortment of words in a number of permissible ways with strikingly different results in meaning. As any small child knows, the justice systems of the world care very deeply about the difference between the sentences *Susie punched Billy* and *Billy punched Susie*. It's the tight coupling between the syntax of the language and its semantics that makes these distinctions possible.

As it turns out, syntactic structure is extraordinarily complex—so much so that the efforts of many linguists over numerous decades have not succeeded in exhaustively describing the patterns of even *one* language (English being the most heavily studied to date). It is one of the least-understood aspects of linguistics. And yet small children master the bulk of it in their first few years of life, before they are seen fit to learn even the most rudimentary aspects of arithmetic. Figuring out how they do this, and what they know about syntactic structure at any given stage, is fraught with theoretical and methodological challenges. It takes a sturdy soul to venture into this particular research domain.

The combinatorial nature of language becomes evident in children's speech at a fairly young age. Typically, they move beyond the single-word stage and combine two words at around 18 months of age, and by about 2 years they can speak in **telegraphic speech** (see Box 6.1), which preserves the correct order of words in sentences but drops many of the small function words such as *the*, *did*, or *to*. At this stage, "kid talk" sounds a bit like the compressed language used in telegrams or text messages: *Mommy go store now. Doggie no bite finger.* Six months later, by the age of two and a half, most children are speaking in full sentences, though still simple ones. By the time they enter kindergarten, they have mastered virtually all of their language's syntax.

What's most interesting is that at no point do kids seem to entertain the possibility that sentences can be made by throwing words together in just any order. Even when their output is restricted to two words, they never seem to violate the basic word order patterns of their language. For instance, English-speaking children always place the subject in front of the verb and the object after the verb; so even in the speech of toddlers, *Susie eat* means something different than *Eat Susie*.

But figuring out what's inside children's heads as they learn to combine words is no easy matter. In this chapter, we'll explore some ideas about what it is that children have to learn about their language's structure, what their syntactic knowledge might look like at various stages, and what it is that allows them to ultimately learn such an intricate and complicated system.

6.1 The Nature of Syntactic Knowledge

Compositionality

In introducing this chapter, I presented you with a tossed salad of words as a way to illustrate how impossible it is to understand words that are combined together without any structure. But you don't have to look at long and complicated sentences to get a sense of why language *has* to have some underlying structure in order for it to get off the ground. You just have to look at compound nouns.

In Chapter 5, we talked about how simple nouns can be combined into more complex compound nouns such as *houseboat* or *insurance policy*. It turns out that this is one of the very rare instances in language where units get combined in a *non-compositional* way. **Compositionality** is the notion that there are fixed rules for combining units of language in terms of their *form* that result in fixed *meaning* relationships between the words that are joined together. At heart, the

BOX 6.1
Stages of syntactic development

Children begin to combine words shortly after their first birthday, when they have about 50–60 words in their vocabulary. Their utterances gradually become more complex, and some regular patterns show up across children in their progression to more complex syntactic forms. One of the most detailed studies of language development was published in 1973 by Roger Brown, who based his analyses largely on data from weekly recordings of three children taken over a period of several years.

Brown found that children of the same age varied a good deal in terms of the syntactic elements they produced in their own speech. For example, by the age of 2 years and 2 months, a little girl named Eve was already producing various prepositions and complex words like *mommy's*, *walked*, and *swimming*, while at the same age, her peers Adam and Sarah were still eking out telegraphic speech consisting only of content words unembellished by grammatical morphemes.

A much better way than sheer age to predict a child's repertoire of grammatical markers is the measure of **mean length of utterance (MLU)**. This refers to the average number of morphemes in a child's utterances measured at a given point in time (i.e., at a specific age). Here are some examples of how MLU is computed:

Daddy's porridge allgone.
1 2 3 4 MLU = 4

My mommy holded the rabbits.
1 2 3 4 5 6 7 MLU = 7

Daddy went to the store.
1 2 3 4 5 6 MLU = 6

Based on his analyses, Brown noticed that function words and suffixes tended to emerge in a fairly consistent sequence across children—and that this sequence didn't necessarily match the frequency with which they were spoken by the children's parents. He identified five stages of children's early syntactic development defined by MLU. The table presents a summary of the approximate ages and inventory of grammatical morphemes at each of these stages.

Roger Brown's five stages of syntactic development

Stage	Age in months	Overall MLU	Morphemes present	Examples
I	15–30	1.75	Content words only	More juice.
				Birdy fly.
				Here book.
II	28–30	2.25	Present progressive	I falling.
			in, on	Dolly in.
			Plural -s	Eat apples.
III	36–42	2.75	Irregular past tense	Baby fell down.
			Possessive -s	Mommy's hat.
			Full form of "to be"	Is Daddy sad?
IV	40–46	3.5	Articles	This is the mommy.
			Regular past tense -ed	I holded it. You fixed it.
			Third person regular (present tense)	He likes me.
V	42–52+	4.0	Third person irregular (present tense)	She does it fast.
			Full form of "to be" as auxiliary verb	Was she swimming?
			Contracted "to be" as main verb	He's nice.
			Contracted "to be" as auxiliary verb	He's swimming.

From Brown 1973.

mean length of utterance (MLU) The average number of morphemes in a child's utterances at a given point in the child's development.

idea is similar to the notion of operations in simple arithmetic: once you learn what addition and subtraction *do*, you can, in principle, add any two (or more) numbers together, or subtract any number from another. You may never have seen this equation:

$$457{,}910{,}983.00475386 + 6{,}395{,}449{,}002.03 = x$$

and you can't possibly have memorized such an equation in the way you might have memorized the "fact" that 5 + 2 = 7. But you can easily compute it, just as you can compute any combination of numbers joined together by an arithmetic operator.

The same thing applies to simple sentences like *Susie punched Billy*. The operation of joining a subject (*Susie*) together with the phrase *punched Billy* yields a perfectly predictable meaning result. We know that Susie is the individual who initiated the action of punching, and Billy is the unfortunate individual at the receiving end of the punch. If we replace the word *Susie* with *Danny*, then *Danny* now stands in exactly the same role in the sentence that *Susie* used to; changing which word appears in the subject slot doesn't allow the possibility that the new occupant of that slot (*Danny*) now refers to the recipient of the action rather than its initiator. We can do the same kind of replacement with the object—slide *Fred* into the slot where *Billy* used to be, and now *Fred* is the recipient of the punching action. These examples seem simple and obvious, but this notion of *predictability of meaning from the way the parts are put together* is at the very heart of the combinatorial nature of language. And it works for excruciatingly complex sentences—there's no reason to believe that there isn't the same kind of tight, predictable relationship between the structure and meaning of more convoluted sentences.

But, interestingly, noun-noun compounds don't behave in this rigid, predictable fashion. For example, consider these words: *houseboat, housewife, house guest, housecoat, house arrest, house lust*. Despite the fact that the same *structural* relationship exists between the component words in all of these compounds, there isn't a uniform *semantic* relationship: a houseboat is a boat that *is also* a house, but a housewife is certainly not a wife that *is also* a house. While you could construe both a housewife and a house guest as *living in* a house (at least some of the time), this certainly can't apply to a housecoat, which is something you *use in* a house. And *house arrest* and *house lust* fit none of these.

For the sake of comparison, let's see what happens when you join words that do stand in a compositional relationship to one another, say, an adjective with a noun. Consider the very diverse phrases *red dog, corrupt executive, long book*, and *broken computer*. A red dog is a dog *that is* red. A corrupt executive is an executive *that is* corrupt. A long book is a book *that is* long, and so on. In fact, any time you join an adjective with a noun, the adjective serves the purpose of identifying a property that the noun possesses. For the more mathematically inclined, we could say that a phrase formed by grouping an adjective with a noun corresponds to a set of things in the world that is the intersection of the sets picked out by the adjective and the noun—in other words, *red dog* refers to the set of things in the world that are both red and dogs. Given that this rule holds for any combination of an adjective and noun, the process of joining these two kinds of words is fully compositional. (I'll leave aside for the moment the tricky cases like *small elephant*: does this refer to the set of things that are both small and an elephant?)

To see how useful compositionality can be, watch what happens when you come across a phrase involving a noun that you don't know, let's say *a red dax*. You may not know what a *dax* is, but you know that it's colored red. But what does *house dax* mean? Is it a dax for a house? A dax that lives in a house? A dax that is a house? You can't tell without more information. Even if you *do* know both of the words in a newly coined compound (such as *house book*), it can take a lot of guesswork to figure out the actual relationship between the two parts (for example, try to work out the meanings of the novel compounds in Table 6.1).

You can see from these simple examples how non-compositionality would seriously hinder the communicative usefulness of sticking words together into longer units. Essentially, you'd have to memorize the relationships between words, or at best infer them by running through some plausible possibilities. But you couldn't predictably *compute* them the way you can do a computation in arithmetic. Creating meanings non-compositionally gets squishy enough with combinations of just *two* words—now scale that up to sentences of 10 or 20 or 52 words (the last is probably the average length of a sentence in a Henry James novel) and you can see why compositionality plays such an important role in the expressive powers of language.

Despite the great allure of compositional meaning, the existence of noun-noun compounds points to the fact that it *is* possible to combine words in ways that don't reflect a general rule—in this case, although the combinations are syntactically uniform, always combining a noun with a another noun, they don't result in a fully general semantic pattern. At the same time, though, some regularities do exist, and there are some common semantic relations that tend to occur over and over again. For example, many noun-noun compounds express a part-whole relationship: *computer screen, car engine, door handle, shirt sleeve, chicken leg, wheel rim*, and so on. (Notice that the second noun is always a part of the first; a *chicken leg* is most definitely not the same thing as a *leg chicken*, whatever that means.) Another common semantic relation is one in which the second noun is a thing for the benefit of the first: *baby carriage, dog bed, cat toy, student center, employee insurance*. So, some generalization by analogy is possible, even if it is looser than more compositional kinds of meaning combinations.

One way to think about the distinction between compositional and non-compositional meanings is in terms of the words-versus-rules debate you met in the discussion of past tense and plural forms in Chapter 5. In that chapter, we saw that researchers came to intellectual blows over a controversy about whether complex words marked as regular plural or past-tense forms (such as *dogs* and *walked*) get formed as the result of a general rule that creates larger units out of smaller morphemes, or whether they arise by *memorizing* the complex words and then extending their internal structure by analogy to new examples.

Memorizing meanings and extending them by analogy is exactly what we have to do with the meanings of noun-noun compounds, since there doesn't seem to be one fixed rule that does the trick. So it's important to take seriously the possibility that other ways of combining words might also be achieved by means other than a rigid combinatorial rule over abstract categories.

In fact, some researchers would argue that what looks like rule-based behavior (for example, making combinations like *red ball*) is really just an extension of the memorize-and-extend-by-analogy strategy used in combining units like *coffee cup*. It's just that it's more regular. Others argue that it's extremely unlikely that adult speakers could manage to flaunt the complexity and creativity that they do without the benefit of something akin to rules. But even if we accept that adult speakers accomplish most of their word combinations by means of rules, this doesn't necessarily mean that combinations of units *at all ages* are accomplished by rules. This issue is an important one to keep in mind throughout the chapter.

It's clear, however, that the rule-based approach offers certain cognitive advantages by putting much less burden on the memorization of the complex words or phrases—we can simply apply the rules in a fully general way. This advantage becomes larger and larger as we scale up from two-morpheme words all the way to complex sentences. And, at the upper limits of our language, ab-

TABLE 6.1 Novel noun-noun combinations

Take a stab at describing a plausible meaning for each of these novel combinations of nouns when they are made into noun-noun compounds (and if more than one meaning comes to mind, provide several). What range of different semantic relationships between the two words do you find yourself drawing on?

rabbit phone	paper candy
flower card	rain paint
wallet plant	window cup
book dirt	computer organ

stract rules allow us to create an endless variety of new combinations that seem impossible to form by analogy to existing forms. What might such rules look like, if indeed they do exist? And what kinds of categories would they involve?

Basic properties of syntactic structure

One of the most basic things a set of language rules needs to do is radically constrain the possible combinations of words into *just those that are meaningful*. For example, given the bag of words *cat, mouse, devoured, the,* and *the,* we need to be able to distinguish between an interpretable sentence like *The cat devoured the mouse* and other combinations that don't yield a meaning at all, and are not considered to be possible combinations (marked by an asterisk as not legal or meaningful within a language):

*Cat mouse devoured the the.

*Cat the devoured mouse the.

*The cat the devoured mouse.

*The the cat devoured mouse.

And we'd like our rules to apply not just to this bag of words, but to other bags of words, such as *teacher, kid, loves, every,* and *that,* to yield meaningful combinations like:

Every teacher loves that kid.

and to rule out bad combinations such as:

*Every that teacher loves kid.

*Teacher kid loves every that.

*That every teacher kid loves.

Our rules need to be stated in terms of useful category labels as opposed to individual words like *teacher* or *cat*—otherwise, we'd never be able to generalize across sentences and we'd need to learn a new rule to go along with every new word. So, it becomes useful to identify words as belonging to particular syntactic categories. Let's propose some categories, and some useful abbreviations:

Det—determiner {*the, every, that*}

N—noun {*cat, mouse, teacher, kid*}

V—verb {*loves, devoured*}

The first line means that *Det* stands for the category determiner, which is a set that contains the words *the, every,* and *that*. Noun (N) is a set that contains *cat,* etc.; and V stands for the category verb. Using these categories as variables, we can now propose a rule that would limit the possible combinations of units, as specified by the following template:

Det-N-V-Det-N

This template allows us to indicate the correct way of stringing together these words and to rule out the meaningless ways.

But syntactic rules that merely specify templates for the correct word order wouldn't get us very far. Imagine you're a child learning about the syntactic possibilities of your language, and you've come up with the linear rule as stated above. Now, you encounter the following sentences:

She loves that kid.

Every kid loves her.

There's no way of fitting these new examples into the rule Det-N-V-Det-N. You might conclude that now you need another syntactic category that includes pronouns. And of course you'd have to specify where pronouns are allowed to occur in a sentence. So, you might add to your collection of possible rules:

Pro-V-Det-N

Det-N-V-Pro

where Pro stands for pronoun. And, if you encountered a sentence like *We love her*, you'd have to add a second pronoun:

Pro-V-Pro

So now you have a minimum of four rules to capture these simple sentences. But the fact that there are four separate rules misses something crucial: ultimately, structure is as systematic as it is because we want to be able to predict not just which groupings of words are legal, but also *what* the groupings *mean*. And, it turns out, the relationship between *she* and *loves that kid* is exactly the same as between *the teacher* and *loves that kid*. If there were separate rules that specified what Det + N could combine with and what Pro could combine with, in theory, they should be able to come with different specifications for the resulting meanings.

In other words, it might be possible for *The teacher loves that kid* to mean that the person doing the loving is the teacher, while *She loves that kid* might mean that whoever *she* refers to is the *recipient* of the kid's love. But this sort of thing never happens in languages. And what's more, if you looked at all the places where pronouns are allowed to occur, you'd find that they happen to be the exact same slots where Det + N are allowed to occur. What's needed is some way to show that the syntax of English treats Det + N and Pro as equivalent somehow—in other words, that Det + N can be grouped into a higher-order category, of which Pro happens to be a member. Let's call this a **noun phrase**, or **NP**.

As soon as we catch on to the notion that words can be clumped together into larger units, called **constituents**, our syntactic system becomes extremely powerful. Not only can we characterize the patterns of structure and meaning in the examples above, but we can explain many aspects of syntactic structure that would otherwise be completely mysterious. For instance, the syntax of English allows certain phrases to be shuffled around in a sentence while preserving essentially the same meaning. Some examples:

Wanda gave an obscure book on the history of phrenology to Tariq.

Wanda gave to Tariq an obscure book on the history of phrenology.

An obscure book on the history of phrenology is what Wanda gave to Tariq.

Here the phrase *an obscure book on the history of phrenology* acts as a single clump. It's impossible to break up this clump and move only a portion of it around; the following just won't work:

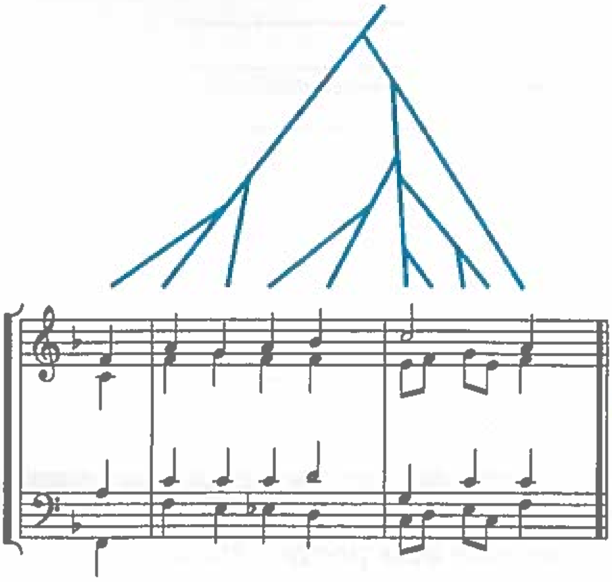
*Wanda gave an obscure to Tariq book on the history of phrenology.

*An obscure book is what Wanda gave to Tariq on the history of phrenology.

The reason these don't work is that the unit that's allowed to be moved corresponds to a higher-order constituent, an NP. The entire phrase *an obscure book on the history of phrenology* is an NP, so it has to move as a unit. (Notice the NP slot, indicated by brackets, could easily be filled by a much simpler phrase, as in *Wanda gave to Tariq [a puppy] / Wanda gave [a puppy] to Tariq*. Again, this illus-

noun phrase (NP) An abstract, higher-order syntactic category that can consist of a single word or of many words, but in which the main syntactic element is a noun, pronoun, or proper name.

constituent A syntactic category consisting of a word or (more often) a group of words (e.g., noun phrase, prepositional phrase) that clump together and function as a single unit within a sentence.



Christus, der ist mein Leben, 1st phrase (J. S. Bach)

Figure 6.1 We intuitively group words into phrases, or constituents. It's been argued that music is structured in a similar way, grouping notes together into constituents based on perceived relationships between pitches and chords. One effect of this is that pausing in the middle of a musical phrase or constituent makes a sequence of notes feel "unfinished" or unresolved. The figure shows an analysis of a musical phrase into its constituent parts. (From Lerdahl, 2005.)

trates the equivalence of NPs in the syntax, regardless of whether they're instantiated as very simple or very complex phrases.) This movement pattern can easily be captured with a notion of constituent structure—we could specify a rule that allows NPs to be moved around to various syntactic positions. But the pattern would seem rather arbitrary if our syntactic knowledge merely specified the linear order of individual words, because such a rule would give no indication of *why* the string of words should be cut exactly where it is. (You might be interested to know that music, like language, is often described as being structured in terms of constituents, as illustrated in **Figure 6.1**.)

One of the best arguments for the idea that sentences have internal structure rather than just linear ordering comes from the fact that the same string of words can sometimes have more than one meaning. Consider the following famous joke, attributed to Groucho Marx:

> Last night I shot an elephant in my pajamas. What he was doing in my pajamas, I'll never know.

Not a knee-slapper, perhaps, but the joke gets its humor from the unexpectedness of finding that the most natural way to interpret *I shot an elephant in my pajamas* turns out to be wrong, creating a jarring incongruity (incongruity being an essential ingredient of many jokes). On a first reading, most people don't group together *an elephant in my pajamas* as an NP unit, for the simple reason that this reading seems nonsensical. So they assume that the phrase *in my pajamas* is separate from the NP *an elephant*. That is, they assign the grouping:

> Last night I shot [$_{NP}$ an elephant] in my pajamas.

rather than the grouping:

> Last night I shot [$_{NP}$ an elephant in my pajamas].

But the joke ultimately requires you to go back and re-structure the sentence in a different (and very odd) way.

The rapid-fire decisions that people make about how to interpret ambiguous structures (see **Table 6.2**) is a fascinating topic that we'll explore in Chapter 8. But for the time being, notice how useful the idea of constituent structure can be. If sentences were defined only by word-order templates rather than being specified in terms of their underlying structure, examples like the Groucho Marx joke would seriously undermine the possibility of there being a systematic relationship between structure and meaning—the same word-order template would need to somehow be consistent with multiple meanings.

In fact, assigning two possible structures to the string *I shot an elephant in my pajamas* also explains an interesting fact, namely, that one of the possible meanings of the sentence evaporates if you do this:

> In my pajamas, I shot an elephant.

Now, it's impossible for the elephant to be sporting the pajamas; it can only be the speaker who's wearing them. This makes absolute sense if you know about constituent struc-

WEB ACTIVITY 6.1

Constituent structure in music
In this activity, you'll explore some parallels between the structure of sentences and the internal structure of music.

TABLE 6.2 Syntactic ambiguities

Can you identify at least two meanings associated with each sentence? Some meanings are clearly more plausible than others.
The children are ready to eat.
You should try dating older women or men.
He offered the dog meat.
What this company needs is more intelligent managers.
Jonathan has given up the mistress he was seeing for three years, to the great dismay of his wife.
Now you can enjoy a gourmet meal in your sweatpants.
Why did Joanie buy the frumpy housewife's dress?

LANGUAGE AT LARGE 6.1

Constituent structure and poetic effect

I've shown that grouping words into larger constituents can account for the equivalence of units like *the elephant in my pajamas* and *she* in the larger frame of the sentence. I've also argued that without constituent structure, we have no way of explaining why a string of words can have two quite different meanings. But there's additional evidence that we chunk words into constituents.

When you utter a long sentence, notice where you're most likely to take a breath or pause slightly. No one, for example, can read the following sentence from a Henry James novel (stripped here of its original commas) without slight breaks:

> He received three days after this a communication from America in the form of a scrap of blue paper folded and gummed not reaching him through his bankers but delivered at his hotel by a small boy in uniform who under instructions from the concierge approached him as he slowly paced the little court.

The slashes show where you're most likely to insert brief pauses:

> He received three days after this / a communication from America / in the form of a scrap of blue paper folded and gummed / not reaching him through his bankers / but delivered at his hotel by a small boy in uniform / who under instructions from the concierge / approached him / as he slowly paced the little court.

Notice that these breaks line up with boundaries between separate clauses or large phrases. As you'll see in this chapter, large chunks, or *constituents*, are in turn made up of smaller chunks. Pauses are most natural between some of the largest constituents in a sentence. The deeper down into the structure you go, the less likely there are to be breaks between words. It would be distinctly odd, for instance, to break up the first few phrases like this:

> He received three days after / this a communication from / America in the form of a /

Happily, Henry James made generous use of commas to help his readers group words into the right constituents.*

In poetry, although commas may be used, line breaks are often used to even more strongly set off phrases from one another, as in this stanza from Wilfred Owen's poem "Dulce et Decorum Est":

> *Bent double, like old beggars under sacks,*
> *Knock-kneed, coughing like hags, we cursed through sludge,*
> *Till on the haunting flares we turned our backs*
> *And towards our distant rest began to trudge.*
> *Men marched asleep. Many had lost their boots*
> *But limped on, blood-shod. All went lame; all blind;*
> *Drunk with fatigue; deaf even to the hoots*
> *Of tired, outstripped Five-Nines that dropped behind.*

* In case you're curious, with James's punctuation in place the sentence reads as follows: "He received three days after this a communication from America, in the form of a scrap of blue paper folded and gummed, not reaching him through his bankers, but delivered at his hotel by a small boy in uniform, who, under instructions from the concierge, approached him as he slowly paced the little court." (Henry James, *The Ambassadors*, 1903, p. 182.)

Continued on next page

LANGUAGE AT LARGE 6.1 *(continued)*

With the exception of the second-to-last line, all of Owen's line breaks segment separate clauses or sentences, carving the stanza at its most natural joints.

But poets can skillfully leverage the expectations their readers have about the most natural carving joints in language by creatively violating these expectations (you know what you've always heard about what you can do with the rules once you know them). The technique of breaking up a line inside of a poem's constituent is called *enjambment*. Sometimes, it has the effect of forcing the reader to jump quickly to the next line in order to complete the constituent. The poet e. e. cummings uses enjambment (often even splitting up words) to enhance the manic feel of his poem "in Just-":

in Just-
spring   when the world is mud-
luscious the little
lame balloonman

whistles   far   and wee

and eddieandbill come
running from marbles and
piracies and it's
spring

In this next poem—a work by William Carlos Williams titled "Poem"—constituents are broken up in a way that creates an effect of freeze-framing the deliberate motions of a cat:

As the cat
climbed over
the top of

the jamcloset
first the right
forefoot

carefully
then the hind
stepped down
into the pit of
the empty
flowerpot.

In fact, the line breaks and the effect they create are the main reason we think of this piece as a poem in the first place.

Finally, in "We Real Cool," Gwendolyn Brooks, by busting up the sentences in an unexpected way, has given extra prominence to the pronoun *we*, lending it a swagger it would never have if each instance of the subject pronoun occurred on the same line as its verb phrase:

The Pool Players.
Seven at the Golden Shovel.

We real cool. We
Left School. We

Lurk Late. We
Strike Straight. We

Sing sin. We
Thin gin. We

Jazz June. We
Die soon.

ture and you also know that the rules that shuffle phrases around are stated in terms of higher-order constituents. Remember that under the pajama-wearing elephant reading, *an elephant in my pajamas* makes up an NP unit, so it can't be split up (much as in the examples earlier about Wanda and what she gave to Tariq). Splitting apart *an elephant* from *in my pajamas* is only possible if these two phrases form separate constituents, as is the case in the more sensible reading of that sentence.

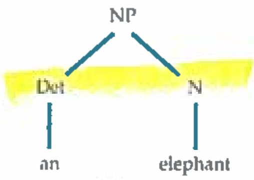
phrase structure rules Rules that provide a set of instructions about how individual words can be clumped into higher-order categories and how these categories are combined to create well-formed sentences.

Phrase structure rules

A useful way of capturing the structured nature of sentences is by means of **phrase structure rules** that provide a set of instructions on how individual words can be clumped into higher-order categories and how these are combined together to create well-formed sentences. For example, to capture the fact that a determiner and a noun can be combined to form a noun phrase, we might write:

NP → Det + N

We can apply this rule to give the following structure to the phrase *an elephant*:



But we also know that a pronoun on its own can correspond to an NP—this is apparent because pronouns and Det + N phrases sit in the same syntactic slots and contribute to a sentence's meaning in the same way:

An elephant sneezed.

It sneezed.

So we also need a phrase structure rule to reflect this structure:

NP → Pro

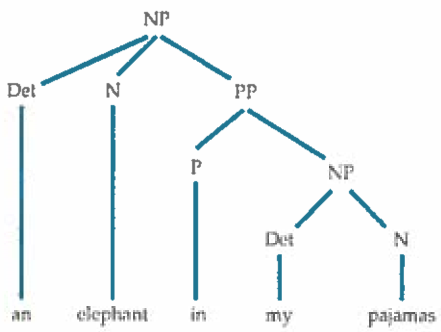
which gives the pronoun *it* this structure:

We can make exactly the same observation about proper names, which also act as NPs (for example, *Jonathan sneezed*). So we now have a third way of building an NP:

NP → name

Now, we've also seen that *an elephant in my pajamas* can be a complex constituent. In fact, it's also an NP, as evidenced by the role it plays in *An elephant in my pajamas sneezed*. This phrase is a bit more complex, involving a **prepositional phrase (PP)** embedded inside the noun phrase (NP). And notice that there's an NP embedded within the PP. So, we get a structure like this (where P stands for the category "preposition"):

prepositional phrase (PP) A syntactic constituent, or higher-order category, that in English, consists of a preposition (e.g., in, under, before) followed by an NP.



BOX 6.2
Rules for constructing sentences of English

Consider a possible set of rules for capturing the sentences we've discussed so far (parentheses indicate optional elements):

1. S → NP + VP
2. VP → V + (NP) + (PP)
3. NP → Pro
4. NP → Name
5. NP → (Det) + N + (PP)
6. PP → P + NP

It takes a bit of practice to get an intuitive feel for how such rules might work. For the following sentences, try using the above rules to create tree structures in which each node of the tree shows the result of applying one of the syntactic rules. If a sentence can have more than one meaning, be sure to show both possible structures, as permitted by the syntactic rules:

The girl with a tattoo loves me.
Bill snores.
Samantha read the book on the Queen's throne.
A life with no love awaits Jenny in the future.
I saw the man with the glasses from my store.

Is that all there is to English syntax? Hardly. If you take note of sentences all around you, you'll quickly run into examples that can't be captured by our current set of rules. Here are a few:

I love the girl who has a tattoo.
Alisha told me that she loved her husband.
Weyman and Stuart drank beer for three hours.
The three little pigs built their houses on the hill.
You believed the outlandish story that she fed you.
Fabian threatened to kill his rival.

You'd need to propose additional rules to build these sentences. Give it a try: suggest a specific set of rules that would produce these sentences, and provide the tree structure that would result. Don't worry about how you would label the syntactic units (any sensible label will do), but do think about which clumps would form a constituent.

Can you think of a few other examples of sentences that go beyond the scope of the rules in this box? (Note: this shouldn't be hard to do!)

patterns of our language, knowledge that might actually be implemented in a very different way. But for the most part, people tend to agree on a number of important points, including the following:

- Our knowledge of structure is **generative**; that is, whatever we know about language structure allows us to recognize and generate new examples of never-before-encountered sentences.
- Our knowledge of language structure must be **hierarchical**—that is, it needs to reflect the fact that words group together into constituents, which in turn can group together with other words or constituents to form larger constituents.
- The generative and hierarchical qualities of language allow for **recursion**, permitting syntactic embeddings that resemble loops that (in theory) could go on forever (see Box 6.3).

We've seen, for example, that a PP can be embedded within an NP—but the PP also can contain an NP. This creates the possibility of a structure that ends up looking a lot like nested Russian dolls, with multiple NPs (each enclosed in brackets in these examples) hiding within embedded PPs:

[The paper on [the desk in [the office]]]

generative With respect to language, a quality of language that allows us to use whatever we know about language structure to recognize and generate new examples of never-before-encountered sentences.

hierarchical Top-down (or bottom-up) arrangement of categories. With respect to language, a quality that involves how words group together into constituents, which in turn can group together with other words or constituents to form ever-larger constituents.

recursion Repeated iterations. With respect to language, refers to syntactic embeddings that nest constituents (such as clauses or NPs) within other constituents in a potentially infinite manner.

To get a structure like the above, we need the following rules:

(a) NP → Det + N + PP

(b) PP → P + NP

(c) NP → Det + N

We can also collapse (a) and (c) into a single rule schema, using parentheses to show that the PP is optional:

NP → Det + N + (PP)

Notice that we now have multiple ways of instantiating an NP. We've seen that any one of the possible expansions for an NP could occur interchangeably as potential subjects of a verb like *sneezed*. The same modular approach is available for any NP slot, regardless of where in the sentence it occurs. For example, consider the second NP in our tree structure, under the PP. Instead of expanding the NP node into Det + N, we could have expanded it as Pro:

[$_{NP}$ An elephant [$_{PP}$ in [$_{NP}$ it]]]

using brackets within brackets to indicate phrases within phrases. And, since an NP can in turn include a PP (which itself must include an NP within), we could also get a much more complex phrase:

[$_{NP}$ An elephant [$_{PP}$ in [$_{NP}$ my pajamas [$_{PP}$ from [$_{NP}$ the store]]]]]

And, iterating the NP rule once more:

[$_{NP}$ An elephant [$_{PP}$ in [$_{NP}$ my pajamas [$_{PP}$ from [$_{NP}$ the store [$_{PP}$ down [$_{NP}$ the street]]]]]]]

Now that we've explored a range of options for creating NPs, we also want to have rules for combining NPs (which serve either as subjects or objects) together with verbs (V) to form sentences (S). For instance, we might have:

S → NP + VP

VP → V + (NP) + (PP)

where VP stands for the category "verb phrase." This fragment of a phrase structure grammar would now allow us to get *both* versions of *I shot an elephant in my pajamas*.

As you can see from **Box 6.2**, we'd need a much more complete set of phrase structure rules to capture the full combinatorial possibilities for English. But even the small set of rules that are described in Box 6.2 can offer up some powerful machinery for creating new combinations.

It's worth noting that there is no consensus about the exact way to represent the syntactic structures of even quite simple sentences. Various frameworks have evolved, each relying on a number of arguments about why the structures should be represented in one particular way or another. These disagreements exist because once you start including a broad enough set of sentences in your data set, there are subtle aspects of linguistic structure that might lead to rethinking the original rules—you may have encountered this process already in considering possible rules for the sentences in Box 6.2. Also, different frameworks might have different ways of lining up the syntactic rules with their semantic effects. There are even disagreements about whether theoretical notions like phrase structure rules really reflect very closely how human minds represent the structures of language—or whether these rules are simply a convenient, shorthand way of expressing our knowledge of the syntactic

There's nothing in the syntactic rules that prevents this from going on in the same vein:

[The paper on [the desk in [the office on [the top floor of [the house]]]]]

and on:

[The paper on [the desk in [the office on [the top floor of [the house down [the street]]]]]]

and on:

[The paper on [the desk in [the office on [the top floor of [the house down [the street outside [the city limits]]]]]]]

Eventually, these recursive structures get to be too much for our memory to keep track of, an issue we'll pick up in Chapter 8. But it's in the nature of a generative syntax like this to allow a handful of phrase structure rules to generate what is, in principle, an infinite number of sentences that adhere to those rules. This makes a language incredibly powerful and creative

Now back to the big question: How do kids ever figure out that their language has the sort of syntactic structure it does? Over the course of hearing many sentences, they would need to register that certain words tend to occur in the same syntactic contexts, and group those words into the same type of syntactic category. They'd then have to notice that certain words tend to clump together in sentences, and that the various clumps occur in the same syntactic contexts, forming interchangeable constituents. They would also have to clue in to the fact that these interchangeable word clumps always have the same relationship to the meaning of the sentence, regardless of the clumps' specific content. Finally (as we'll see in more detail in Section 6.4), they would have to figure out the possibilities for moving constituents around in a sentence while keeping track of relationships between them over long spans of words. Any one of these learning tasks involves hearing and tracking lots and lots of sentences and then drawing quite abstract generalizations from these many sentences.

One of the most hotly debated issues in psycholinguistics is whether kids are helped along in the learning process by genetic programming that outfits them with specific assumptions about how languages work. Such innate programming would no doubt make learning syntax less daunting. And, despite the fact that languages of the world come in a great variety of syntactic flavors, there are a surprising number of ways in which they overlap in terms of their structural properties. Of all the possible systems that languages *might* have evolved to combine words, it seems that only a pretty restricted subset of these options ever turn up. And, interestingly, kids rarely seem to follow dead-end leads that would cause them to posit exotic grammatical systems that stray from what we think of as a human language. As a result, a number of researchers have proposed that children come into the world with a set of constraints that steer them away from ever considering grammars that fall outside the range of human grammars. In other words, children might have in place an innate **universal grammar** that provides them with a set of learning biases that line up nicely with the syntactic systems of human languages.

universal grammar A hypothetical set of innate learning biases that guide children's learning processes and constrain the possible structures of human languages.

Opponents of this view argue that no such extra help is needed. Kids, they suggest, are equipped with robust general-purpose learning machinery that is perfectly up to the task of learning syntactic generalizations from huge quantities of data. The reason children don't work themselves into a tight corner

BOX 6.3
A language without recursion?

In most languages, recursive structures are varied and abundant. For example, in English, recursion makes it possible to create complex phrases like these:

John's brother's house is big.

Frog and Toad are friends.

Cyprian will either marry his childhood sweetheart or he'll spend his life alone.

Ernie thinks that Bert wants to get a pet pigeon.

The letter that Juliet sent to Romeo went astray.

Language researchers have generally assumed that recursion is a universal property of language. But there may be exceptions. Dan Everett, a linguist who worked among the Pirahã people of the Amazonian jungle (see **Figure 6.2**), has claimed that sentences like the above, with their embedded constituents, don't exist in Pirahã (Everett, 2005). Instead of saying *John's brother's house*, speakers of this language would say something like:

Brother's house. John has a brother. It is the same one.

Or, instead of saying *The tiger got Jake and Lisa*, a Pirahã speaker would say:

The tiger got Jake. Lisa also.

One consequence of the lack of recursion is that, unlike English, Pirahã places an upper limit on the length of any given sentence in that language.

Everett's claims about Pirahã have been challenged by other linguists who question his analysis of the language, and because of limited access to this small, remote group of speakers, the controversy has yet to be resolved. But if Everett's claims are accurate, they raise some intriguing questions. Why would a language decline to make use of the powerful device of recursion in its syntax? And, from another angle, if humans can communicate with each other perfectly well without recursion in their language, why is recursion such a massively robust feature of language? What are the communicative costs and benefits of recursion?

Figure 6.2 The Pirahã people live on the banks of the Maici River in Brazil's Amazon Basin. As of 2010, there were about 420 individuals in this small community.

thinking that English sentences are generated by an alien grammar is because they have access to plenty of data that would quickly disabuse them of syntactic generalizations that fall too far from the mark.

What's more, there are alternative explanations for the syntactic similarities across languages, as explored in Chapter 2. Languages may share certain similarities *not* because the human brain is genetically programmed for certain structures, but because certain structures do a better job of meeting communicative needs, or lining up with the brain's strengths and weaknesses when it comes to learning or processing linguistic structure.

> **WEB ACTIVITY 6.2**
>
> **Discerning the rules** This activity will give you a taste of the task facing children who have to learn about the right syntactic generalizations for their language. You'll see sets of data from various foreign languages, as well as made-up "data" from invented alien languages. Can you spot the alien languages? What is it about them that makes their syntax seem un-human-like?

The disagreements between the nativist and data-driven views of syntactic learning tend to come down to several issues: First, what information *is* there in the input that a child hears—that is, is there actually enough evidence in the input to promote the right generalizations and circumvent the incorrect ones that kids might be tempted to make? Second, what kinds of learning mechanisms are available to kids throughout their learning trajectory? Third, what is it that children know about language structure anyway? Does their knowledge actually correspond to the kinds of abstract rules and representations that we've just posited for an adult's knowledge of syntax? In the rest of this chapter, we'll see that these questions have yet to be definitively answered. On the other hand, it's clear that we have a growing set of theoretical and methodological tools with which to address them.

6.2 Learning Grammatical Categories

How do children know about grammatical categories?

To crawl back into the mind of a young child in the early stages of language development, it always helps to consider an unfamiliar language. Let's look at Zapotec, an indigenous language of Mexico that you probably haven't encountered before. Using the following sentences as a basis, take a stab at drawing some syntactic generalizations, at least at the level of basic word order. What regularities can you see? (Warning: it may take a little while to work through these examples.)

ytaa'az gyeeihlly li'eb
bgu'tya' bzihny
ytoo'oh li'eb ca'arr
naa'ng banguual
naa li'eb banguual
gwua'ilreng li'ebr
rcaa'za ygu'tya' bzihny
binydyang dolf ytoo'oh pa'amm ca'rr
re'ihpy pa'aamm laa'reng gwua'llreng li'ebr

Got it? Good. Now, what is the correct way to order the following three words: *juaany, be'cw,* and *udiiny?* (Seriously, give this a try before reading any further.)

Actually, I suspect you can't tell—this wasn't an entirely fair exercise. You simply didn't have enough information in this little language sample to be able to figure out much of anything. But things change quite dramatically if I give you a bit of information about how the individual words map onto meanings. Using the mappings below, try again to derive some syntactic generalizations and figure out the correct order of *juaany, be'cw,* and *udiiny.*

ytaa'az—beat
bgu'tya', ygu'tya'—kill
ytoo'oh—sell
naa'ng, naa—be

gwua'llreng—read
rcaa'za—want
binydyang—hear
re'ihp—tell
bzihny—mouse
ca'arr—car
li'ebr—book
banguual—old
udiiny—hit
be'cw—dog
gyeeihlly, li'eb, pa'amm, dolf, and juaany—all proper names

This information makes all the difference. You should now be able to tell that *udiiny juuany be'cw* would give you a well-formed sentence.

It's worth thinking about *why* the information about how the individual words map onto meanings is so useful. It's useful because of the following assumptions you were implicitly making: (1) words that map onto the same general kinds of meanings (for example, actions rather than people or things) will occupy the same slots in the syntax, and (2) the syntactic patterns are affected by the role that entities play in a sentence (distinguishing, for example, the agents that instigate actions from the entities that are acted upon).

Where do these assumptions come from? Quite possibly from your knowledge of English, in which syntactic categories do tend to be made up of words that are similar in meaning. But what's interesting is that these assumptions turn out to be universally true of languages. Given that this is the case, and that they turn out to do so much work in breaking into the syntax of a new language, it seems sensible to ask whether children come into the world preprogrammed with certain basic preconceptions about the relationship between language structure and meaning. In order for these assumptions to be accessible at the very outset to a child learning a *first* language, they'd have to be innate.

This is the position taken by proponents of the **semantic bootstrapping hypothesis**. The idea is that the child comes equipped with innate expectations of certain grammatical categories as well as built-in mappings between key concept types and grammatical categories. For example, children might jump-start syntactic learning with the innate knowledge that nouns tend to be used to refer to objects, or that the subject of a sentence is typically the agent of the action that's being described.

As useful as it might be for babies to have such preprogrammed expectations, this doesn't necessarily mean that they have them. It might instead be the case that babies start off innocent of even the most basic facts about how meaning and syntax relate to each other, that they have to build from the ground up the notions that words fall into specific grammatical categories, and that this constrains not only aspects of their meaning, but also how they combine with other words in sentences. The arguments for innately based semantic bootstrapping become weaker if we can show that babies are able to learn these things easily and without any built-in assumptions.

To get a feel for how this might happen, consider again the Zapotec sentences you saw earlier. What would happen if you *didn't* assume from the get-go that words for people and things could be grouped together into a coherent syntactic category that patterned systematically in the language's structure?

semantic bootstrapping hypothesis The idea that children come equipped with innate expectations of certain grammatical categories, as well as built-in mappings between key concept types and grammatical categories.

distributional evidence The tendency of words or types of words to appear in certain syntactic contexts, allowing extrapolation of these tendencies to newly learned words.

What if you had to rely *only* on the evidence that was there in front of you in the sentences themselves, without the benefit of any preconceptions that words that are similar in meaning might behave similarly in the syntax of a sentence? How would you form a concept of a grammatical category, and how would you figure out which words belong to it?

You likely could eventually break into the system, but it would take you many more example sentences than just the few that you were offered here. But eventually, with enough data, you might notice, for instance, that certain words like *ytaa'az* or *re'ihpy* only ever occur at the beginnings of sentences, so you'd begin to group them as belonging to the same class of words (let's call this Class A for the moment). You would then also notice that only certain words can appear immediately after the Class A words, words like *li'eb* or *bzi-hny*, which we'll call Class B words. Given enough **distributional evidence** of this sort—that is, evidence about the tendencies of words to appear in certain syntactic contexts—you could come up with some generalizations about word order. Once you moved on to learning the meanings of some of these words, you might then notice that Class A words tend to be words that refer to actions, and that Class B words tend to refer to people, things, or animals. This would then allow you to make very reasonable guesses about whether new words you meet should belong in the syntactic categories of Class A or B, once you had some idea of their meanings. From this point on, you would be in a position similar to the one you started the exercise with—that is, with certain ideas firmly in place about the mapping between syntactic categories and meanings. It just would have taken you a while to get there. Moreover, paying attention to distributional evidence, rather than meanings alone, would ultimately lead you to a more accurate understanding of these categories (see **Box 6.4**).

Is distributional evidence powerful enough?

To make a convincing case that youngsters can form syntactic categories by tracking distributional evidence, we'll need to answer several questions. The first of these is: If we look beyond just a small sampling of language, how reliable would this distributional evidence be? And if distributional evidence does turn out to be a reliable source of information for forming grammatical categories, our second question would be: Is there any evidence that small children are able to track distributional evidence and group words into categories accordingly?

Let's start with the first question regarding the reliability of distributional evidence. In arguing for the semantic bootstrapping hypothesis, which takes the position that some assumptions about syntax are innate, Steven Pinker (1987) noted that a child could run into trouble if he were relying solely on distributional patterns in sentences such as these:

(1a) John ate fish.

(1b) John ate rabbits.

(1c) John can fish.

If our alert toddler happened to notice that *fish* and *rabbits* occur in exactly the same syntactic environments in examples 1a and 1b, he'd be in danger of concluding that the following sentence is also perfectly good:

(1d) *John can rabbits.

The problem here is that a single word, *fish*, actually falls into more than one syntactic category; it can act as either a noun or a verb, while *rabbits* can't. But

BOX 6.4
Science is *not* a verb

Michael Shermer, editor of *Skeptic* magazine, is a fervent advocate for science. In a 2006 interview with Kevin Berger of *Salon* magazine, he said:

> We've got to get past this idea that science is a thing. It isn't a thing like religion is a thing or a political party is a thing. It's true that scientists have clubs. They have banners and meetings and they drink beer together. But science is just a method, a way of answering questions. It's a verb, not a noun.

Shermer's point is that, rather than treating science as a collection of facts or ideological beliefs, we should understand that it's a process that demands active engagement.

I appreciate Shermer's take on science. Truly, I do. But I can't get on board with the idea that science is a verb. The notions of verb-hood and noun-hood are, at their core, notions of *structure* and not meaning, defined in terms of which slots in a sentence they can occupy. And it's plain to see that the word *science* occupies noun-y slots rather than verb-y ones:

Science is not a verb.

Rice is not a liquid.

Ink is not a food.

The science of language is cool.

The hair of kittens is soft.

The president of the company is foolish.

The word *science* sits in the same slots as the words *rice*, *hair*, and *president*—all of which are, uncontroversially, nouns. If *science* were a verb, we'd be able to say things like:

I think we should have scienced that.

Did you remember to science your theory?

He presented his findings after sciencing in solitude for a year.

Common "knowledge" is that nouns refer to "things" and verbs to "actions." And, as argued in this chapter, the match between the *syntactic* notions of nouns and verbs and the kinds of *meanings* nouns and verbs tend to capture is potentially very useful for language learning. This match-up of syntax and meaning provides clues about the syntactic categories of new words we encounter. But the match is only a tendency; ultimately, learners of a language have to cope with the fact that actions and events can be nouns (*destruction; revenge; flood*) and that not all verbs involve actions (*need; correspond; negate*). So, the meaning of a word can sometimes provide misleading or unhelpful information about its syntactic category. When this happens, distributional cues need to prevail.

how could our child tell that *fish* in 1a is in a different category than *fish* in 1c? If he knew that it referred to a thing in 1a but an activity in 1c, and if he knew that these meanings mapped to different categories, he'd manage not to go astray. But without these assumptions in place, he could be misled.

Distributional evidence might be messy in other ways that would preclude our child from learning the right categories for words. For example, if he were paying attention to which words can occur at the beginnings of sentences in English, he'd come across examples like these:

(2a) John ate fish.

(2b) Eat the fish!

(2c) The fish smells bad.

(2d) Whales like fish.

(2e) Some of the fish smells bad.

(2f) Quick, catch the fish!

(2g) Never eat fish with a spoon.

lexical co-occurrence patterns Information about which words tend to appear adjacent to each other in a given data set.

bigrams Sequences of two words (i.e., word pairs).

trigrams Sequences of three words.

In each of the above examples, the first word of the sentence belongs to a different syntactic category. Assuming that children aren't fed a carefully regimented sample of speech that manages to avoid these problems until they've properly sorted words into categories, how useful could distributional evidence be?

It's obviously going to matter *what* distributional evidence we consider. Looking just at the left edge of a sentence doesn't produce great results for English, but other patterns could be much more regular. For example, looking at **lexical co-occurrence** patterns (that is, at information about which words tend to appear adjacent to each other in a data set) could prove to be more promising. For instance, in English it turns out to be fairly easy to predict which category can occur in which of the following slots in these word pairs, or **bigrams**:

(3a) the __

(3b) should __

(3c) very __

Chances are, if asked, you'd supply a noun in 3a, a verb in 3b, and an adjective in 3c. And lexical co-occurrence patterns get even more useful if we look at sequences of *three* words, or **trigrams**:

(4a) the __ is

(4b) should __ the

(4c) very __ house

Note, for example, that either an adjective or a noun can occur after *the*, but only a noun can occur between *the* and *is*. Similarly, examples 4b and 4c are more constrained with trigrams than are the bigrams in 3b and 3c.

Pursuing this idea, Toby Mintz (2003) set out to measure the reliability of sequences, or "frames," such as *the __ is* or *should __ the*. He looked at a large database containing transcriptions of the recorded speech of parents talking to toddlers and pulled out the 45 most frequent sequences of three words in which the first and last words were the same. So, for instance, the sequences *the doggy is* and *the bottle is* count as two instances of the frame *the __ is*. He then measured, for each of these frames, how accurate it would be to make the assumption that the middle words in the trigrams would always be of the same grammatical category. This measure of accuracy reveals just how predictive the frame is of that intervening word's category.

Mintz found that, by relying just on these frequent frames in the database, it would be possible to correctly group the words that occurred in the frames with an accuracy rate of better than 90%. Hence, it seems there's pretty sturdy statistical information to be had just on the basis of distributional evidence (provided, of course, that small children are able to tune in to just this type of handy statistical evidence while perhaps ignoring other, less useful distributional evidence).

Using Mintz's "frequent frames" is only one possible way of capturing statistical regularities for words of the same grammatical category. Other formulations have also proven useful to greater or lesser degrees, and the usefulness of various statistical strategies may vary from language to language. Leaving aside for the moment the issue of exactly how best to capture distributional regularities, it's at least fair to say that information from lexical co-occurrence is reliable enough to allow kids to make reasonably good guesses about the category membership of a large number of words, once they've been exposed to enough examples.

So, what about our second question: Are very small children able to make use of these regularities? We saw in Chapter 4 that babies as young as 8 months can track statistical patterns over adjacent syllables and are able to use this information to make guesses about word boundaries. So, there's already good reason to suspect that they might also be attuned to statistical patterns in inferring the grammatical categories of words. And some recent studies of babies in their second year provide more direct evidence for this view. For instance, Mintz (2006) studied how 12-month-olds would react upon hearing novel words in frequent frames within sentences such as *She wants you to deeg it* or *I see the bist*. Some words, like *deeg*, consistently appeared in verb slots, while others, like *bist*, appeared in noun slots during a familiarization phase. Mintz then measured looking times during a test phase and found that the babies distinguished between grammatical versus ungrammatical sentences that contained these novel words. That is, they looked longer at the "ungrammatical" sentence *I bist you now* than at the "grammatical" *I deeg you now* (see **Figure 6.3**).

WEB ACTIVITY 6.3

Finding syntactic categories Are you smarter than a 15-month-old? Try this activity to see if you, too, can pull syntactic categories out of a sample of data from an artificial language. The stimuli come from a study by Rebecca Gómez and Jessica Maye (2005).

Other studies, using artificial grammars, also suggest that beginning at about 1 year of age, very small children can sort novel words into categories based solely on distributional information. In fact, they seem to be able to do this even when the distributional information is not perfectly consistent, suggesting that small doses of problematic examples like those flagged by Pinker in examples 1a through 1c above might not present a dire problem for learning. The exact nature of the statistical information that children can draw on is still not clear, though. This may well change over the course of development; for example, it may be that information involving adjacent words in the form of bigrams is easier for kids to clue in to earlier in their development than information involving non-adjacent elements, as in the frequent-frames hypothesis.

Finding evidence that children can form categories on the basis of distributional evidence alone does not, of course, rule out the possibility that they *also* tap into innate preconceptions about the relationship between grammatical categories and their typical contributions to meaning. It's perfectly possible that children lean on both kinds of information.

(A) Familiarization phase

Group 1

Verb Frame Sentences	Noun Frame Sentences
She wants *to deeg it.*	I see *the gorp in* the room.
She wants *to lonk it.*	I see *the bist in* the room.
You *can deeg.*	That's *your gorp.*
You *can lonk.*	That's *your bist.*
Can *you deeg the* room?	I put *his gorp on* the box.
I lonk you now!	Here's *a bist of* a dog.

Group 2

Verb Frame Sentences	Noun Frame Sentences
She wants *to gorp it.*	I see *the deeg in* the room.
She wants *to bist it.*	I see *the lonk in* the room.
You *can gorp.*	That's *your deeg.*
You *can bist.*	That's *your lonk.*
Can *you gorp the* room?	I put *his deeg on* the box.
I bist you now!	Here's *a lonk of* a dog.

(B) Test phase

Grammatical for Group 1; ungrammatical for Group 2

Can you lonk the room?
I deeg you now!
I put his bist on the box.
Here's a gorp of a dog.

Grammatical for Group 2; ungrammatical for Group 1

Can you bist the room?
I gorp you now!
I put his lonk on the box.
Here's a deeg of a dog.

(C) Results

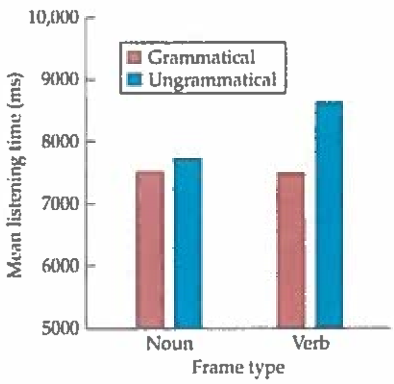


Figure 6.3 Mintz used the head-turn preference procedure (see Method 4.1) to measure the ability of infants to infer syntactic categories from distributional evidence. (A) Infants heard nonsense words within bigrams (word pairs, shown in italic) or trigrams (three words) in either noun-supporting or verb-supporting sentence frames, unaccompanied by any semantic context. To make sure some nonsense words weren't intrinsically more "noun-y" or "verb-y" than others, the subjects were divided into two groups. The words presented to Group 1 as nouns were presented to Group 2 as verbs, and vice versa. (B) In the test phase, the words were presented to the infants in grammatical and ungrammatical sentences. (C) Results, showing mean listening times to grammatical and ungrammatical sentences, by frame type. (Adapted from Mintz, 2006.)