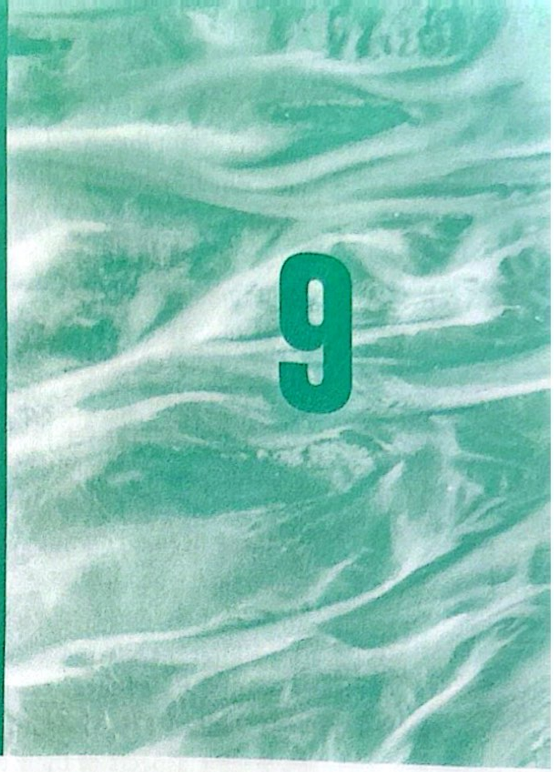


9

Language Development in Infancy and Early Childhood

Babies are like people, only smaller.

PROBABLY STEVEN PINKER

The vast majority of children learn the language that their parents and peers speak without special instruction, threat, or reward. Children sometimes experience difficulty mastering some of the more subtle aspects of their native language, but unless they have particular forms of genetic defect, such as those that result in severe cognitive impairments or *developmental language disorder*, children will achieve a high degree of skill understanding and producing language. The apparent ease with which children spontaneously acquire these abilities disguises the extraordinary challenges that they face in mastering language and the difficult problems they overcome in the process. This chapter details some of these challenges and reviews studies that give us hints about the strategies children use to overcome them.

Prenatal Learning

Behaviorist approaches to learning (e.g. Skinner, 1957) have viewed children as a *tabula rasa* ("blank slate"). The child's experience fills up the slate with knowledge about everything, including language.

Introduction to Psycholinguistics: Understanding Language Science, Second Edition.
Matthew J. Traxler.

Companion Website: www.wiley.com/go/traxler/psycholinguistics2e

While children have general-purpose learning mechanisms, they are born knowing nothing about the world in which they will function. Behaviorists in the classical period subscribed to the hypothesis that *babies are dumb*.

More recent research in child development has shown that, rather than being blank slates, young infants have an innate appreciation of important aspects of the world, including important principles of visual perception such as occlusion, the physical properties of objects and substances, and numbers (Aguiar and Baillargeon, 1999; Hespos et al., 2009; Izard et al., 2009). In light of these newer findings, many cognitive and developmental psychologists subscribe to some version of the *babies-are-smart* hypothesis, and many language scientists are now engaged in trying to find out just how smart babies are with respect to language skills.

Nativist approaches to language (e.g. Lenneberg, 1967) view language abilities as resulting from adaptation and natural selection. Researchers in this tradition have proposed that, when babies are born, they do not possess knowledge of a specific language—no one is born speaking French or Italian—but they do have innate learning mechanisms that allow the child to figure out how the adult language works. "There is a special cognitive faculty for learning and using language" (Pinker, 1984, p. 33; see also Chomsky, 1981; Pinker, 1984, 1994b). These learning mechanisms cause children to pay attention to specific aspects of their environment and to organize their perceptions in ways that lead to a uniform understanding of the adult language system. The nativist position falls within the smart-baby tradition, in that it views the infant as being born with certain kinds of knowledge already in place.

Just because young infants appear to have knowledge about number, occlusion, and some principles of physics, that does not automatically mean that they have comparable knowledge about language. Whether children are born with knowledge of language therefore needs to be investigated. If we can determine what kinds of language-related knowledge infants and small children have at different points in time, that will help us decide whether that knowledge was gained as the result of experience (as the behaviorists claim) or whether that knowledge was gained by associating perceptual experiences with pre-existing knowledge structures (as the nativists claim).

One problem in discriminating between the behaviorist and nativist positions arises because learning can start before birth. Just because a baby is born with a particular kind of knowledge, that does not mean that the knowledge is *innate* (if by *innate knowledge* we mean *instinctive*, *self-generated*, or *not caused by experience*). For example, newborn infants are able to tell the difference between recordings of someone speaking their native language, the language that their parents speak, and the same person speaking a different language. At 2 days old, children of French-speaking parents can detect when a bilingual female speaker is speaking French and when the same speaker is speaking Russian (Mehler et al., 1988). They prefer to listen to French, their native language, rather than Russian. Does that mean that knowledge of French is innate? Of course not. French people speak French because they are exposed to French-speaking models, not because they have "French language" genes. An alternative hypothesis says that French babies learn what French sounds like after one day of experience with French, most of which they spend asleep. But that, too, is highly unlikely. There is a third possibility, however, one that involves learning before birth—*prenatal* learning. How might this prenatal learning happen?

The fetus' auditory system is capable of perceiving environmental input in the third trimester of pregnancy, the last 90 days or so. Pregnancies normally last for about nine months. During that period of time, environmental sounds reach the fetus' ears and are processed by the developing auditory system. The fetus is encased in a sack of amniotic fluid and insulated from the outside world by the mother's body. Because the mother's voice is generated in and propagated through the mother's body, it is, on average, the loudest thing the fetus is exposed to. When babies are born, they show a strong preference for their mother's voice compared to other female voices, but not their father's compared to other male voices. This is probably because, from the fetus' perspective, mom's voice is a lot louder than dad's before birth (DeCasper and Fifer, 1980). The amniotic fluid and maternal tissue act as a filter that reduces the amplitude (loudness) of environmental sounds that originate outside of the mother. Low-frequency sounds (bass range) are less affected as they pass from the air, through the mother, to the fetus. The amplitude of high-frequency (treble range) sounds is greatly reduced. The speech sounds that make one phoneme different from another, and therefore make one word different from another, are mostly carried in the high-frequency portion of the acoustic signal. As a result, fetuses are not exposed to those parts of the speech signal that would enable them to start learning words. However, prosodic characteristics of speech (relative loudness or *accent*, fundamental frequency, tempo, pauses, and so forth) are largely carried by the lower frequency components of the acoustic signal, so they *are* available to the fetus.

A third-trimester fetus has a functioning perceptual system that can take advantage of auditory stimuli that are present in the intra-uterine environment. Specifically, the fetus is exposed to auditory stimuli that provide the basic outline for the prosodic features of the fetus' native language. As we will see shortly, prosodic features can help children identify important components of the speech signal, and a fetus that learns something about the prosody of their native language will benefit when it comes time to learn their first language. But do fetuses respond to prosodic information? And do they retain any memory of their prenatal experience after being born? How could we find out?

Babies suck

Babies really do suck. It is one of the few things they can do well when they are born. Babies suck to eat, but they also engage in *nonnutritive sucking* in between feeding times. It turns out that you can train even newborn infants to suck hard by rewarding them for doing so.

The *high-amplitude sucking* (HAS) research technique takes advantage of this fact. In HAS experiments, babies, even newborn infants as young as 2 days old, are connected to a device called a *pressure transducer* that measures how much pressure babies exert when they suck, and how often they suck. A pacifier with a very small hole in the end is inserted into the baby's mouth. During a baseline period, the machine measures how hard and how often the baby sucks when no stimulus is presented. The measurement of how hard the baby sucks is referred to as *amplitude* (higher amplitude means the baby exerts more pressure). Then, during a training period, the baby is rewarded with an *appetitive stimulus* (something the baby likes), such as the sound of the mother's voice, when the baby exerts more sucking pressure than the average amount, or sometimes when the baby sucks faster than normal.[1]

When the baby's sucking amplitude drops back down to baseline, the baby is said to have *habituated* to the stimulus. The baby is acting as though they are bored with the stimulus. The more a baby likes the stimulus, the longer the infant will produce high-amplitude sucks to keep the stimulus going. If two different stimuli are played, you can figure out which one babies prefer by seeing how long they are willing to keep their sucking amplitude high. You can also use this technique to find out whether an infant can tell the difference between two stimuli. If you play a stimulus to babies long enough, they will habituate and their sucking amplitude will drop to baseline. If you change the stimulus at that point, and babies notices the change, they will *dis-habituate*. That is, the babies' sucking amplitude will increase until they get tired of the new stimulus. If a baby does not notice the difference between the old and the new stimulus, and treats them as being identical, the baby's sucking amplitude will stay low when you change from the original stimulus to a new one.

HAS experiments have been used to investigate the way prenatal exposure to speech affects fetuses and newborn infants (DeCasper et al., 1994; DeCasper and Spence, 1986; Krueger et al., 2004). In one set of studies, pregnant mothers recited a short story (such as *The Cat in the Hat*) two times a day every day during the last six weeks of their pregnancies. During this training period, fetuses had the opportunity to become familiar with one particular story. After the babies were born, they were tested using a version of the HAS technique. This testing period was completed before the babies were 2 days old (the youngest infants were 44 *hours* old). All of the test babies listened to the familiar story (e.g. *Cat in the Hat*) as well as a new story read by the same person. Some of the babies were tested with recordings of their own mothers reading a familiar story, but half of the babies listened to a recording of an unfamiliar female reading the familiar story. If babies learned anything about the familiar story and retained that information over time, then they should behave differently when they heard the familiar story compared to when they heard the new, unfamiliar story.

No matter who read the story, the baby's own mother or somebody else, newborn infants worked harder to hear the familiar story than they did when the unfamiliar story served as the reinforcer. These results show that fetuses did learn something about language from their environment prior to being born. Because high-frequency information was filtered out of the acoustic stimulus during the training phase (before the babies were born), the babies' preference could not be based on the specific words in the familiar story. Similarly, because babies preferred the familiar story no matter who read the story, their behavior during testing could not reflect a preference for one specific voice (i.e. the mother's).

What caused the newborn infants to prefer the familiar story over the unfamiliar one? The most likely cause is the prosodic cues available in the speech signal. Although prenatal fetuses could not hear the specific words in the story, prosodic information, including alternating patterns of loud and soft sounds, patterns of high and low tones, pauses, and so forth would have been available. Those patterns are consistent enough across speakers that babies could detect the familiar prosodic pattern, even when the story was read by someone other than their own mother. These results show, then, that fetuses do respond to prosodic cues, and that they retain information about prosodic patterns. Therefore, prenatal learning of prosodic features could lay the foundation for further language learning after the infant is born.

Could it be possible that babies actually experience very rapid learning after birth? Could the preceding experimental results reflect a form of super-fast acquisition? It's unlikely, given that the baby's exposure to the familiar story occurred only before birth, but additional experiments really lay this issue to rest (DeCasper et al., 1994; Krueger et al., 2004; see also Barajas et al., 2021; Krueger and Garvan, 2019). These studies also used a study-test design, similar to the original DeCasper and Spence experiment. But instead of waiting for babies to be born, the researchers tested them while they were still inside their mothers. This prenatal testing took advantage of another fact about fetus physiology: In the last trimester, the fetus' heart rate changes when they process acoustic information. Specifically, the onset of an acoustic stimulus causes the heart rate to slow down (*cardiac deceleration*). Fetuses also respond to a change in the acoustic stimulus with cardiac deceleration (Lecanuet et al., 1988). If fetuses fail to notice when one acoustic stimulus is replaced by another, then there should be no change in heart rate. Thus, heart rate can be used to test whether fetuses can spot the difference between two stimuli.

Taking advantage of infants' heart rate response to different stimuli, researchers assessed how third-trimester fetuses responded to recorded speech. In one such study, mothers recited brief nursery rhymes three times a day while their fetuses were in a quiet (*quiescent*) state. This training period lasted for about a month. Then, while the fetuses' heart rate was measured, they listened to a recording of a different female reading either the same nursery rhyme or a new, unfamiliar nursery rhyme. Fetuses showed greater cardiac deceleration for the familiar rhyme as compared to the new rhyme, and these effects were larger for older fetuses than for younger ones (most likely reflecting maturation of the nervous system). Thus, fetuses learn and retain information about speech before they are born. Most likely, they learn about and remember the prosodic qualities of speech. These results also largely rule out the "super-fast post-birth" learning hypothesis.

The reality of prenatal learning complicates the debate between the nativists and the behaviorists. Nativists have relied on abilities that appear in very young infants to argue for the existence of instincts or innate (presumably genetically driven) abilities. If an ability appears in a newborn infant, before that infant has had much exposure to the environment, then it is unlikely that the infant's behavior reflects learning based on exposure to the environment. If babies were born knowing how to ride a bicycle, we would have to consider the possibility that their genetic endowment wired their brains in just the right way to produce that ability. Similarly, if babies were born knowing something about how language works, we would have to consider the possibility that those abilities result from genetically driven neural organization. However, because we know that fetuses learn, we have to rule out prenatal learning before we conclude that abilities present in newborns reflect a genetically driven brain mechanism. In the case of voice preference (newborns prefer their own mother's voice over others), language preference (newborns prefer their native language over others), and story preference (newborns prefer familiar speech patterns over unfamiliar ones), prenatal learning because of exposure to specific stimuli suffices to explain the results, and appeals to innate, genetically driven, language-specific mechanisms are unnecessary. However, ruling out innate sources of these particular preferences does not rule out all innate knowledge. It just means that learning from the environment has to be considered an alternative hypothesis any time anyone appeals to innate knowledge to explain some aspect of an infant's behavior.

Infant Perception and Categorization of Phonemes

Infants learn about their first language by listening to people talk. To learn their first language, infants need to solve a number of different puzzles. They have to figure out which speech sounds (phonemes) occur in the language (and which do not), how those phonemes go together to make words, what words mean, and how words go together to make sentences. They have to do all of this without any deliberate instruction.

Let's start by reviewing infant speech perception. Phonemes constitute the building blocks of words in the baby's new language. One basic task that the infant must solve is to figure out how the inventory of speech sounds is organized—which differences between sounds matter, and which can be safely ignored. This task is more complicated than it might seem at first. One complicating factor is that, although different speech sounds may have physical similarities, no two speech sounds are identical. For example, every speaker's voice is unique because of physical differences between different vocal tracts. Some speakers have higher pitched voices than others, so their speech sounds will contain relatively more high-frequency energy. Some people speak faster than others, so their speech sounds unfold differently in time. Despite these differences between different speech sounds, adults "shoehorn" widely diverse acoustic signals into a fixed number of categories (about 40 in English). Adults are so good at this that we do not normally notice differences between sounds that belong to the same category, even though detailed physical analysis shows that those sounds have different physical properties.

In some ways, young infants respond to speech sounds very much like adults do. For instance, studies of infants show that they experience categorical perception, just like adults. Using a variant of the HAS technique, Eimas and colleagues (Eimas et al., 1971) presented 1- and 4-month-old infants a syllable (i.e. /ba/) that has a short *voice onset time* (or VOT, the amount of time that elapses between the very beginning of the speech sound and the time when the vocal folds start to vibrate). After infants habituated to that syllable (as evidenced by a reduction in how fast they sucked on their pacifiers), the original stimulus /ba/ was replaced by one of two other stimuli. Sometimes, the syllable was replaced by another stimulus with another short VOT. Adults would perceive both of these stimuli as being /ba/, despite the fact that the two stimuli had slightly different VOTs. Sometimes, the original stimulus would be replaced by another stimulus that had a long VOT. Adults would perceive the second stimulus as being /pa/, rather than /ba/. The question was: Would infants notice the change in stimuli when the two came from the same category (as defined by adults' perception of the two)? And would they notice the change when the two stimuli came from different categories (again as defined by adults' perception)?

When the original and replacement stimuli both had relatively short VOTs, infants treated the two as being identical. That is, after they habituated to the original stimulus, their rate of HAS stayed the same when the short VOT stimulus was replaced by another short VOT stimulus. But when a long VOT stimulus replaced a short VOT stimulus, babies noticed the change, as evidenced by an increase in HAS just after the new stimulus replaced the old one. The results of this study suggested that infants, like adults, treat speech sounds as belonging to discrete categories. When two speech sounds had similar voice onset properties, infants treat them as being identical, even though the two sounds had different physical properties (different VOTs). When two speech sounds come from categories that adults perceive as separate (such as /ba/ versus /pa), infants also perceive these sounds as being separate.

Findings like these have been used by nativists to support the claim that *speech is special*. That is, they claim that humans perceive speech sounds using specialized mechanisms, possibly genetically determined, and that people treat speech differently than other kinds of sounds (e.g. Eimas et al., 1971; Liberman et al., 1967; Pardo and Remez, 2006). This claim has been challenged by researchers who note that other species besides humans treat speech sounds as categorical (Kluender and Kiefte, 2006)[2] and that people perceive nonspeech sounds categorically as well. Because chinchillas and Japanese quail can be trained to categorize speech sounds similar to the way people do, categorical speech perception does not result from a unique human adaptation for speech processing. If categorical perception resulted from a special-purpose speech-processing system, we have to assume that chinchillas and Japanese quail have the same special system, which doesn't make sense because they gain no fitness benefit from categorizing speech sounds.

Categorical perception of nonspeech sounds has been demonstrated in experiments using synthesized sounds (Cutting and Rosner, 1974; Jusczyk et al., 1980, 1977). These synthesized sounds mimic some of the characteristics of musical instruments. For example, plucking a string on a violin creates the equivalent of longer VOT because there is a lag between the onset of the sound and the point in time where the violin string starts to vibrate in a steady state.

When synthetic sounds that mimic the properties of naturally occurring nonspeech sounds are played to adults and 2-month-old infants, both adults and infants perceive physically different stimuli as belonging to the same category. Stimuli with short lags between sound onset and steady-state vibration are perceived as sounding like someone dragging a bow across a string. Stimuli with longer lags between sound onset and steady-state vibration are perceived as sounding like someone plucking a string. As the lag between onset and vibration changes from short to long, there is an abrupt transition in the way the sound is perceived. At one lag, the sound is perceived as a "bowed" sound. At a very slightly longer lag, perception changes to "plucked."

Because infants treat both speech and nonspeech sounds as belonging to discrete categories, and because small changes in the physical characteristics of both speech and nonspeech sounds can lead to large and abrupt changes in the way infants perceive those sounds, we can either conclude that children are born with two very similar "special" mechanisms, one that applies to speech sounds and one that applies to nonspeech sounds, or that categorical perception of speech results from a more general, "nonspecial" sound-processing system. While categorical perception may be the product of a biologically determined perceptual mechanism, the mechanism that produces categorical perception does not seem to have been selected for in humans specifically because it gives children an advantage in processing speech sounds.

Categorical perception of speech cannot be used as evidence for a genetically determined speech-processing mechanism (because other sounds produce categorical perceptions and because other animals besides humans perceive speech categorically). However, other evidence points toward a genetic basis for some aspects of speech processing. In particular, infants appear to have an innate (unlearned) preference to listen to speech (or attend to language-related gestures; Jusczyk, 1997; Jusczyk and Bertoncini, 1988; Krentz and Corina, 2008; Stone and Bosworth, 2019; and particular kinds of speech are more attractive to infants than others; Cooper and Aslin, 1990).

Because some aspects of speech processing (e.g. categorical perception) seem to reflect general-purpose perceptual mechanisms, while other aspects appear to be innate (e.g. preference to listen to speech), Jusczyk and Bertoncini hypothesized an *innately guided* learning process to explain how children acquire their first language. The innate part of the innately guided process is twofold. First, the infant has an instinct (innate drive) to pay attention to specific aspects of the environment, speech sounds in particular, and to undertake especially detailed processing of speech sounds. The initial state of the innate learning system must be general enough that the system can learn any possible human language. If the initial state were too narrowly focused, then the child might be born into a community with a language that fell outside the preset state of the learning mechanism. To ensure that the mechanism is capable of learning any human language, "The infant is innately prewired with broad categories that may develop in one of several different directions" (Jusczyk and Bertoncini, 1988, p. 233). The learning part happens when the infant is exposed to a specific language, and this exposure causes the child to move from the initial state, with little knowledge specific to the language, to the adult state, with fully developed phonological, lexical, morphological, and syntactic knowledge.

To determine whether this account is plausible, we could look for evidence that young infants have broad speech categories and that those categories become refined as the child is exposed to one specific language. Such evidence can be found in studies of phonological processing in newborn and very young infants. For infants to be able to learn any human language, they must be capable of identifying important distinctions between different speech sounds (phonemes). A child could be genetically prewired to recognize specific *phonological contrasts* (differences between speech sounds, such as the characteristics that make the /p/ sound different than the /b/ sound). But this prewiring would be unnecessary if the adult language did not make use of that distinction.

English, for example, does not make use of *aspiration* (whether a burst of air comes out while a phoneme is being produced),[3] but Hindi does. So, while aspiration or the lack thereof does not create a meaningful difference in English phonology, it does in Hindi. Similarly, adult Japanese speakers do not hear the difference between the *liquid* phonemes /r/ and /l/, because that distinction is not meaningful in Japanese. English adults do hear the difference, because that phonological contrast is meaningful (root is different than loot). It is less harmful to be born being able to perceive a non-helpful contrast (such as the aspiration contrast in English). Being born with such a nonfunctional contrast may cause you to believe that there are two speech sounds in your native language when there is only one, but that will not prevent you from learning differences between different words. However, if you are born with phonological categories that blur or eliminate the difference between two different phonemes, then you will not be able to learn some words (you would treat *root* and *loot* as being the same word).

Evidence from infant speech perception studies suggests that infants are born with the ability to recognize most phonological contrasts (Streeter, 1976; Werker et al., 1981; Werker and Tees, 1983, 2002; see also Dehaene-Lambertz and Dehaene, 1994). Very young infants have the ability to detect the difference between phonemes even when they have had no opportunity whatsoever to hear the difference. In the Kikuyu language, there is only a single labial stop consonant (roughly equivalent to the English consonant /b/). *No* corresponding Kikuyu consonant differs from /b/ only in VOT. So, Kikuyu speakers do not have a labial stop consonant comparable to English /p/. Because there is no distinction between /b/ and /p/ in Kikuyu, children reared in a Kikuyu-speaking environment do not have the opportunity to learn that distinction via exposure to naturally occurring language. However, when 2-month-old infants were tested on stimuli that were the English equivalents of /b/ and /p/, they could detect the difference. Kikuyu infants in a HAS experiment listened longest to the phoneme /b/, which did sound like a phoneme that they would have heard before (this reflects a kind of familiarity preference that is often observed in infant language studies). The infants dishabituated to two different phonemes with longer VOTs than the preferred stimulus. They listened longer to the two long-VOT stimuli than they did when the same stimulus was presented before and after the change point. Similar results have been obtained for 7-month-old infants of English-speaking parents tested on Hindi aspirated/non-aspirated phoneme contrasts, and for 6-month-old infants of English-speaking parents tested on contrasts that are present in Thompson (a language used by First Nations British Columbia residents).

In all of these cases, infants perceive contrasts between phonemes that do not exist in their native language, that they have not had the opportunity to learn via exposure, and that adult speakers of those languages are incapable of perceiving. Results like these show that children are born with a categorical organization of phonology that enables them to detect contrasts that may or may not be important in their native language. Because infants are sensitive to a huge variety of phonological contrasts, infants never face the problem of having too few phonological categories, and any contrast that is important in the adult version of the language can be preserved. The main problem infants need to solve, then, is to figure out which of the multitude of possible contrasts actually do matter, and to organize their phonological categories so that different versions of the same phoneme are mapped together into the same category (while still maintaining distinctions between speech sounds that really do represent different phonemes).

Experimental data show that the way infants organize phonological categories, and therefore the way they respond to speech stimuli, changes dramatically in the first year of life. Infants' ability to perceive non-native phonological contrasts is greatly reduced by their first birthday, and the way they respond to native language phonological contrasts also changes (Barajas et al., 2021; Jusczyk, 1997; Kuhl et al., 2006; Werker and Tees, 1983, 2002). Thus, exposure to language in the environment appears to reinforce some phonological contrasts while eliminating others. How is that accomplished?

Some theorists suggest that phonological *prototypes* emerge from the infants' experience (Kuhl et al., 1992; Reh et al., 2021). These prototypes represent a kind of perceptual average of specific instances of a given phoneme.[4] The idea is that the infant is exposed to many repeated instances of a given phoneme, because different speakers are all providing input to the infant, and because the same speaker will pronounce a given phoneme differently on different occasions. By mapping phonemes together on the basis of physical or perceptual similarity, and by treating the "average" as a perceptual "magnet," the infant learns to de-emphasize minor variations in pronunciation and to super-emphasize larger variations. This leads to sharper distinctions between phoneme categories that are used *contrastively* (changing the sound changes the meaning of the word it appears in). Additionally, phoneme contrasts that are not meaningful will eventually drop out of the system of phonological representations, because different versions (such as aspirated versus non-aspirated /p/) will all be attracted to the same perceptual "magnet."

To test this possibility, researchers have looked at phoneme prototypes in different languages. For instance, Swedish and English both have phonemes that sound like "ee," but there are minor variations in pronunciation across the two languages so that the English "ee" has a different prototype than the Swedish "ee." When 6-month-old Swedish and English infants were exposed to minor variations from each prototype, they responded to those minor variations in different ways. Swedish infants treated minor variations of the Swedish prototype as being the same phoneme, but they treated equally minor variations of the English prototype as being completely different phonemes. English infants did the same thing, just in reverse. They treated minor variations from the English prototype as being the same as the prototype; equally minor variations from the Swedish prototype led to big changes in the infants' response.

Solving the Segmentation Problem

To learn their first language, infants need to figure out where the words are. They need to accomplish that task even though fluent speech does not provide them with clear and unambiguous indications of where words begin and end much of the time (Cutler, 1996; Jusczyk, 1997). Babies do not typically hear single words spoken in isolation. Only about 10% of the words young infants hear are spoken without any other words in the same utterance (with different mothers producing as few as 5% or as much as 17% of their total infant-directed output as isolated words; Brent and Siskind, 2001; Fernald and Morikawa, 1993). In experimental situations where mothers are asked explicitly to teach new words to their children, the frequency of isolated words in the mother's speech is still a relatively low 28% (Woodward and Aslin, 1990, in Jusczyk and Aslin, 1995).

Nonetheless, some researchers (e.g. Bortfeld et al., 2005; Brent and Siskind, 2001; R. L. Frost et al., 2020) hypothesize that infants learn some words by hearing them spoken in isolation. Many of a child's first 50 spoken words are in fact those that have been heard in isolation, and the more often a word occurs in isolation, the more likely it is that this word will occur among those first 50. Further, based on extended naturalistic observation, caretakers appear to produce between about 6 and 60 isolated words per hour of interaction with small infants. Thus, it is at least plausible to think that at least some words enter the child's vocabulary because the child has heard that word spoken all by itself on a number of different occasions. However, infants appear to learn some words that they have never heard spoken in isolation, and some studies suggest that embedding a word in a fluent string of connected speech can actually help infants to learn that word, rather than hinder them (Fernald and Hurtado, 2006; Saffran, 2001). Because infants learn some words that they hear only as part of longer utterances, either they have an innate ability to identify individual words heard among other words in fluent speech or they must develop this ability via experience.

If you have ever listened to an unfamiliar foreign language, you will have an idea of how difficult it is to identify individual words in fluent streams of speech. Figure 9.1 shows that the places where people perceive boundaries between words do not correspond to silent parts of the speech signal. Speech does not have the equivalent of the white space in between words that helps us identify individual words when we read, so babies cannot rely on silence marking the beginnings and ends of words.

Natural speech therefore presents infants (and anyone else unfamiliar with the language) with the *segmentation problem*: The message consists of collections of words, but the speech signal does not provide obvious cues as to where one word ends and the next one starts. Before an infant can start to learn words, that is, to identify collections of sounds that make up a word and then associate meanings with those collections of sounds, the infant must *segment* the stream of speech, mentally chopping it up into word-sized chunks.

Experimental evidence suggests that the ability to segment speech into word-sized chunks does not appear until the infant is 6–7½ months old (Bortfeld et al., 2005; Jusczyk and Aslin, 1995; Jusczyk and Hohne, 1997). Evidence for young infants' segmentation ability comes from the *conditioned head turn* procedure. In this procedure, a baby is seated on a caregiver's lap in front of an apparatus that has a central light and two additional lights, one on the baby's left and one of the baby's right. The central light is flashed to gain the baby's attention, and then one of the side lights is flashed. When the baby looks at the side light, sound is played through a speaker located adjacent to the side light. When the baby looks away from the side light for more than a couple of seconds, the sound is turned off and the trial is terminated.

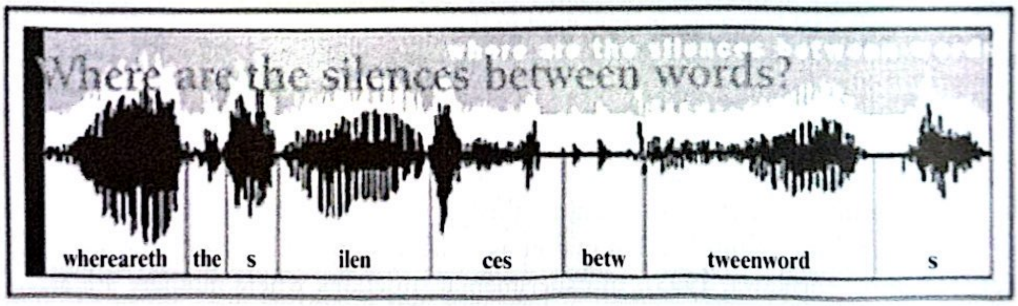


Figure 9.1 Sonogram of the question *Where are the silences between words*? Note that there are no silent gaps between words in the signal. *Source:* Saffran (2003), Association for Psychological Science

Experimenters can use this procedure to determine whether babies learn anything from prior exposure to a stimulus. During a training period, infants are exposed to some stimulus, such as an isolated word. Then, during the test period infants hear the familiar word or a novel, unfamiliar control word. If the baby can remember the familiar word, and if they can tell the difference between the familiar and the novel, unfamiliar word, then they should spend more time looking and listening when the familiar word is played than when the unfamiliar word is played. Young infants often display a preference for the familiar stimulus, looking and listening longer in that condition, but sometimes they show an opposite preference for novel stimuli. At 6 months of age, infants who are familiarized with a word by being exposed to that word spoken in isolation (by itself) do *not* listen longer to short sentences containing the familiar word. However, by 7½ months, infants who undergo the same training procedure *do* listen longer to short sentences that contain familiar words than sentences that contain unfamiliar words.[5] Result such as these show that 7½-month-old infants do remember something about the familiar word and that they are able to recognize that familiar word when it is part of a continuous stream of fluent speech. Thus, 7-½-month-olds have begun to solve the segmentation problem.

The fact that older infants but not younger infants show evidence of segmentation ability suggests that this ability is not innate, and is instead built from the infants' experience listening to the native language. If segmentation ability is not innate, there must be some pre-existing abilities (*precursors*) that the infant capitalizes on to develop segmentation ability. Researchers have identified two major classes of precursors that may provide the tools that infants need to solve the segmentation problem: prosodic cues and *phonotactic knowledge* (phonotactic knowledge refers to the patterns of phonemes that occur in the language, see later in the chapter) (Cutler and Norris, 1988; Gerken and Aslin, 2005; Jusczyk, 1997; Morgan and Demuth, 1996).

The *prosodic bootstrapping* hypothesis proposes that infants pay attention to prosodic features of their native language, and that they use those features to identify candidate words.[6] Prosody can support segmentation of the speech signal because prosodic features correlate with word boundaries. Although this correspondence is not perfect, it may be consistent enough for infants to start identifying candidate words from the speech signal.

Prosody is plausible as the entry point to segmentation because newborn infants can detect the difference between native and non-native utterances on the basis of prosodic differences, and because infants as young as 2 months old can detect differences between prosodic patterns, even when the phonological content of two utterances is identical, or nearly so (Hohne and Jusczyk, 1994). In the study, 2-month-old infants listened to someone saying *nitrates* or *night rates* (if you say those two things, and listen carefully, you may be able to hear slight differences in the way you pronounce them), which have the same phonological content but different prosodic qualities. When infants habituated to (or got bored with) one of the utterances, they dishabituated when the other one was played.

Although young infants are sensitive to differences in prosody between different utterances, it takes them some time to learn about some of the basic prosodic patterns that occur frequently in their native language. For example, about 90% of the *bisyllabic* (two-syllable) words in English have a *trochaic stress pattern* (Jusczyk et al., 1999). In trochaic stress, the first syllable is spoken a little bit louder than the second syllable. The English words *cookie*, *baby*, and *bottle* all have trochaic stress (try pronouncing them with the second syllable louder than the first—that will sound strange). Some English words have *iambic stress*, where the second syllable is louder than the first. *Guitar*, *debate*, and *pursuit* all have iambic stress. If babies pay attention to stress, and if they assume that a stressed syllable is an important unit in the language, they will be able to identify the beginnings of many words in the language by assuming that a stressed syllable is the beginning (or *onset*) of a word.[7] Researchers have labeled this version of prosodic bootstrapping the *metrical segmentation strategy*.

Young infants may not be sensitive at all to differences in stress patterns between different utterances (Weber et al., 2004). When 4-month-old infants were presented with a speech stimulus consisting of a series of iambic words, with an occasional trochaic word inserted into the list, their brain wave activity was the same for words with different stress patterns. But older infants (5 months old) appeared to detect the different stress patterns. Their brain waves changed when they heard a trochaic (strong–weak stress) word that was preceded and followed by several iambic words.[8]

Sensitivity to stress patterns emerges over time in young infants, and is not part of the infants' innate package of language-learning tools. This makes sense if you believe that innate language skills must be broad enough to enable infants to learn any language that they are exposed to. Different languages have different stress patterns. The iambic stress pattern is more common in French, for example, than the trochaic stress pattern. Thus, French babies would need to deploy the opposite strategy and assume that *un*stressed syllables are word onsets to successfully apply the prosodic stress hypothesis. Infants may be born with a predisposition to listen to speech and pay attention to prosodic features in speech, but it would be counterproductive for them to be born with a specific version of the metrical segmentation strategy prewired. French babies would be disadvantaged if they were innately wired to assume that stressed syllables appear at the beginnings of words. English babies would be disadvantaged if they were innately wired to assume that words begin with unstressed syllables. The most likely theory, then, is that infants take some time to determine the dominant stress pattern in their native language before they start relying on stress patterns to form hypotheses about where words begin.

By 7½ months of age, infants who hear short sentences containing words with trochaic stress recognize those trochaic words when they are later tested using the conditioned head-turn procedure (Jusczyk et al., 1999). Thus, 7½-month-old infants appear to use the dominant stress pattern in English (strong–weak) to hypothesize about where words begin. Infants appear to apply the metrical segmentation strategy even to samples of languages that they have not been exposed to before. For example, Houston et al. (2000) showed that 9-month-old infants from English-speaking families used the metrical segmentation strategy to identify candidate words in Dutch sentences. This is possible because Dutch has the same predominant trochaic stress pattern that English has. Having picked up the prevailing stress pattern in their native language, 9-month-olds behave as though they believe that stress pattern will apply to all kinds of speech, even if they are unfamiliar with the details of the new language.

Of course, the metrical segmentation strategy will not always succeed because many words that infants hear will have the opposite stress pattern. If a two-syllable word starts with a weak syllable instead of a strong one, then infants who apply the "strong-syllable-equals-word-onset" strategy will mis-segment speech that contains iambic words. So, for example, if an infant hears *My guiTAR is out of tune* (where TAR indicates a stressed syllable), that infant may extract the sequence *taris* and try to treat that like a word. In fact, experiments on 7½-month-old infants suggest that they make exactly this kind of mistake when they listen to short sentences containing iambic (weak-strong) words. They treat *TARis* as though it were a word. However, when 10½-month-old infants were tested, they were able to recognize iambic words, such as *guitar*, after listening to short sentences containing iambic words. These results suggest that younger infants rely more heavily on the metrical segmentation strategy than older infants do, and that older infants rely on cues other than stress patterns to segment iambic words (otherwise, they would continue to treat *taris* as a word). What information might cause infants to shift away from exclusively relying on the metrical segmentation (strong = initial) strategy?

Two sources of information should be available to 10-month-old infants. First, they will have begun to build up a vocabulary based on the metrical segmentation strategy. They can use familiar trochaic words as "landmarks" to identify segments of speech for further analysis. So, if an iambic word is sandwiched between words that have stressed syllables (e.g. *JENny's guiTAR is LOVEly*), the metrical segmentation strategy is not going to come up with a neat way to carve up the utterance, and infants may start looking for an alternative. In fact infants as young as 6 months old can identify words they have never heard before when those words appear immediately after a word that the child already knows. If the infant knows the word *mommy* (as many 6-month-olds do), and if they hear the utterance *mommy's cup* (where *cup* is a word that the child does not yet know), they will recognize the word *cup* as being familiar in a subsequent testing phase (Bortfeld et al., 2005; see also Tincoff and Jusczyk, 1999).

Older infants may begin to incorporate *phonotactic information* in their search for word boundaries in their search for a way to supplement the metrical segmentation strategy. Phonotactic information refers to the patterns or combinations of phonemes that occur in different parts of words in a given language. For example, the phonotactic properties of English prevent a word from starting with the consonant cluster /gd/. But that kind of cluster is fine in Polish (e.g. *Gdansk*). Likewise, English words cannot end with /spl/, but they can begin that way (as in *splatter, splendid*, and *split*). Infants who notice where different combinations of phonemes occur in utterances can develop phonotactic knowledge that can help them segment speech.

For example, a child who knows that words can start with /spl/ but cannot end with /spl/ will have an advantage when segmenting the utterance *This place is dirty*. Instead of treating it as *Thispl ace is dirty* (because no words end in /spl/), they are more likely to hypothesize a boundary between the /s/ and the /pl/.[9] Likewise, a child would not segment *bigdog* as *bi gdog* or as *bigd og* because the sequence /gd/ does not occur at either the beginning or the end of any English words.

Sensitivity to the phonotactic properties of the native language appears to emerge between 7 and 9 months of age. For example, younger infants do not have a preference for listening to native language speech, but by 9 months of age, children do prefer to hear their native language instead of a prosodically similar foreign language. English and Dutch have similar prosodic systems, and Dutch babies do not care whether they are hearing Dutch or English until they get to be about 9 months old, at which point they prefer Dutch (Jusczyk et al., 1993). Prosodic features of Dutch cannot account for this preference (because they are very similar to English), so the preference that emerges at 9 months most likely reflects details of the phonological/phonotactic system of Dutch. Further evidence that 9-month-old infants know about their native-language phonotactic system comes from experiments showing that children prefer to listen to pseudowords (fake words) whose phoneme sequences occur often in their native language rather than pseudowords whose phoneme sequences occur rarely in their native language (Jusczyk et al., 1994). Younger infants did not show a preference either way, suggesting that knowledge of the native-language phonotactic system is not fully developed until somewhat later.

Phonotactic knowledge can support infants' ability to segment speech and identify individual words. But if infants do not know where the words begin and end, how do they ever learn that particular sequences of sounds are more common at the beginning of the word than the end, or that particular sequences of sounds are not possible at the beginning of a word? Where does phonotactic knowledge come from? One possibility is that infants begin to learn the phonotactic system by paying attention to the beginnings and ends of entire utterances. The beginning of an entire utterance has to be the beginning of a word, and the end of an utterance has to be the end of a word. By attending to the front and back end of utterances, infants can learn about the way phonemes are distributed in different parts of words.

When mathematical modeling is used to simulate the development of word segmentation ability, the models "learn"[10] much more quickly when they are told where utterances in the training set begin and end (Christiansen et al., 1998). Infants may also use prosodic stress to learn about syllable onsets and offsets, by giving special attention to stressed syllables. Infants who are just learning to speak tend to reproduce accented (stressed) rather than unaccented (or unstressed) syllables, independent of the semantic content those syllables convey. In languages where root morphemes are stressed, children produce root morphemes before they produce other types. In languages where derivational or inflectional morphemes are stressed, children tend to produce those first (Pye, 1983).

Infant-directed speech

If you spend any time around infants, you will notice that when a baby shows up, adults turn into blithering idiots.[11] Babies have that effect on us. The pitch of our voice increases. We speak in shorter sentences. We speak more clearly and distinctly, and we vary our pitch and our loudness much more than we do when we speak to adults. This collection of strange speech properties goes by the name *infant-directed speech* (IDS), *child-directed speech*, or *motherese*,[12] and it is the object of study of very serious, tough-minded language scientists.

Why do adults use this special style of speaking when they address infants? One reason is that babies like it. Newborn infants prefer to hear the sound of a female voice speaking IDS over the same voice speaking adult-directed speech, and it helps infants stay in a good mood (Cooper and Aslin, 1990, 1994; Werker and McLeod, 1989). Beyond mood effects, IDS may help infants solve the segmentation problem. Because IDS has exaggerated prosodic features, it may provide clearer indications of important boundaries between words, phrases, and clauses. Further, IDS utterances are relatively short, which lightens the memory load that utterances impose on infants. Critical-topic words also tend to appear in highly prominent positions within IDS utterances, often at the end, and topic words tend to be marked with special prosodic features (Fernald and Mazzie, 1991). IDS utterances also engage infants' attention, boosting even further the salience of speech stimuli that infants appear to have an innate drive to attend to.

Although not all cultures have IDS, Quiche Mayan appears to be one that does not (Pye, 1983), infants appear to benefit when adults use IDS while speaking to them (Kalashnikova et al., 2018; Liu et al., 2003; Thiessen et al., 2005). Infants aged 6 months to 1 year whose mothers spoke more clearly, as measured by physical differences between different vowel sounds, demonstrated superior ability to discriminate between similar sounding words. Thus, exposure to more clearly enunciated IDS appears to instill in infants the phonological contrasts that are important in their native language.

In one direct test of segmentation skill, infants (6–8 months old) exposed to IDS outperformed infants who were exposed to adult-directed versions of the exact same test materials (Thiessen et al., 2005). In addition, studies of depressed mothers show that their infants lag behind their peers in speech-processing ability—at least through the early stages of speech segmentation and word learning (Kaplan et al., 2002). The core problem in this instance appears to be that depressed mothers' speech does not contain the prosodic features that help support the infants' positive affect, mark important boundaries in the speech stream, and emphasize key elements. While infants of depressed mothers have difficulty learning new words when listening to their own mothers' speech, these same infants perform at about the same level as their peers (infants of non-depressed mothers) when they listen to another infant's non-depressed mother speak. Thus, the problem is not that infants of depressed mothers cannot learn, the problem is that the input they are getting lacks the IDS-related cues that could support and facilitate their learning.[13] To sum up, research on IDS indicates that, although IDS may not be necessary for children to learn their first language, it helps them master some aspects of speech comprehension and word learning, and it certainly does no harm. And it's fun.

PRACTICAL ADVICE CORNER

You may become a parent one day. If you don't become a parent yourself, you may know someone who becomes a parent. If so, you might be interested in whether parents' behavior influences a child's language development. The short answer is that, yes, parents have a great deal of influence on the course of a child's language development (Fernald et al., 2006, 2001; Hurtado et al., 2007; Marchman and Fernald, 2008; Pan et al., 2005; Swingley et al., 1999; Tsao et al., 2004). Very young infants treat words as perceptual wholes, and they sometimes fail to discriminate between two similar sounding words. However, by the age of 24 months most infants have become very efficient and very accurate at recognizing familiar words. Like adults, 24-month-old children process familiar words incrementally. That is, they can recognize a familiar word before the entire word has been spoken. However, some babies are better at this than others. What makes the difference is the size of the baby's vocabulary and the speed with which the baby can process incoming speech. Further, these two abilities are in a symbiotic relationship of sorts: The more words the child knows, the faster that child can recognize familiar words; the faster the child can recognize familiar words, the more resources that child has left over to pay attention to and learn from other parts of the speech signal. Further, these differences have long-term, cumulative effects. Children who are faster at recognizing familiar words at 6 months old have bigger vocabularies at 24 months old than children who are slower at recognizing familiar words at 6 months old. Speed of speech perception predicts long-term growth of both comprehension and production. This relationship holds in the United States both for middle- and upper-middle-class children learning English and for lower- and lower-middle-class children learning Spanish. Differences between fast and slow processors can be explained largely as a kind of practice effect. The more language infants hear, the more efficient they get at recognizing familiar words; and the more efficient they become at recognizing familiar words, the more they can learn about new words. Differences in processing efficiency between different infants can also be attributed in part to how much adults talk to them, and there are vast differences between and within different socioeconomic groups in the amount of speech that gets directed toward infants. Maternal education also appears to matter a great deal. Independent of socioeconomic status, mothers with more formal education speak more to their children than mothers with less formal education. Bottom line: Talk to babies. They will thank you later.

Statistical Learning and Speech Segmentation

Infants may make use of prosodic features to identify boundaries between words in fluent speech, but an alternative, the *statistical learning* approach, suggests that speech segmentation skill has little or nothing to do with the metrical segmentation strategy or other forms of prosodic bootstrapping (Saffran, 2001, 2003; Saffran et al., 1996; Thiessen and Saffran, 2003; see McMurray and Hollich, 2009; Smith et al., 2018). According to the statistical learning approach, infants notice patterns in complex stimuli and use those patterns to analyze speech stimuli to identify important subcomponents, including words. As Saffran (2003, p. 110) notes, "Infants can rapidly capitalize on the statistical properties of their language environments, including the distributions of sounds in words … to discover important components of language structure." How might this work for learning words?

Saffran and colleagues (e.g. Saffran, 2001, 2003; Saffran et al., 1996) note that, in English, some syllables are highly predictable from other syllables. Consider the word *pretty*. It has two syllables: "pri" (/prI/ in the English phonetic alphabet) and "tee" (/ti/). In IDS, /ti/ follows /prI/ about 80% of the time (it's not clear what percentage of the time /ti/ is preceded by /prI/, but if a baby hears /prI/, it will very likely hear /ti/ right away).[14] The syllable /ti/, however, provides very little information about what syllable will come next. Just about anything could come after *pretty* (*baby, doggie, mommy, apple, …*), so there is no way the baby can predict what comes after /ti/. Notice that there's a clear difference in probabilities here that an infant could exploit to help figure out how the language works. If the likelihood of /ti/ given /pri/ is very high, maybe those two syllables should be considered as a package.

Think of it this way: If you have a male friend, and every time you see your male friend, he is with the same female, you might soon deduce that those two are boyfriend and girlfriend, and you might start to think of them as a couple (kind of like Brangelina).[15] If, on the other hand, you have two other friends, but you never see them together at the same time, you would not assume that they are a couple. So, if some pairs of syllables co-occur more than others, an infant might start to think that the co-occurring syllables make up a word (e.g. *pretty*), and syllables that do not co-occur, or do so rarely, do not go together to make a word. If so, babies could use co-occurrence information to segment the speech stream, identifying words by treating high-probability pairs of syllables as words and low-probability pairs as separate.

Evidence that babies are capable of detecting patterns in syllables comes from experiments where babies listen to an artificial "mini-language" during a training phase, and get tested on what they learned later on (Saffran et al., 1996; Saffran, 2001, 2002; see Mirman et al., 2008 for similar experiments on adults). The mini-languages are made up of sets of nonsense syllables, such as *jik, pel*, and *rud*. The languages are designed such that some syllables always occur immediately after others. If infants hear *jik*, they always hear *pel* next. Other combinations of syllables occur much less frequently. For example, *rud* might appear immediately after *pel* a third of the time, but the other two thirds of the time different syllables, such as *mib* and *lum*, would follow *pel*. This set of rules is used to create long strings of syllables, such as *jik pel rud neb jik pel mib vot loke hep jik pel lum*. A synthesizer, rather than a human speaker, generates the strings of syllables so *no* prosodic cues indicate where any word boundaries might be. So, the only information that the infants have to work with is the *transitional probability* between different syllables—the likelihood that one syllable will be followed by another.

When infants are tested using a version of the conditioned head-turn procedure, they listen for different amounts of time to pairs of syllables that are "words" in the language than pairs of syllables that have lower transitional probabilities (and therefore are not "words" in the language). The results are about the same when infants are exposed to real words from an unfamiliar real language. English-learning babies recognize Italian words when they are exposed to a version of mini-Italian where pairs of syllables from real words have high transitional probability, and pairs of syllables that cross word boundaries have low transitional probability (Pelucchi et al., 2009).

Infants are capable of picking up statistical information present in the speech stream and using that information to segment it, even when there are no prosodic cues such as accent, changes in pitch, or pauses to help them identify individual parts of the speech stream that could be words. Further studies indicate that older infants (17–18 months old) can learn about phonotactic patterns by paying attention to statistical cues (Chambers et al., 2003). They can also learn to associate high-probability pairs of syllables with pictures of unfamiliar objects (Graf-Estes et al., 2007), which suggests that infants are treating the segments that they identify as being word-like, rather than just as interesting snippets of sound. Further, statistical learning may be the result of a general-purpose learning mechanism in children (Hamrick et al., 2018; Monroy et al., 2019; Smith et al., 2018). They can learn simple correspondences between visual stimuli (e.g. pictures of dogs), as in Figure 9.2. When children are exposed to sets of pictures based on the ABA pattern (e.g. Husky, German Shepherd, Husky), they spent more time

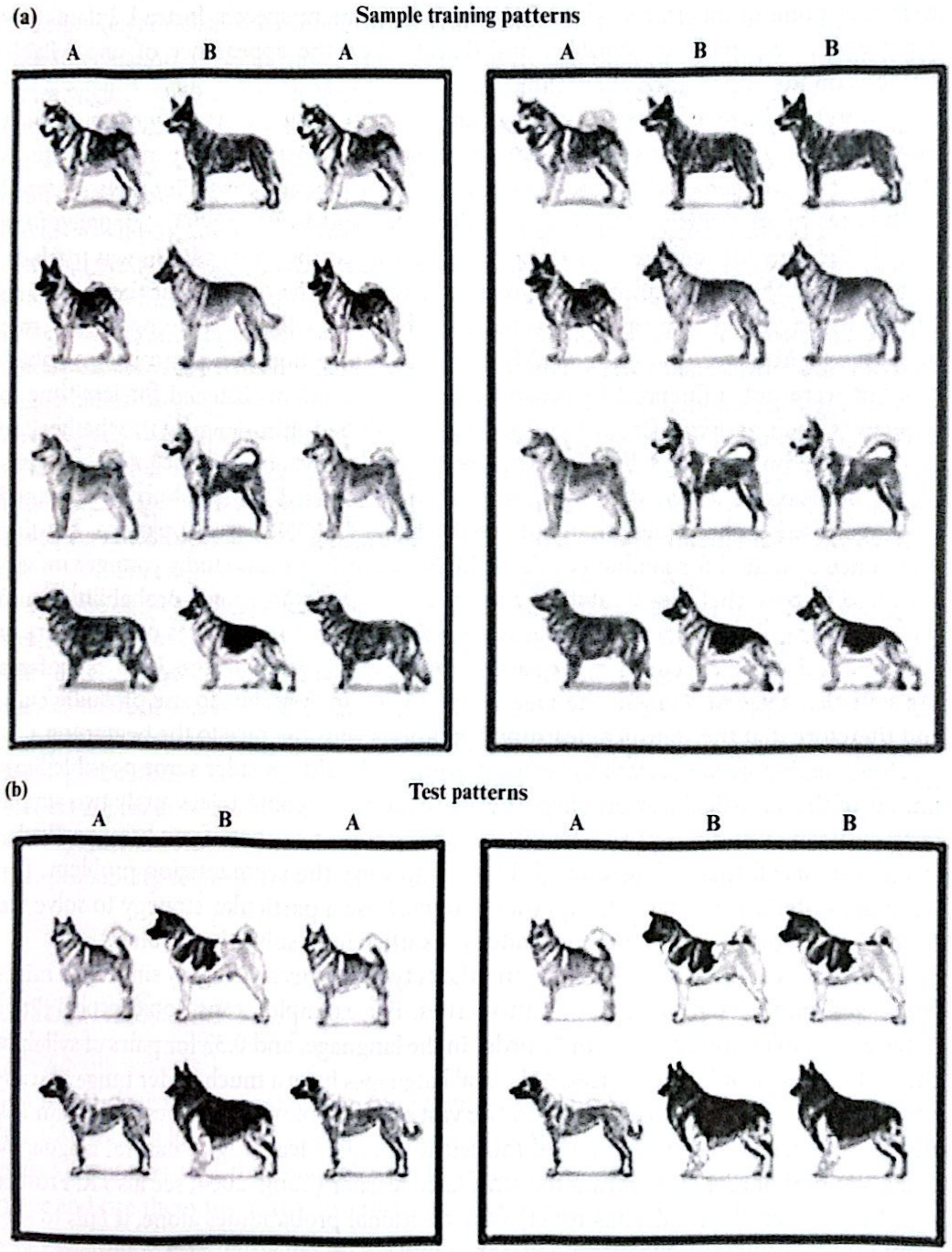


Figure 9.2 Patterns of dogs used to train and test 7-month-old infants. *Source:* Saffran et al. (2007), with permission from Elsevier

looking at a novel AAB pattern (e.g. Dachshund, Dachshund, Labrador). The existence of statistical learning for both language and nonlanguage visual stimuli could indicate that general-purpose learning mechanisms, rather than innate, genetically determined, language-specific learning mechanisms, are responsible for aspects of speech segmentation and word learning.

Language scientists seeking to explain how infants "break into" the speech stream want to know which cue or cues infants use first to begin solving the segmentation problem. The metrical-segmentation-strategy hypothesis proposes that children can begin to segment words without hearing them in isolation, and without knowing anything about transitional probabilities. The isolated-word hypothesis proposes that children begin to segment speech by first learning about isolated words, then by recognizing previously learned words in fluent speech streams, using known words to detect the beginnings and ends of words that surround the familiar words. The statistical learning hypothesis suggests that none of this is necessary for infants to segment speech. Instead, infants pay attention to sequences of syllables and detect when the appearance of one syllable depends on the appearance of another.

Researchers have been testing younger and younger infants to try to figure out which cues babies rely on at the very beginning of speech segmentation. In one such study, researchers provided both prosodic cues and transitional probability cues to word boundaries in an artificial mini-language (Thiessen and Saffran, 2003). In some of the stimuli, the prosodic cues mirrored those in English (i.e. the stress pattern was trochaic, strong–weak). In other stimuli, the prosodic cues were the opposite of the prevailing pattern in English. The transition probabilities between syllables were the same across the two sets. Younger (6–7-month-old) infants paid attention to transitional probabilities, but were not influenced by prosodic cues. These infants listened for less time to sequences such as *jikpel* (from the previously described mini-language), whether the emphasis fell on the first syllable (*JIKpel*) or the second syllable (*jikPEL*). Older infants (8–9 months old) showed a different pattern. They preferred familiar words (as defined by transitional probabilities) that had the trochaic (STRONG–weak) pattern, but that preference switched for iambic (weak–STRONG) words. In this study, younger infants appeared to base their segmentation choices entirely on transitional probabilities, and not prosodic information. Older infants appeared to base their choices on a mixture of prosodic and statistical cues. On the basis of these results, we might conclude that infants use statistical cues to segment the speech stream before they start to use prosodic cues, and therefore that the statistical learning hypothesis wins the race to the beginning.

However, before we declare the contest over, we should consider some possible limitations of the statistical learning hypothesis. To judge the contest, let's apply two simple tests: sufficiency and necessity. The sufficiency criterion says that, if you use a particular strategy by itself, that will be enough by itself to solve the segmentation problem. The necessity criteria is stronger. It says that you *must* use a particular strategy to solve the problem (and presupposes that the strategy is sufficient to solve the problem).

Experiments on statistical learning in infants typically involve highly simplified mini-languages with very rigid statistical properties. For example, transitional probabilities between syllables are set to 1.0 for "words" in the language, and 0.33 for pairs of syllables that cut across "word" boundaries.[16] Natural languages have a much wider range of transitional probabilities between syllables, the vast majority of which are far lower than 1.0. Researchers have used mathematical models to simulate learning of natural languages, using samples of real IDS to train the simulated learner (Yang, 2004; see also R. Frost et al., 2019). When the model has to rely on transitional probabilities alone, it fails to segment speech accurately. However, when the model makes two simple assumptions about prosody—that each word has a single stressed syllable, and that the prevailing pattern for bisyllables is trochaic (STRONG–weak)—the model is about as accurate in its segmentation decisions as 7-month-old infants. This result casts doubt on whether the statistical learning strategy is sufficient for infants to learn how to segment naturally occurring speech (and if the strategy is not sufficient, it cannot be necessary either; see Swingley, 2020 for a contrary viewpoint).

The modeling work by itself is suggestive, but what we really need are controlled experiments with infants exposed to unfamiliar speech streams that contain a range of transitional probabilities that closely match those that occur in natural speech. Such studies have not yet been conducted. Some experimental work, however, suggests that substantial exposure (10 hours) to a language that has an unfamiliar prosodic pattern does not enable infants to segment the speech (Jusczyk, 1997). Because infants do segment unfamiliar speech that has a familiar prosodic pattern (e.g. English and Dutch) after minimal exposure, but do not segment a language with an unfamiliar prosodic structure (e.g. English and Mandarin), this also suggests that prosody, rather than statistical information, is responsible. The pattern of transitional probabilities in Dutch syllables *might be* higher than in Mandarin (a possibility that has not yet been addressed in the literature), otherwise, it is tempting to think that, at least for 9–10-month-old infants, the availability of prosodic cues make a bigger difference to their word-segmentation ability than the availability of information about how likely one syllable is given another, at least when probabilistic information approximates what is available in natural speech.

Other studies suggest that statistical learning may not be necessary for segmentation to occur. (To be fair, the outcomes of statistical learning experiments suggest that other strategies may also be unnecessary for infants to segment speech, at least when transitional probabilities between syllables in words are high.) The isolated word and metrical segmentation strategies appear to be sufficient for the successful identification of at least some words in fluent speech. Further, the cross-language segmentation experiments reviewed earlier (Houston et al., 2000) indicate that infants do not need to know which syllables go together frequently in a language before they can identify words in that language. As Houston et al. (2000, p. 507) note, "The ability to extract familiar words from fluent speech is not dependent on familiarity with the phonetic structure of the input." Even some of the statistical learning experiments suggest that high transitional probabilities are not necessary for learners to identify important components of artificial mini-languages (Saffran, 2002). In some experiments, there is little difference in performance between learners who are exposed to mini-languages with high transitional probabilities and learners who are exposed to mini-languages with much lower transitional probabilities. In other experiments (e.g. Mirman et al., 2008, Experiment 1), learners appear to segment "words" composed of low-probability sequences of syllables (with transitional probabilities of 0.33) before they segment "words" composed of high-probability sequences (with transitional probabilities of 1.0).

Interim Summary

Can infants learn new words by hearing them spoken in isolation? It appears that they can. Can infants learn new words by paying attention to prosodic patterns in the language? It appears that they can. Can children detect statistical regularities in the language and use them to identify important components of the language? It appears that they can. Unless we assume babies are dumb, there is no reason to believe that they are incapable of taking advantage of whatever information speakers make available to them.

Thus, the safest bet is to conclude that infants rely on a variety of cues to solve the segmentation problem. This conclusion is reinforced by testing on real infants, and mathematical modeling of the learning process also points to the benefits of paying attention to multiple cues. When a neural network seeks to accurately identify words in fluent, child-directed speech, the network is much more accurate when it uses multiple cues (Bagou and Frauenfelder, 2018; Christiansen et al., 1998). One cue by itself may be unhelpful, but the value of that same cue can rise dramatically when it is combined with other cues. Infants, like good scientists, appear to be flexible, pragmatic learners rather than dogmatic followers of a single, narrow strategy. Babies are smart, so if a cue is available, there is a good chance that they will figure out a way to make use of that cue to help them learn.

Learning Word Meanings

Children start to produce words at about the age of 12 months ... if we stick to the more conservative estimate of 60,000, (this) equates to about 10 new words a day up until the end of high school.

PAUL BLOOM, *HOW CHILDREN LEARN THE MEANINGS OF WORDS*

Children are highly efficient word learners. As Paul Bloom notes, "They achieve this feat without any explicit training or feedback" (Bloom, 2000, p. 26). Segmenting fluent speech into word-sized chunks of sound gets infants going on the road to mastering their first language, but before they can start to communicate effectively, they need to associate concepts with packages of sound (words) that they pull out of the speech stream. Children as old as 14 months who are still trying to master the phonological system of their native language have difficulty associating sounds and meanings, but children get better at learning words the longer they spend learning the language and the more skill they develop at discriminating between similar speech sounds (Fisher et al., 2001; Mills et al., 2004; Stager and Werker, 1997). It takes most infants about 18 months to learn their first 50 words (as measured by how many different words they say in daily life), but after that children experience a *word spurt*, during which time the rate at which children learn new word meanings increases dramatically (McMurray, 2007). What is most amazing about children's word-learning abilities is that they are able to deduce new word meanings simply by hearing the word used a couple of times. Older children can deduce a word's meaning after hearing it used only once.

To accomplish these amazing feats of deduction, infants must overcome a set of serious challenges. The *poverty of the stimulus* problem is one big obstacle (Brown, 1957; Quine, 1960). *Poverty of the stimulus* as it applies to word learning refers to the fact that the environment does not provide the information that the child needs to identify one and only one meaning for a word. Quine asks you to imagine traveling to a distant, foreign land that has rabbits in it. One morning, you and your local guide encounter a rabbit running across the path. You point at the rabbit and say *What's that?* The local guide says, *gavagai*. You might *assume* that *gavagai* means "rabbit," but your assumption could be wrong. The guide might really be saying "furry," or "long ears," or "mammal," or "lagomorph," or "above ground," or "running," or "anything with its back toward us," or "not a snake," or "that sucker's fast," or "tastes great," or "less filling." Maybe the guide is just saying, "What?" Briefly, there is nothing in the environment that rules in or rules out any particular interpretation of the utterance *gavagai*. This issue is highlighted in an urban legend (probably false) that describes the way the word *kangaroo* entered the English language. According to this legend, some of the first English people traveling in the Australian outback encountered kangaroos and asked their aboriginal guides *What's that?* The guides said *kangaroo*, which in their language means "I don't know."

So, what is the poor infant to do? Folk theories of language acquisition, and also behaviorist theories, assume that children learn word meanings by playing a language game, sometimes called *point and say* (Bloom, 2000; Clark, 2009; Skinner, 1957).[17] In the point-and-say method, the parent points at something and makes a noise, like *bunny!* The child sees the object that is being pointed at and associates the noise with the object. Indeed, nouns in young children's vocabularies tend to refer to concrete objects, like dogs, cats, and balls, that could be the object of the point-and-say game (Brown, 1957). However, as Quine's analysis shows us, pointing and speaking, by themselves, do not provide enough information for the child to deduce the speaker's intended meaning.

Before children can learn word meanings, they have to somehow narrow down the range of possible meanings. However, even if infants somehow come up with a successful strategy for object names, many of the words that they need to learn do not refer to discrete objects. For words such as *milk* and *plastic* (substances), or *thought* (a process), there is no object that someone could point to. Further, some of the earliest words that children produce are social interaction terms like *hi* and *bye-bye* that also lack corresponding objects. How do infants learn those? Finally, if point and say were the primary word-learning mechanism, we would expect blind children to be horribly disadvantaged when it comes to learning word meanings, but they learn at about the same rate as other children (Landau and Gleitman, 1985).

What about labels for actions, like verbs? The point-and-say hypothesis supposes that children learn words because they simultaneously see an object and hear a label. But people rarely label an action while they are doing it (Fisher et al., 1991; Gleitman and Gleitman, 1997). People don't say, *I'm drinking* while they are drinking. They don't say, *I'm closing the door now*, while they are closing a door. Action verbs such as *chase* and *flee* are especially tricky, because every event of *chasing* is also an example of a *fleeing* event. Every time someone chases, something flees, and vice versa (Fisher et al., 1991; Gleitman and Gleitman, 1997). If someone points to a scene while an action of chasing (and fleeing) is going on, and says *blicket*, how does the child know whether to interpret *blicket* as *chase* or *flee?* Both are going on at the same time, and the meaning of the word *blicket* depends on the point of view of the person speaking about the event, not on the point of view of the child observing the event.

Considerable research in language development has been devoted to figuring out exactly how children narrow their search for word meanings and arrive at a set of meanings that very closely match the standard meanings used by everyone who knows the language. To understand how this trick is done, we have to start by abandoning the idea that point and say is the only, or even the most important, strategy that children use to learn what words mean. What are the alternatives? One solution to the poverty of the stimulus problem proposes that children have innate categories, such as *noun* and *verb*, which they seek to populate with specific words from their native language. According to this *genetically guided learning* hypothesis, infants populate those categories with specific words by attending to salient bits of speech and trying them out as nouns or verbs (Pinker, 1984, 1994b, 1996). This approach eliminates many of the possible interpretations of *gavagai* (those involving actions or complex relationships), but it does not get the child all the way home. For example, even if the child only considers object labels as meanings, *gavagai* could mean *animal, mammal, rodent*, or *rabbit*. Or *gavagai* could be a proper name for that specific bunny, equivalent to *Flopsy* or *Peter*. To solve that problem, the child could store each instance of *gavagai*, as well as an episodic memory of the context when the word was spoken. That way, if *gavagai* occurs when the child sees a dog, cat, and horse, the *rabbit* and *Flopsy* interpretations are ruled out, and something like *animal* or *mammal* becomes more likely.

Alternatives to the genetically guided learning hypothesis appeal to general-purpose learning and memory skills. According to approaches such as these, children have a general ability to pick up and remember linguistically conveyed information (Markson and Bloom, 1997; see also Novack et al., 2021; Swingley and Fernald, 2002) rather than a special mechanism that just does word learning. Research on learning and memory shows that both children and adults appear to have better recall for arbitrary information when that information is conveyed in language than by other means. For example, children and adults in Markson and Bloom's study were exposed to a set of objects. For one of the objects, they were told either that its name was *koba* or that it was given to the experimenter by her uncle. Both of these facts about the object are arbitrary, and both are conveyed using language. A different set of subjects were shown the same set of objects, and the experimenter placed a sticker on one of the objects (but did *not* name it or say where it came from). Up to one month later, children and adults in the two linguistically conveyed knowledge conditions (the *koba* and uncle conditions) were able to remember which object the arbitrary knowledge was associated with. However, people in the sticker condition could not remember which object was associated with the sticker. This result could indicate that learning new words uses the same mental equipment as learning new arbitrary facts conveyed by language, and that learning word meanings does not depend on a special-purpose word-learning mechanism. The explanation of why, exactly, language is more effective than other methods of conveying arbitrary information requires further investigation, however.

To overcome the poverty of the stimulus, infants could begin by assuming the most general possible interpretation of a new word (i.e. a new word would be treated as though it applied to everything, kind of like the word *thing*). Alternatively, the child could assume that a new word applied only to the specific object that the child was attending to when the word was spoken (i.e. a new word would be treated as a proper name, like *Flopsy* or *Evelina*).

It appears that children use neither of these strategies. Instead, infants appear to have a pre-existing bias to treat new labels like *gavagai* as names for *basic-level* categories (Masur, 1997). When people are asked to name objects, they tend to come up with labels that are not too specific and not too general, but that are just right to discriminate between different kinds of objects and to include a wide variety of examples of the concept (Rosch and Mervis, 1975). These "just right" labels are known as *basic-level* terms, to discriminate them from *superordinate* and *subordinate level* words. If you see a picture such as one of those in Figure 9.2a or b, and you are asked to say what the picture shows, you will use a basic-level name like *dog*, rather than a subordinate-level name like *Poodle*, or a *superordinate* category label like *mammal* or *animal*.[18]

Infants (and adults) also have a pre-existing bias to treat novel words as labels for entire objects, rather than parts of the objects, substances, colors, or other features of the object (Markman and Hutchinson, 1984). So, if you and an infant are looking at Flopsy, and someone says, *Gavagai!*, both you and the infant are likely to assume that *gavagai* is a word that applies to the whole shebang (not just part of the thing), and also that the word is a basic-level term referring to the category of things that physically resemble Flopsy.

Plate 1 Patterns of neural activity in response to actual body movements (left side) and words referring to face (smile), arm (throw), and leg (walk) actions (right side). Neural activity related to face movement appears in green, finger and arms movement in red, and foot and leg movement in blue. *Source:* Hauk, Johnsrude, and Pülvermüller, 2004, p. 304.

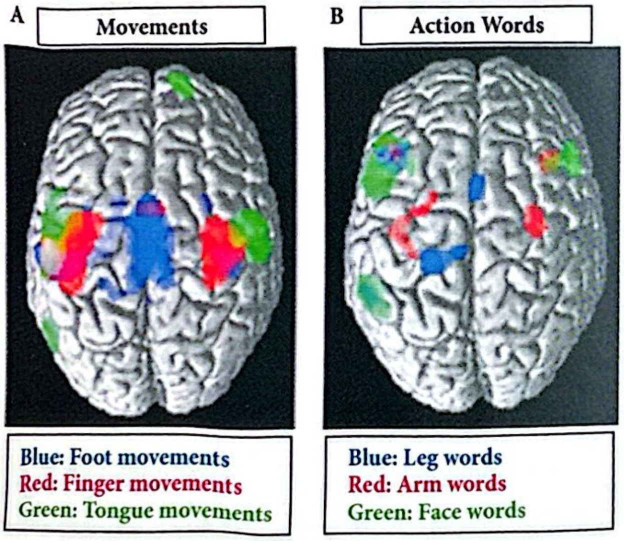


Plate 2 Voxel-based lesion–symptom mapping (VLSM) results for nonlinguistic tasks (top) and reading comprehension (bottom). Deficits in action perception are more strongly correlated with damage to frontal regions. Deficits in reading about actions are more strongly correlated with more posterior regions. *Source:* Saygin, Wilson, Dronkers, and Bates, 2004, p. 1797.

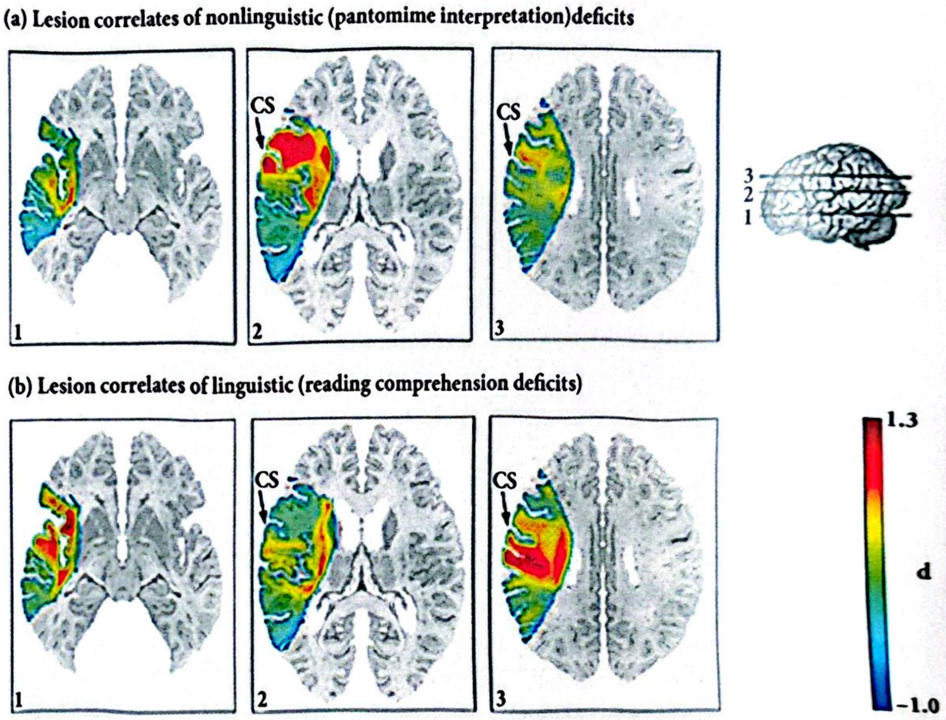


Plate 3 PET imaging data. Subjects performed similarity judgments on words or pictures. Top: common areas that were activated for both words and pictures. Middle: areas that were activated for words but not pictures. Bottom: areas that were activated for pictures but not words.

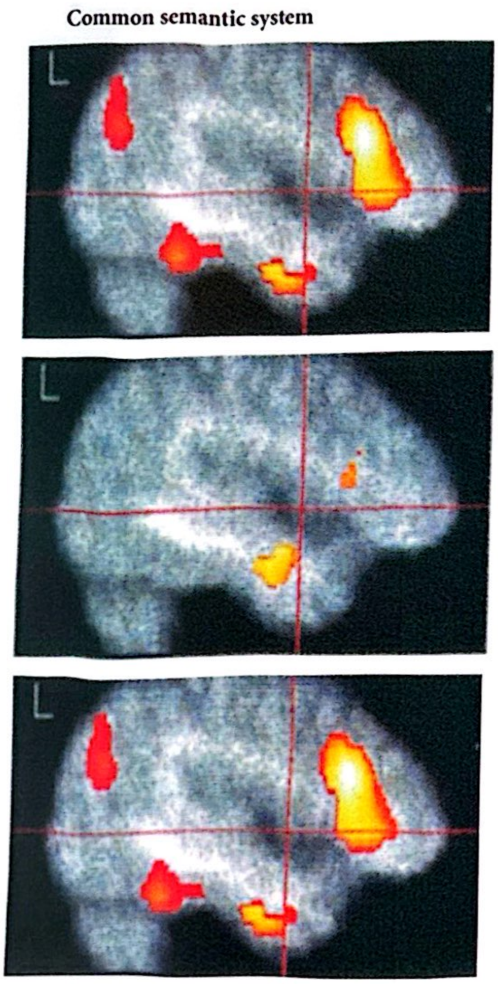


Plate 4 PET neuroimaging results. The top half shows greater occipital lobe activity during covert naming of animals versus tools. The bottom half shows greater inferior frontal lobe activity during covert naming of tools versus animals. *Source:* Martin, Wiggs, Ungerleider, and Haxby, 1996, p. 651.

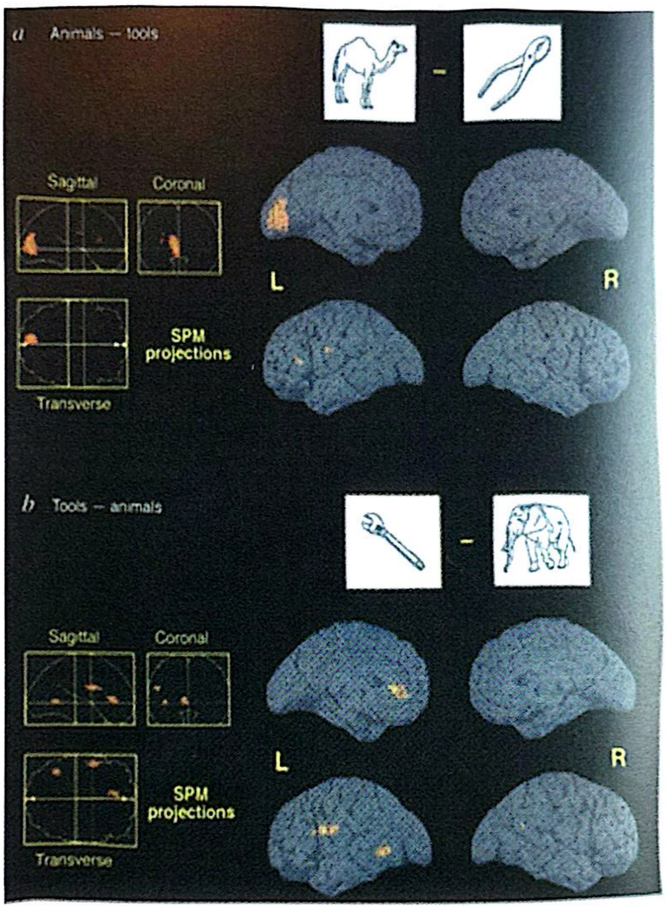


Plate 5 fMRI data showing greater left-lateralized frontal activity for questions tapping abstract versus perceptual properties of animals (shown in orange). Questions that tapped visual features of animals led to increased activity in right parietal lobe (dark blue). *Source:* Goldberg, Perfetti, Fiez, and Schneider, 2007, p. 3796.

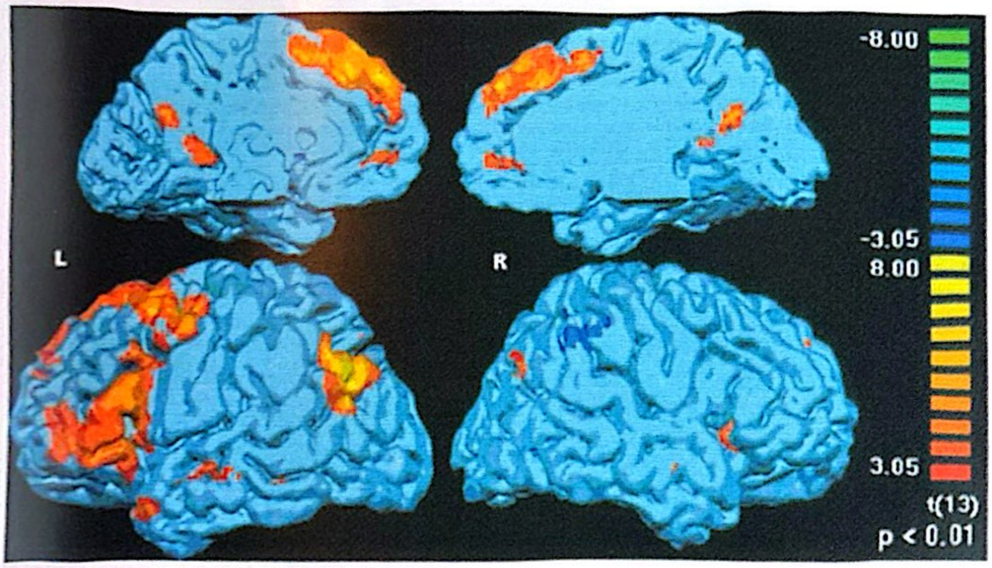


Plate 6 Lesion–performance correlations. *Source:* Damasio, Grabowski, Tranel, Hichwa, and Damasio, 1996, p. 501.

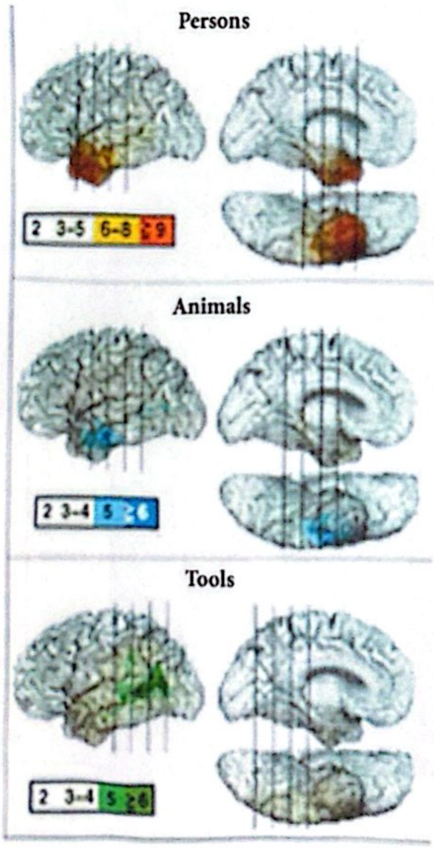
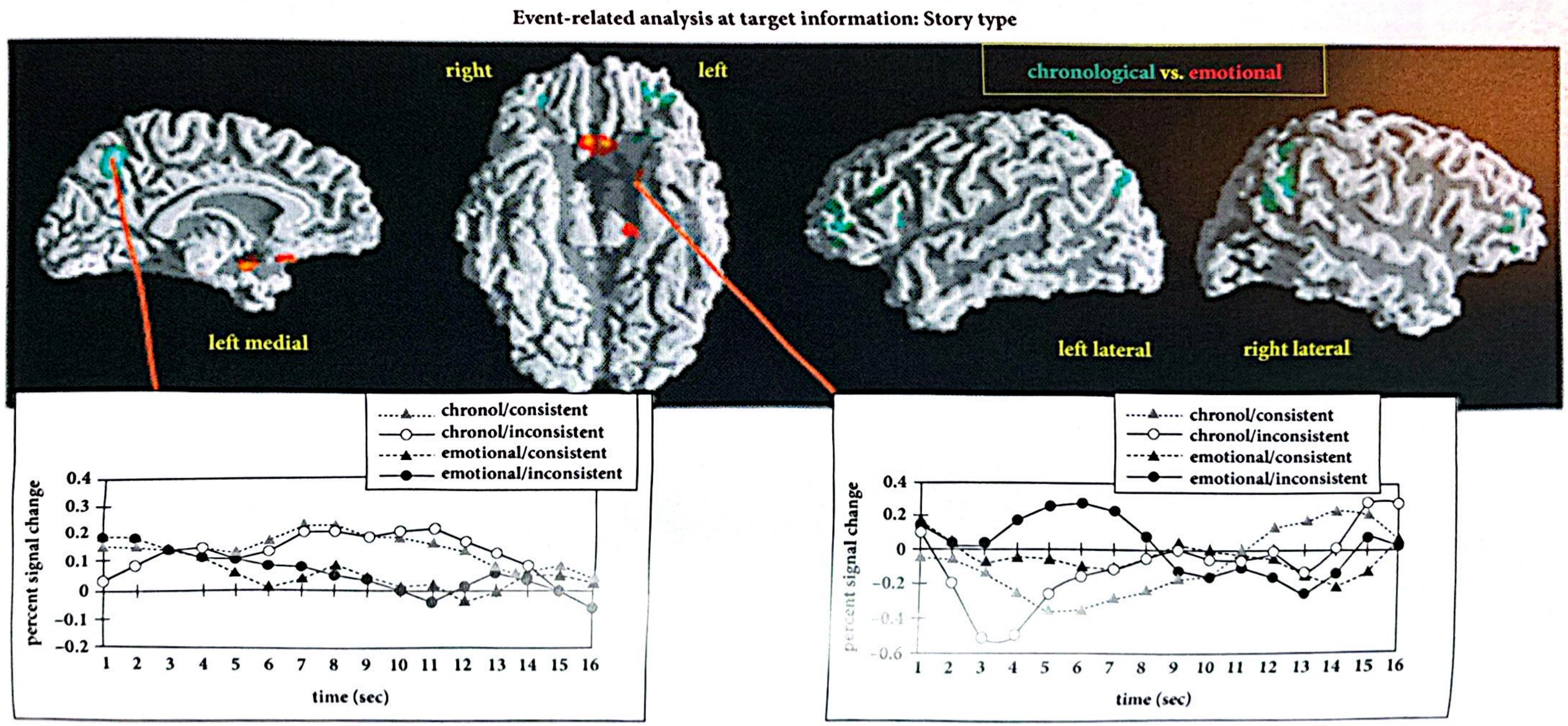


Plate 7 fMRI results comparing response to chronological (green) and emotional (yellow and red) information in stories. *Source:* Ferstl, Rinck, and Von Cramon, 2005, p. 728.



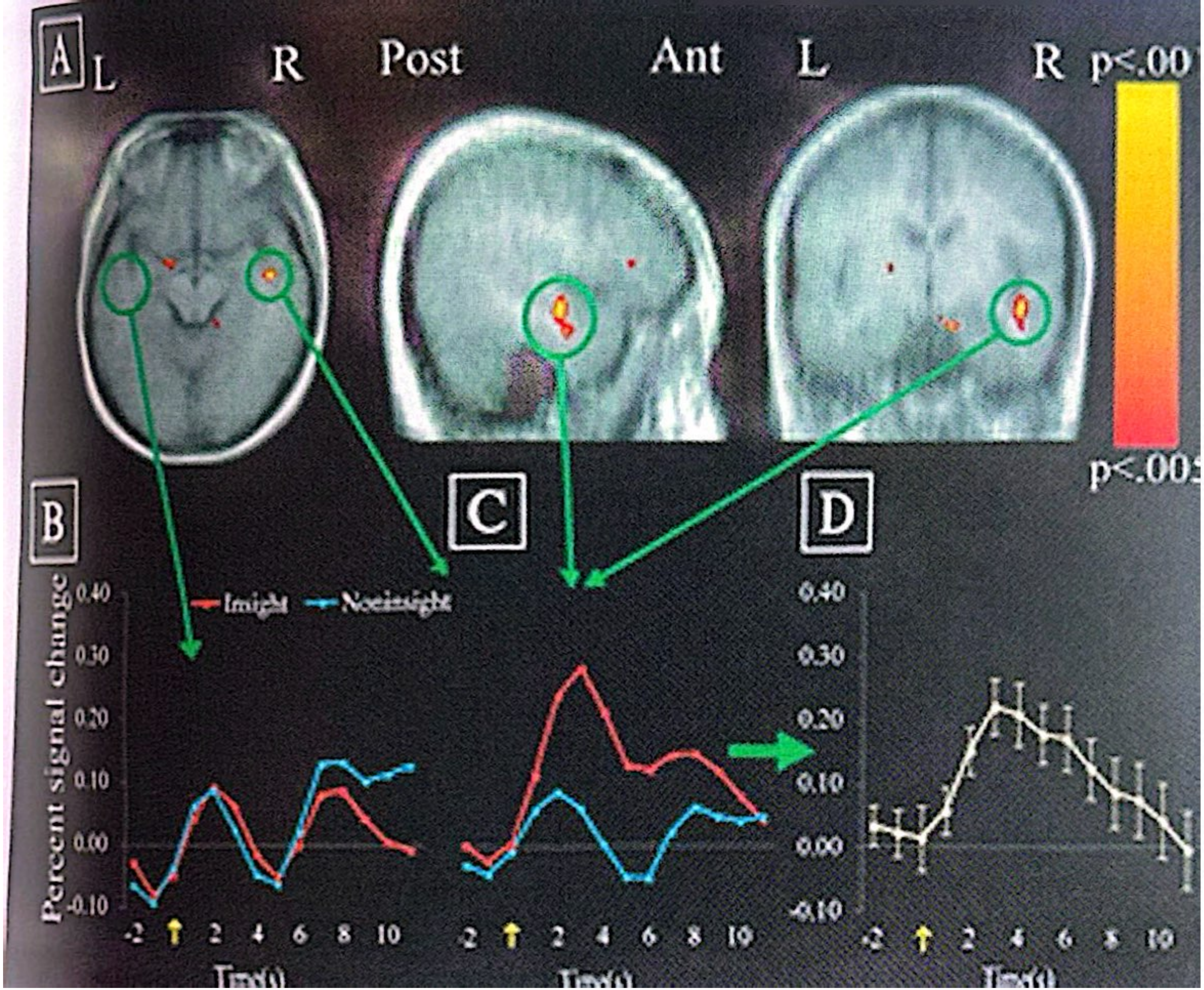

A
L
R
Post
Ant
L
R
p<.001
p<.005
B
C
D
Insight
Noninsight
Percent signal change
0.40
0.30
0.20
0.10
0.00
-0.10
-2
2
4
6
8
10
Time(s)

Plate 9 fMRI results. The areas in yellow and red, primarily in the right superior temporal lobe, showed greater activity when the text implied, rather than explicitly stated, that the focused character engaged in an action. *Source:* Virtue, Haberman, Clancy, Parrish, and Beeman, 2006, p. 107.

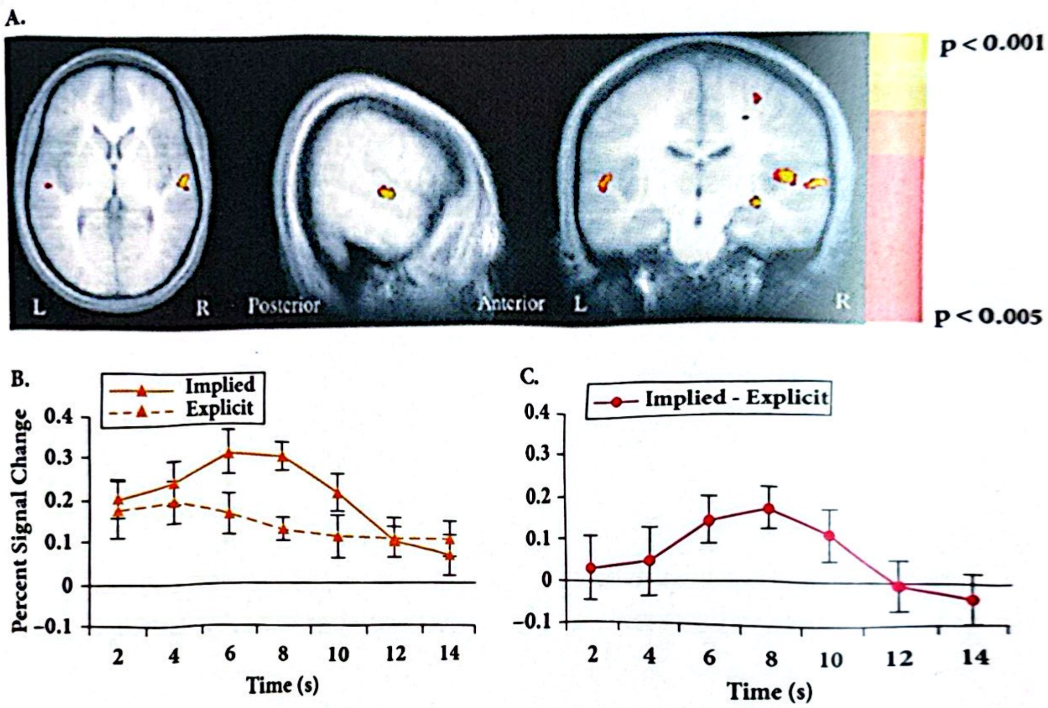


Plate 10 fMRI activation results. The *right* hemisphere is pictured on the *left* hand side of the picture. *Source:* St. George, Kutas, Martinez, and Sereno, 1999, p. 1320.

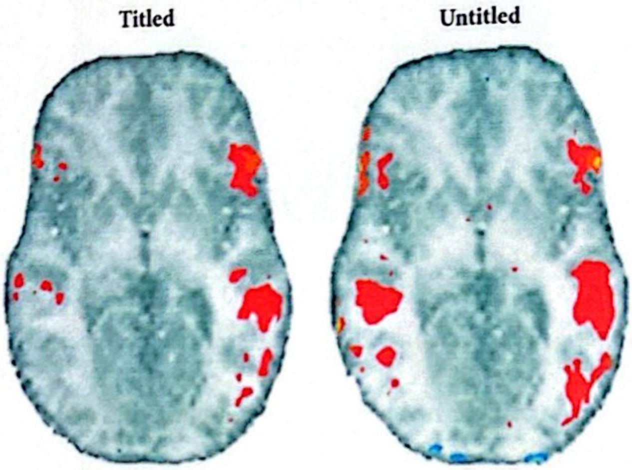


Plate 11 Whole-brain image of titled (top) vs. untitled (bottom) stories. *Source:* St. George, Kutas, Martinez, and Sereno, 1999, p. 1322.

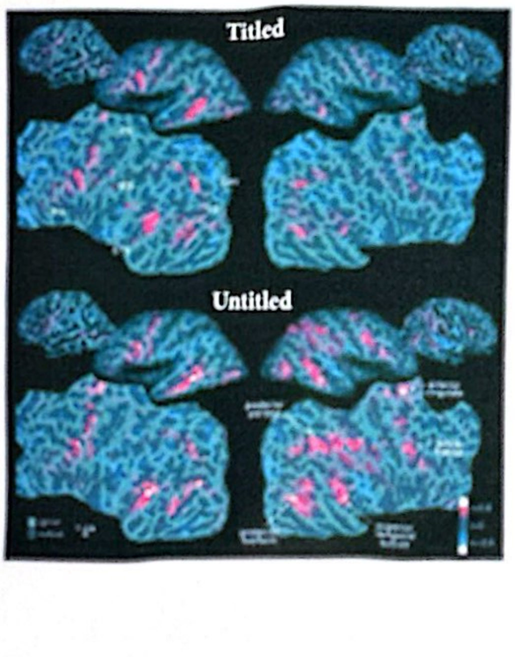


Plate 12 Right-hemisphere brain activity is greater for unrelated pairs of sentences than for closely related pairs of sentences. *Source:* Kuperberg, Lakshmanan, Caplan, and Holcomb, 2006, p. 357.

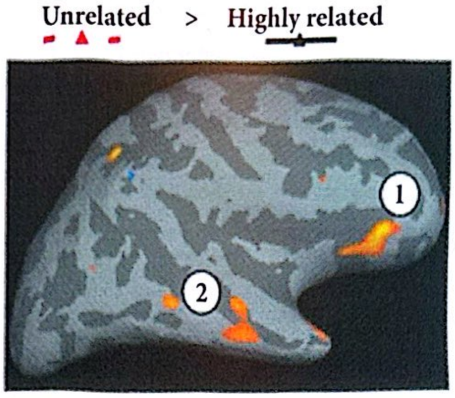


Plate 13 The brain responds differently to highly coherent versus incoherent stories. This is true of both the left hemisphere and the right hemisphere. Areas of the brain that respond more strongly to moderately coherent stories than to either highly coherent and incoherent stories are marked in yellow. *Source:* Kuperberg, Lakshmanan, Caplan, and Holcomb, 2006, p. 354.

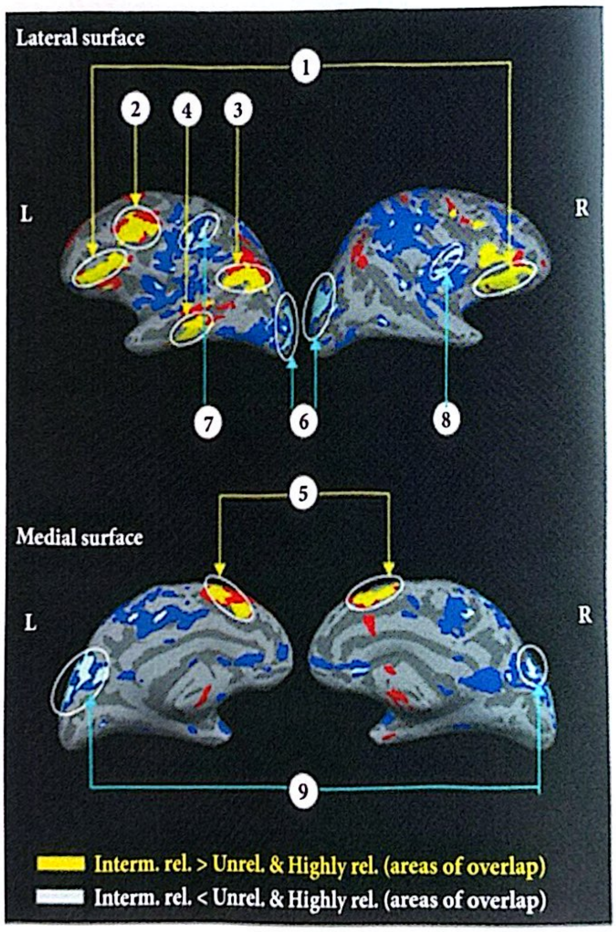


Plate 14 PET results. The scans indicate greater blood flow in right-hemisphere brain regions for metaphoric sentences compared to literal sentences. The left-hand brains show the right hemisphere. The top brains show the medial (middle) surface of the brain. The bottom brains show the lateral (outside) surfaces. *Source:* Bottini et al., 1994, p. 1246.

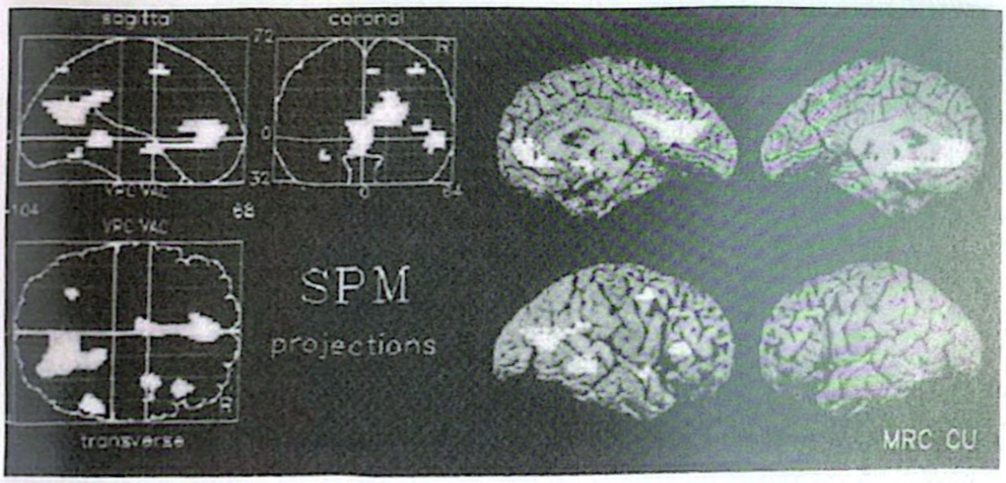


Plate 15 fMRI results. Metaphoric sentences produced greater neural response in the left hemisphere when compared with literal sentences. No differences based on sentence type (literal vs. metaphoric) were observed in the right hemisphere. (Compare Plate 16 to Plate 15. The clash of data could hardly be more extreme.) *Source:* Rapp, Leube, Erb, Grodd, and Kircher, 2004, p. 399.

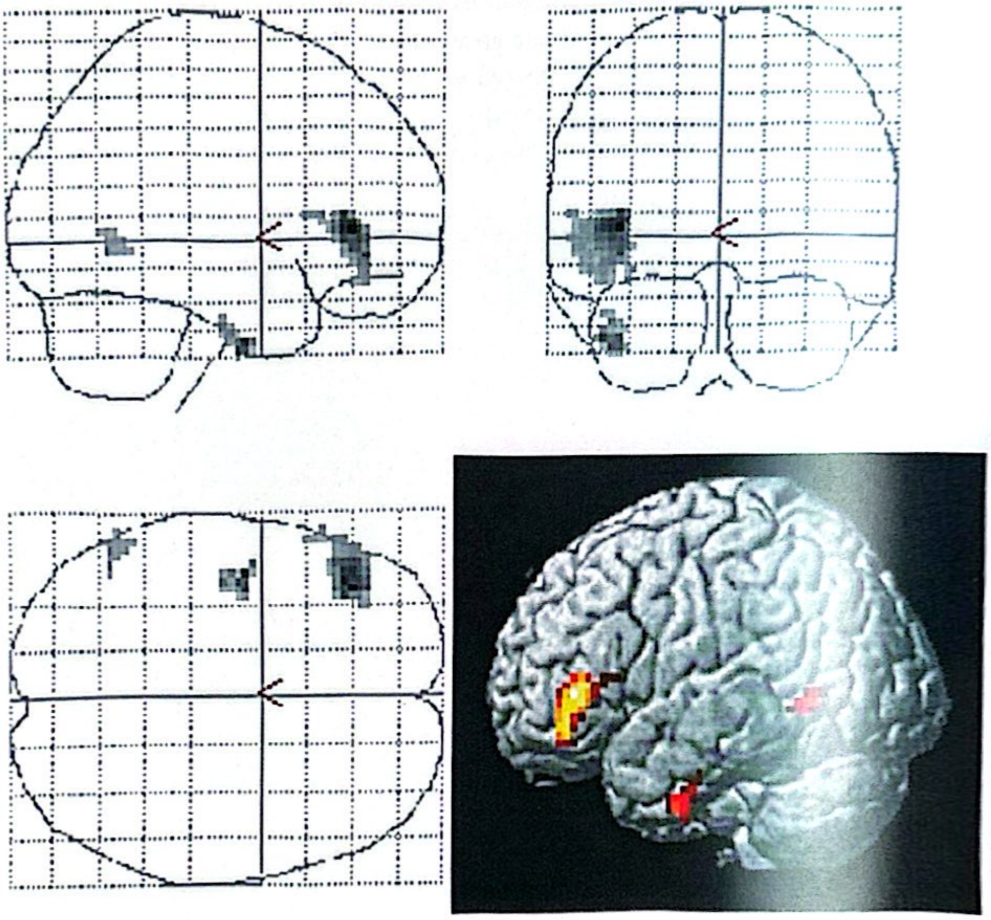


Plate 16 fMRI results. The orange areas represent parts of the brain that responded with greater activity to novel metaphors compared to conventional/familiar metaphors. The circled area is the right homologue of (counterpart to) Wernicke's area. *Source:* Mashal, Faust, Hendler, and Jung-Beeman, 2007, p. 123.

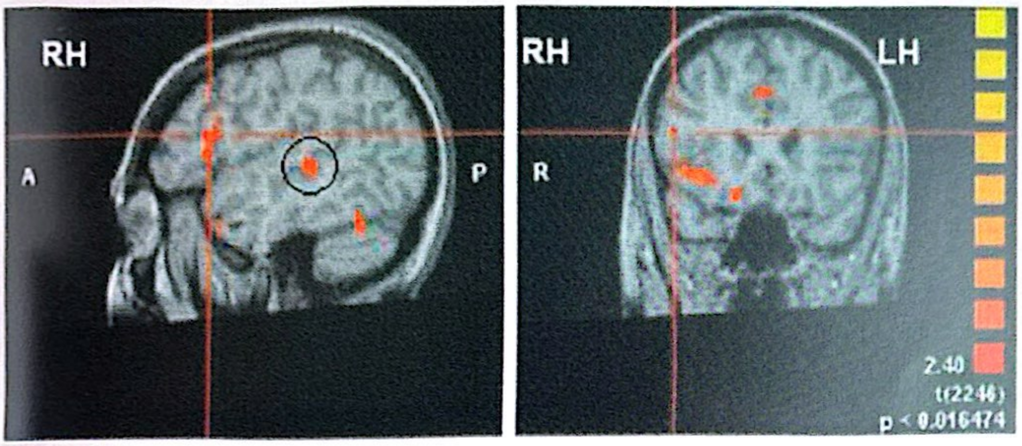


Plate 17 Neuroimaging data. Left: A comparison of neural activity for novel pseudowords (left) and familiar words (right). Right: Activity associated with real word reading subtracted from activity in pseudoword reading. The left hemisphere is shown on the right; the right hemisphere is shown on the left. *Source:* Dietz, Jones, Gareau, Zeffiro, and Eden, 2005, pp. 86, 88.

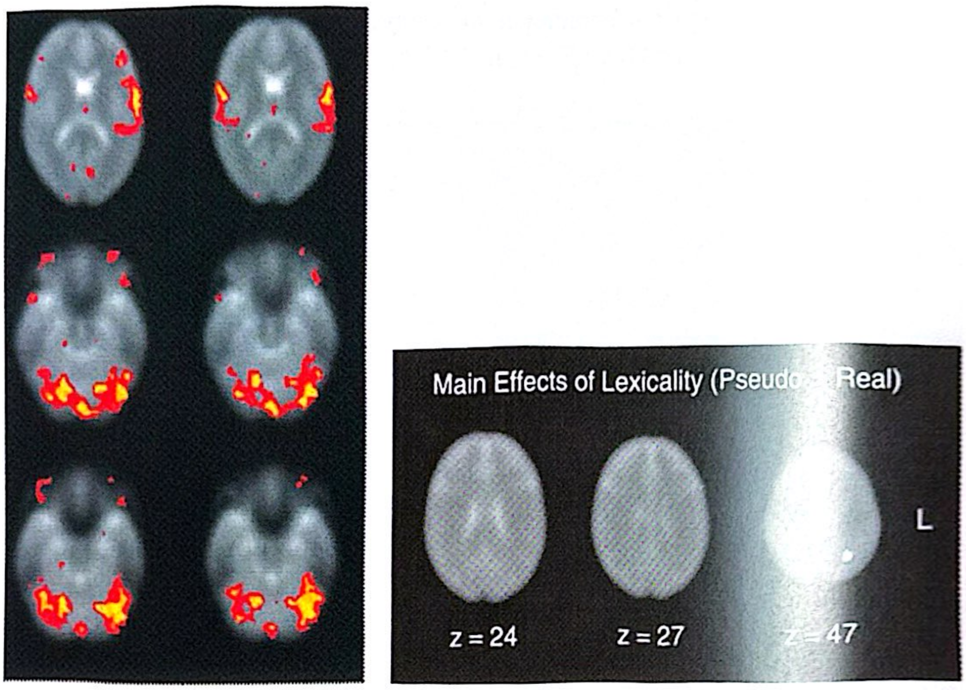


Plate 18 Brain response to viewing ASL sentences. The graph shows the difference between the brain's response to meaningless and meaningful ASL gestures. The left hemisphere appears on the left-hand side. The right hemisphere appears on the right. The top two pictures are of hearing people who do not know sign language. Unsurprisingly, their brain's response to meaningless and meaningful signs is the same. The middle pictures show that deaf signers activate both hemispheres to a greater extent when seeing meaningful signs. The bottom pictures show that hearing signers have a bilateral response to meaningful signs as well, but they do not activate all of the right-hemisphere regions that deaf signers do. Red means very large difference in activation between meaningful and meaningless signs. Yellow means small, but still significant, difference in activation. *Source:* Neville et al., 1998, p. 924.

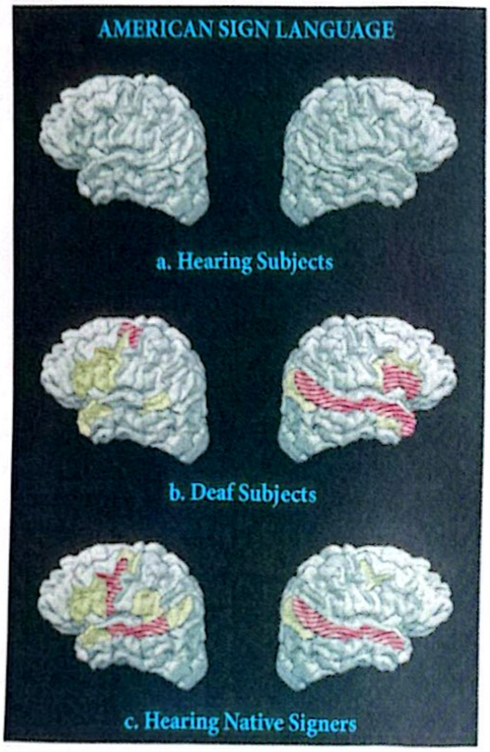


Plate 19 The arcuate fasciculus. *Source:* Catani, Jones, and Ffytche, 2005.

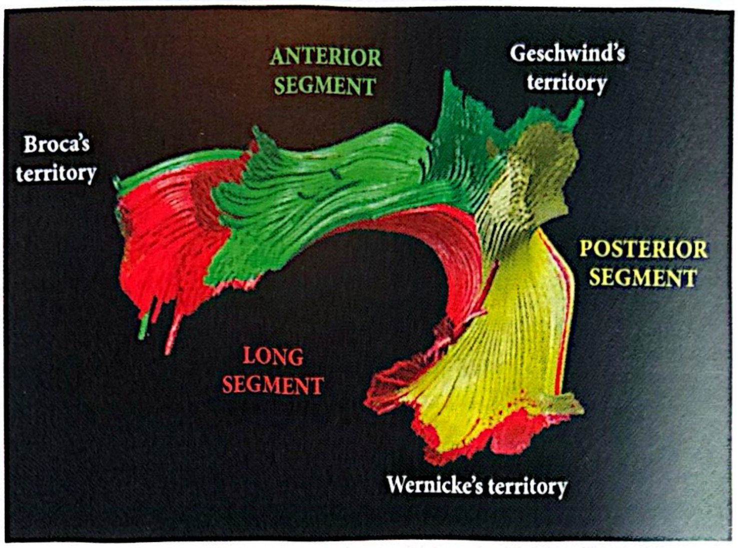
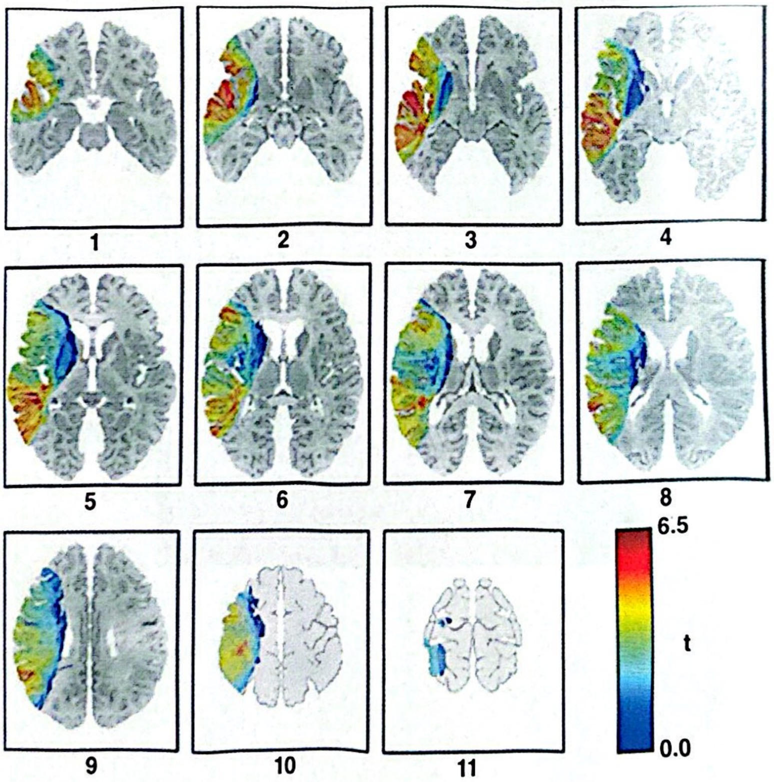


Plate 20 VLSM plot of positive t-values obtained by comparing patients with and without lesions at each voxel on the CYCLE-R sentence comprehension measure.



Children sometimes misconstrue the meanings of new words. When children learn a new word, they tend to *extend* that term by using it as a label for physically similar objects, especially those that have a similar shape (Markson et al., 2008).[19] Sometimes, children *overextend* basic-level terms, by using a word like *bunny* to refer to all small, furry creatures, including cats and dogs. When they overextend, they apply a known label to a category for which competent adult speakers use a different label.[20] Children also sometimes *underextend* terms. In that case, the child fails to use a familiar label for other members of the same category. Bloom (2000) reports a child who, for a time, believed that only cars that he could see through his front window could be called car. All other cars were called something else. Another child reportedly interpreted *shoe* as referring only to the members of one set of shoes in one particular closet. While these types of errors represent a small minority of children's word output, they offer further evidence that Quine was on to something when he suggested that lining up sounds and meanings is not a trivial problem.

While children sometimes misconstrue the meanings of novel words, all normal children eventually come to understand words about the same way everyone else does. To achieve this feat, children apply some further strategies that allow them to identify and refine word meanings. These strategies include the *mutual exclusivity assumption* and the related *principle of contrast* (Clark, 2009; Markman and Wachtel, 1988).

Mutual exclusivity involves the assumption that no two words in the language have exactly the same meaning. If a child already knows the name of a concept, that child will reject a second label as referring to the same concept. Children can use this principle to figure out the meanings of new words, because applying the principle of contrast rules out possible meanings. If you already know that *gavagai* means "rabbit," and your guide points at a rabbit and says, *blicket*, you will not assume that *gavagai* and *blicket* are synonyms. Instead, you will consider the possibility that *blicket* refers to a salient part of the rabbit (its ears, perhaps) or a type of rabbit or some other salient property of rabbits (that they're cute, maybe). In the lab, children who are taught two new names while attending to an unfamiliar object interpret the first name as referring to the entire object and the second name as referring to a salient part of the object. For somewhat older children (3–4 years old), parents often provide an explicit contrast when introducing children to new words that label parts of an object (Saylor et al., 2002). So, an adult might point to Flopsy and say, *See the bunny? These are his ears*. Children do not need such explicit instruction, however, as they appear to spontaneously apply the principle of contrast to deduce meanings for subcomponents of objects (e.g. *ears*) and substances that objects are made out of (e.g. *wood, naugahyde, duct tape*).

The *principle of contrast* also helps children learn how to organize words when more than one label can apply to a given concept (Clark, 2009). The principle of contrast, like the mutual exclusivity assumption, says that two labels should not apply to the same object, but if they do, there must be some difference in meaning between the two labels. For example, the words *dog*, *mammal*, and *animal* can all appropriately apply to the same mutt, but they do not mean the same thing. The choice of which exact term a speaker uses reflects the speaker's stance toward the named object and particular concepts that the speaker wishes to highlight. If the speaker chooses the word *mammal*, they may wish to draw attention to the similarity between Fido and physically diverse members of the mammal category. If the speaker uses *dog*, they may wish to draw attention to differences between Fido and other kinds of pets.

If a child lacked the principle of contrast, the child could assume that the words *dog*, *mammal*, and *animal* are interchangeable. Further, nobody would ever correct the child if they said *mammal* when looking at Fido. (At least, it would be really strange if somebody said, *No, that's a dog, not a mammal.*) Worse still, without the principle of contrast, the child might assume that new words were just synonyms of words the child already knows. Of course, children do not make this kind of mistake. In the lab, children apply the principle of contrast to pick up new word meanings very rapidly (Markman and Wachtel, 1988). When children are shown a familiar object (e.g. a spoon) and an unfamiliar object (e.g. a whisk), and someone uses a novel word, as in *bring me the fendle*, children pick up the unfamiliar object, which indicates they have associated the new word with the unfamiliar object. It is as if children think, "Here are two objects. I know that one of them is called *spoon*. I don't know what the other one is called. I need to get the *fendle*. Because the spoon is called *spoon*, and *fendle* is different than *spoon, fendle* must be the name of the new object."

Because learning word meanings involves associating sounds and concepts, children cannot learn the meanings of words until they have some appreciation of the concepts that the words refer to. Thus, the kinds of thoughts that a child can entertain should have some effect on the way the child learns new words. Children's perceptual systems carve the world into discrete objects, and they have intuitive notions about how these objects should behave (Aguiar and Baillargeon, 1999; Hespos et al., 2009). Given that children have substantial knowledge about objects early on, we might expect them to learn names for objects before they learn names for other kinds of concepts. Since object names are primarily or exclusively nouns, we should expect infants to learn nouns before they learn other types of words, including verbs which generally convey information about relationships between objects or relationships between objects and events. Some researchers therefore believe that early word learners should show a *noun bias* (a tendency to learn nouns before other kinds of words) regardless of the way adults speak to them.

Characteristics of English make nouns relatively prominent in IDS. In other languages, including Italian, Japanese, and Chinese, nouns tend to be much less frequent, and they appear in less prominent places in IDS (Caselli et al., 1999). Does this difference in the relative frequency of nouns and verbs lead to differences in the rate at which infants in different cultures learn nouns and verbs? If concepts precede word learning, noun bias should apply to Italian, Japanese, and Chinese, just as it does to English. If the frequency with which words occur in prominent positions drives word learning, then infants in different cultures should learn nouns and verbs at different rates.

Some studies do suggest that, in languages where verbs are more salient than nouns in IDS, infants learn verbs at an elevated rate compared to nouns (Fernald and Morikawa, 1993; Tardif et al., 1999, 1997). However, studies that show faster verb than noun learning tend to use short-term observation and a limited range of tasks that seek to elicit spoken-word production. Studies that use naturalistic observation or parental report tend to show greater consistency across languages in the relative proportions of nouns and verbs early on in toddlers' vocabularies. Large-scale cross-linguistic studies on different languages that emphasize nouns and verbs to different degrees show that nouns make up a greater percentage of infants' and toddlers' early vocabulary than verbs (Bornstein et al., 2004; Setoh et al., 2021). While more work needs to be done to settle the question definitively, it appears as though noun bias could be a general property of early word learning, and so the idea that conceptual development leads vocabulary development is still plausible.[21]

Learning words also involves deducing that people intend to refer to concepts when they speak, that is, that people have *referential intent*. Knowing that other people have thoughts that differ from yours, and that they might wish to convey their thoughts to you, is part of your *theory of mind*. Some children, including children with autism, lack some aspects of the theory of mind, and as a result, they sometimes associate meanings to sounds inappropriately (Bloom, 2000). Other children, such as those with Williams syndrome, have large deficits in general intelligence but an apparently intact theory of mind, and they appear to assign meanings to words much as normal children do.

In addition to knowing that other people have private thoughts and wish to convey those thoughts when they speak, children appear to track detailed aspects of other people's mental states and abilities, and use that information when assigning meanings to words. Children can observe where speakers are focusing their attention, and they use that knowledge to infer what objects people are labeling when they speak (Baldwin et al., 1996). However, children do not simply look where the speaker is looking and assume that a new word refers to something at that location. If two objects are visible to the child, but only one of the objects is visible to a speaker, children understand that the speaker can use a novel word to refer to the hidden object, even if the speaker is looking at the visible object (Nurmsoo and Bloom, 2008). Children also pay attention to speakers' general knowledge and reliability as a source of information (Birch and Bloom, 2002; Sabbagh and Baldwin, 2001). Some speakers are more knowledgeable than others, which means that some speakers are more likely to know the meaning of a word. Children capitalize on this to learn new words (Birch et al., 2008). Children are more likely to believe an adult than a child when the adult produces a name for a novel object. Children also pay more attention to speakers who have been reliable in the past. If children are exposed to an individual who gives the wrong names for familiar objects (objects for which the child already has words), and the individual names a novel object, the child will not reflexively adopt the spoken name as the label for the novel object. All of these effects show that children do not blindly associate sounds with objects. Children engage in sophisticated deduction, and weigh multiple factors when acquiring new vocabulary.[22]

Syntactic bootstrapping

Children have a number of tools that they can deploy to figure out noun meanings, but what about verbs? As Cindy Fisher and Leila Gleitman note, young children may have more difficulty learning verb meanings than noun meanings because verbs convey more complex concepts than nouns. While nouns can refer to concrete, directly observable properties, interpreting a verb requires the child to understand the speaker's perspective on an event (Fisher et al., 1991; Gleitman and Gleitman, 1997). While observing an event involving a duck chasing a rabbit, the speaker could focus on the agent and say *The duck blickets the rabbit*, in which case, *blicket* means "chase." If the speaker focuses on the rabbit, and says *The rabbit blickets the duck*, *blicket* means "flee." The point-and-say method does not explain how children learn the meanings of verbs such as *chase* and *flee*, because the context of the event does not provide enough information, by itself, to specify what perspective the speaker is taking.

Children can take steps toward overcoming this version of the poverty of the stimulus problem by paying attention to syntactic cues to meaning. This hypothesis forms the core of the *syntactic bootstrapping* hypothesis (Brown, 1957; Fisher, 1996; Gleitman, 1990; see Pinker, 1994a, for a dissenting viewpoint).

Syntactic characteristics of utterances can support meaning inferences in a number of different ways. First, syntactic properties of utterances could help children figure out whether a new word is a noun or a verb. When children look at a still picture of someone cutting cloth with an unfamiliar tool, and they hear *In this picture, you can see sibbing*, they infer that *sibbing* is a verb and refers to the cutting action.[23] If they hear *Can you see a sib* or *some sib*, they infer that the new word is a noun. The syntactic properties of the utterance in the noun case provide further clues as to what the noun might be. Nouns can be classified as being in the *count noun* category, which is used for items that we view as being individuals (e.g. *I see a cat*, but not **I see a pile of cat*). Nouns can be classified in a different way as belonging to the *mass noun* category, which is used for substances whose components are not treated as individual items (e.g. *I see a pile of dirt*, but not **I see a dirt*). When the new word follows the article a, children infer that *sib* refers to a count noun, so they pick out the tool in the picture. When the new words follows the article *some*, they infer that *sib* refers to a mass noun, and they interpret *sib* as being the cloth in the picture. So, syntax can provide cues to word categories.

Children also use syntactic properties of verbs to figure out what verbs mean (Fisher, 2002; Fisher et al., 1991; Goksun et al., 2008; Lee and Naigles, 2008; Song and Fisher, 2005; see Dittmar et al., 2008, for an alternative outlook). In particular, children pay attention to the *subcategorization frame* to home in on a novel verb's meaning (recall that *subcategorization* refers to the number and kinds of partners that verbs have in a particular sentence). Consider the following two sentences:

(1) She blicked!

(2) She blicked her!

Even though you do not know the meaning of the verb *blicked* yet, you can draw some conclusions about what the meaning *might* be just by knowing how many arguments (partners) and adjuncts (optional partners) go along with the verb in the sentence. In sentence (1) *blicked* only has a subject argument (*She*). So, in this case, *blicked* is an *intransitive verb*. In sentence (2) *blicked* appears with both a subject and an object argument. In sentence (2) *blicked* is used as a transitive verb.

Children can take information about transitivity into account when they interpret novel verbs. For example, in an experiment by Sylvia Yuan and Cynthia Fisher (Yuan and Fisher, 2009), children listened to sentences similar to (1) or (2) and then later watched a pair of videos (see Figure 9.3). Children who had heard intransitive sentences similar to (1) pointed to the right-hand picture when asked to "Find blicking!" Children who had heard the transitive version of the training sentence pointed toward the left-hand picture. Thus, children interpret verbs presented in a transitive frame as actions that relate two different actors, but they interpret verbs presented in an intransitive frame as an action with only a single participant. By combining information gleaned from the syntactic frame (transitive vs. intransitive) with specific information provided by the video, children infer that *blicking* means *waving your arm* or *waving your right arm above your head* in one case, and *pulling someone's leg* in the other case. Note that subcategory information, by itself, is not enough to specify a unique meaning for a verb, because different verbs sometimes have the exact same subcategorization possibilities. However, children *can* figure out verb meanings by combining subcategory information, information from the physical environment, and possibly inferences about the mental state and communicative goals of the speaker as the verb is being spoken.

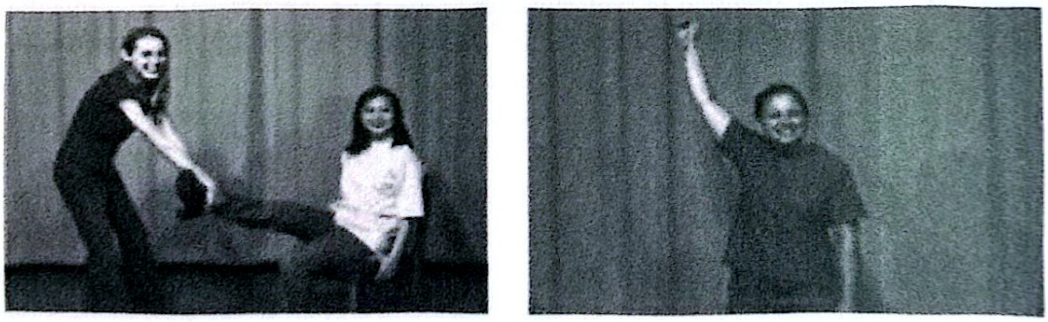


Figure 9.3 Stills from a video depicting a two-participant event (left) and a one-participant event (right) used to test young children's interpretation of the novel verb *blicking*. Source: Yuan and Fisher (2009), Association for Psychological Science

How do children figure out these syntactic cues in the first place? Cynthia Fisher and colleagues propose that when children view an event, they develop an organized conceptual representation of that event (Fisher, 2002). When someone describes that event, children associate linguistic units with elements of the (non-linguistic) structured conceptual representation. Consider the left-hand picture in Figure 9.3. In perceiving this event, children will note that someone is initiating the action (the *agent*), someone is being acted upon (the *patient*), as well as details of the action. When children hear a description of the event, such as *She's blicking her*, they map the subject of the sentence (she) to the conceptual agent, the object of the sentence (*her*) to the patient, and the verb to the action.

Fisher refers to this process of associating words in sentences to concepts in event representations as *alignment*. In her terms (Fisher, 2002, p. 56), "Children might arrive at a structure-sensitive interpretation of a sentence by structurally aligning a representation of a sentence with a structured conceptual representation of a relevant scene." Of course, if the child does not know the difference between *she* and *her*, it will not be possible to do this mapping. But given that infants start to recognize nouns fairly early in development, it is not unreasonable to assume that knowledge of nouns could bootstrap the acquisition of syntactic form. If, for example, the child knows the names of both of the young women in Figure 9.3, and heard a number of sentences describing actions involving those two, they might detect a stable pattern in the linguistic descriptions of the events. That is, that causal agents tend to come before the verb and patients come after.[24] This kind of knowledge would help the child interpret subsequent utterances that mention different sets of actors and actions.[25]

Acquisition of Morphological and Syntactic Knowledge

Every "theory of learning" that is even worth considering incorporates an innateness hypothesis.
NOAM CHOMSKY

There is a special cognitive faculty for learning and using language.
STEVEN PINKER

At around 2 years of age, children begin to create utterances that have more than one word in them. Children's ability to combine words into larger units is often indexed using a simple measure called the *mean length of utterance* (or MLU). This measure counts the number of morphemes or words that the child produces in a single utterance (which is in turn assessed by looking at pauses in the speech stream). Children at about age 2 have MLUs of just over 1, because they express ideas using single-word utterances (*No!*, *More!*, and so on).

As children mature, the mean length of utterance steadily increases and children develop more sophisticated ways to express increasingly complex thoughts. As they acquire skills, children become much more flexible in the way they use language. Instead of repeating snippets of what they hear, they craft utterances that nobody has ever said before. In other words, children become more *productive* in their language use. Many young children's utterances are ungrammatical according to adult language standards (e.g. *I want see my bottle getting fix; Mommy I poured you*), but many of them are fully grammatical (*I'm going to show you where Mr. Lion is;* Clark, 2009), and children's production performance gets closer and closer to adults'. One of the central questions in language development is: How do children acquire the skills they need to form grammatical phrases and sentences? In other words, how do they acquire adult-like knowledge of their native language grammar?

Answering this question requires a detailed analysis of the input that is available to the child learner as well as detailed analysis of the child's production and comprehension skills. This investigation can focus on any number of different types of knowledge and skills. To organize the discussion, let's consider three kinds of grammatical knowledge that children acquire: word categories, morphology, and phrase structure.

Different languages have different categories of words—for example, some languages lack adjectives (Stoll et al., 2009). Thus, a child learning a language must learn what categories the language has and where specific words fit into the system of categories. In addition, children must learn how those categories of words are expressed within phrases and sentences. Do verbs come before grammatical objects (as in English and Mandarin), after (as in Japanese), or is word order flexible (as in Russian)? The child must also learn how word categories are organized within phrases.

In addition to discovering categories and phrase organization, children must also learn aspects of *morphology*—the different forms that a word can take. Morphological marking plays a variety of important roles in language, and different languages have different morphological systems. Some kinds of morphology are used to express different flavors of meaning, such as the difference between present and past tense (e.g. *walk* vs. *walked; sing* vs. *sang*). Other aspects of morphology show how words in sentences relate to one another, such as agreement between subjects and verbs (e.g. *The cats were* ... but not *The *cat were* ...) or between determiners and nouns in languages like French and Spanish (*el burro* but not *la burro*). Morphological marking is also used in many languages to identify a word's *case*, and this helps identify what grammatical and semantic roles a word is playing in a sentence. (English has very little case marking, but other languages, such as Russian, Finnish, and Hindi, make extensive use of case marking. In English, we mark most of our pronouns for case—as in *he* vs. *him, she* vs. *her, I* vs. *me*—but not other words.) Thus, to learn a language, a child must master its morphological system.

Finally, phrase structure knowledge is an important aspect of child language learning because there are some ways of combining words that are consistent with the adult grammar, while many ways of combining words in longer expressions are possible but not consistent with the adult grammar.

The investigation of child language acquisition has been shaped by two general philosophies and theoretical frameworks. On the one hand, the *nativist approach* has assumed that an innate or instinctual *universal grammar* plays a central role in word category knowledge, as well as the acquisition of morphology and phrase structure knowledge (Chomsky, 1965; Pinker, 1994a; Wexler, 1998). That is, children are born with some of the knowledge that they will need to develop adult language skills already in place. Stephen Crain and his colleagues explain it like this (Crain et al., 2006, p. 31, emphasis mine):

> Children are born with a set of universal linguistic principles and a set of parameters *that account for variation among languages ... These innate linguistic parameters define a space of possible human languages—a space that the child explores, influenced by her environment, until she stabilizes on a grammar that is equivalent to that of adults in her linguistic community.*

One advantage of this approach is that it solves the poverty of the stimulus problem as it relates to category knowledge and phrase structure. Scientists working in the nativist tradition have argued that the input available to children is not sufficient for them to rapidly and accurately infer what the adult grammar allows because the language that children hear is full of fragments of phrases and sentences and false starts, which means they are exposed to numerous ungrammatical expressions. Further, even if the input were in perfect agreement with the adult language, nativists argue that the input is consistent with more than one grammar, but that children nonetheless invariably adopt the one grammar on which the language is actually based.

On the other hand, developments in psychological research and mathematical modeling have led some scientists to conclude that there is no such thing as a universal grammar (Evans and Levinson, 2009; Everett, 2016), that children acquire linguistic knowledge gradually and in a piecemeal fashion rather than setting parameters (Theakston et al., 2002), that the input to children is more systematic than had been previously assumed (Huttenlocher et al., 2007; Stoll et al., 2009), and that knowledge of word categories, morphology, and phrase structure can be learned by children even when the system is not "seeded" with pre-existing, innate knowledge (Redington and Chater, 1997; Westermann et al., 2009). These alternative viewpoints and models go by a number of different names—connectionism, subsymbolic computing, statistical learning, usage-based grammar, and so on. Let's keep things simple by lumping them together under the heading *probabilistic learning*, which emphasizes children's opportunistic use of many different sources of information to converge on likely solutions to complex learning problems, including the acquisition of first-language grammar. The following sections contrast the nativist and probabilistic learning approaches to category acquisition, morphology, and phrase structure.

Acquisition of word category knowledge

Language scientists are strongly divided on the question of where a child's knowledge of grammatical categories comes from. On the one hand, scientists following in the nativist tradition believe that knowledge of grammatical categories is innate (Chomsky, 1965; Pinker, 1996). Specifically, children are born with knowledge of grammatical categories, such as *noun* and *verb*.[26] Exposure to language stimuli causes children to populate their innate categories with specific words, and children exposed to different languages wind up with different sets of words and, depending on the input, different sets of categories.

The categories are populated via a learning process called *semantic bootstrapping*. According to semantic bootstrapping, learning is based on the child's ability to distinguish between physical objects, actors (agents), and actions, independent of any linguistic labels for those concepts. The tricky bit, then, is for the child to line up names referring to different kinds of concepts (e.g. physical objects vs. actions) with different components of the linguistic system (e.g. nouns and verbs). Infants solve the category learning problem by using "semantic notions as evidence for the presence of grammatical entities in the input" (Pinker, 1996, p. 40). Although this process may lead to the occasional error, the correlation between semantic characteristics and abstract grammatical categories is strong enough in child-directed speech that a significant fraction of the child's early vocabulary will be assigned to the correct categories. The same learning procedures can cause the infant to acquire knowledge of other grammatical functions, such as which parts of sentences are subjects (e.g. *The baby*), which are argument-taking predicates (e.g. *ate*), which are objects (the *oatmeal*), and which are parts of prepositional phrases (e.g. *with the spoon*).

Not everyone believes that infants come prepackaged with significant amounts of linguistic knowledge, however. To start, children's knowledge of word categories may differ in important ways from adults' knowledge of those same categories.

First, children may not have abstract, generic categories of words, such as *noun*, before they have substantial experience with the language. As Eve Clark notes (Clark, 2009, p. 167):

> *we tend to take for granted that [young children] are making use of adult-like word classes. But this assumption is far too strong. Even in such combinations as* hot + X *or* big + X, *where* X *is almost always a noun, this is not because children already have a category "noun" but because the meanings expressed by these patterns call for reference to an entity in the* X *slot.*

If children had an adult-like category of *noun*, then they should be willing and able to replace any noun with any other noun, as long as the resulting substitution leads to a meaningful, plausible sentence. However, in young children's spontaneous speech, they are very selective in the way they combine words. Rather than treating a word like *dog* as a member of a generic *noun* category, and therefore inserting that word in a variety of different appropriate locations, children start by combining such "nouns" with only a small number of verbs denoting a restricted range of actions.

Similar effects are found with young children's use of verbs. While they are willing to add inflections to verbs that belong to some semantic classes (e.g. verbs that describe activities such as *run* or *play*), they are not willing to add inflections to verbs belonging to other semantic classes. If young children had a generic category of *verb*, and treated all members as fully interchangeable, then all verbs should be equally likely to be subject to inflections. Thus, it appears as though young children's categories are based more on concrete semantic properties (e.g. *person* vs. *animal*, *activity* vs. *state*) than on abstract grammatical properties (e.g. *X* can replace any other member of the category X, regardless of the specific meaning of an individual word; any procedure that can be done with one member of category *X* can be done to all members of category X).

Second, the probabilistic learning approach challenges the idea that innate, pre-existing knowledge is necessary for children to develop category knowledge. These accounts propose that the category structure that children develop reflects the kinds of language stimuli that the child is exposed to, and the likelihood of different words appearing in different contexts, rather than on a predetermined category structure (Elman, 1993; Kolodny et al., 2015; MacWinney, 1998; Onnis et al., 2008; Waterfall et al., 2010). To support these claims, probabilistic learning advocates have developed connectionist models that take as their input the kinds of simple sentences that are prevalent in child-directed speech. Their outputs vary. In some models, the output is the prediction of the next word in a sentence given a particular sentence fragment. In other models, the existence of grammatical categories is inferred on the basis of similarities between different sentences that contain the same words. An important feature of these models is that they do not need explicit feedback to learn the structure of the language to which they are being exposed. Children most often do not receive feedback or correction from caregivers when they make grammatical mistakes, and they typically ignore such correction when it occurs (Pinker, 1996). While it remains a very open question which of the available models, if any, most closely resembles the actual mental process by which young children acquire knowledge of word categories, the existence of such models shows that word categories can be inferred from the kinds of language that infants are exposed to, which undercuts the argument that category knowledge must necessarily be in place prior to the onset of language learning.

Acquisition of morphological knowledge

As noted earlier, morphology is a central component of a language's grammatical system. Let's focus here on an aspect of inflectional morphology that has received a lot of attention: the acquisition of tense marking on verbs in English. (Acquisition of morphological systems in other languages can be analyzed along the same lines of innate vs. acquired knowledge (e.g. Gerken et al., 2005; Onnis et al., 2018).)

English verbs change from one form to another as tense and aspect change, and the specific form also depends on the person and number of the subject noun phrase. In the present tense, the verb *kick* appears as *kick* with a first-person singular subject noun (*I kick*), but as *kicks* with a third-person singular subject (*He kicks*). Regardless of the person and number of the subject noun, the past tense form is always *kicked* (*I/we/he kicked*). Many English verbs take the same suffix (-ed) in the past tense, but some don't. Adult English speakers say *I go*, but not, *I goed*. Instead, they say *I went*. Verbs like *go* are highly frequent in adult and child speech, and their morphological characteristics make them *irregular*. *Dodge, duck, and dip* are regular, because their past tense form is the stem plus the suffix -ed. *Sing, ring, think*, and *stink* are irregular because none of them ends in *-ed* in the past tense. (*Dive* is an oddball because sometimes people say *dove*, but sometimes they say *dived*.)

The question is: How do children learn the past tense forms of verbs? Perhaps they simply memorize the past tense form independently for each verb that they know. This hypothesis runs into trouble immediately on two grounds. First, if children need to hear the past tense form of a verb before they can use it, then they should have trouble coming up with past tense forms for new verbs. However, when children between 4 and 7 years old are given novel verbs, such as *trink*, and are asked to produce the past tense form, they do so easily, most often producing a regular past tense form, such as *trinked*, rather than the irregular analog to the verb *think, trought* (Berko, 1958). After a certain amount of exposure to English, children appear to acquire a procedure that says: To form the past tense, add *-ed* to the present tense stem. Such productive ability shows that children know more than just memorized forms.

The second problem is that, if children are memorizing past tense forms before they use them, then they should almost never make errors because the language model that they are relying on for input will almost never have the wrong past tense form (i.e. adults speaking would almost never say *ringed* instead of *rang* or *thinked* instead of *thought*). While children in the early parts of the multi-word production stage, beginning around the age of 2, make almost no errors with irregular past tense verbs, over-using the regular form (e.g. saying *thinked* instead of *thought*) emerges during later stages of development (at about 3 years old for some children). Gradually, children learn to differentiate between regular past tense verbs and verbs that require special past tense forms. Children producing forms that never appear in the input (*thinked, singed, goed*) in addition to producing regular past tense forms for verbs that they have never heard before shows that they are not just memorizing and repeating things that they hear.

Pinker's *words-and-rules* account explains children's use of past tense forms and other aspects of morphology, such as plural nouns and possessives (Pinker, 2000). According to this account, infants begin by categorizing words. Early in development, the verb category is sparsely populated and children treat each individual verb as an independent entity. As they are exposed to more and more language, they recognize similarities between different versions of the same verb (*kick* and *kicked, tickle* and *tickled*), and they develop an insight: The past tense can be generated from the present tense form by a rule (add -ed). After they have this insight, they apply the rule willy-nilly and all verbs, even the irregular ones, are subjected to the add *-ed* treatment. This over-application of the rule, or *over-regularization*, causes children to make errors until they create a separate list in their long-term memory of irregular verb forms.

Thus, the mature past-tense production system consists of two components. One is a rule-based system that says "look up the present tense stem and add -ed." The other is a list of exception words that needs to be searched any time a past tense verb is produced. The list of exception verbs is compiled as children are exposed to a sufficient number of examples of each exception verb. The words-and-rules account explains why children have a "u-shaped" learning curve for past tense morphology. They start out correctly copying independently memorized forms, they make errors once they have sufficient experience to notice and overuse a rule that relates past tense and present tense forms of verbs, and finally they compile a list of past tense verbs that don't play by the rules, thereby eliminating over-regularization errors.

Proponents of the probabilistic learning approach have challenged the words-and-rules formulation on a number of different grounds and this facet of development continues to keep language scientists awake long into the night (Joanisse and Seidenberg, 1999; Seidenberg and Joanisse, 2003; Westermann and Jones, 2021). Part of the problem revolves around the description of how, exactly, children behave as they are learning to produce past tense forms for verbs. While the words-and-rules approach argues for the sudden onset of over-regularization, consistent with sudden insight into the existence of a rule, the actual data may not support this claim.

First, studies of child language development are often plagued by sparse data problems. That is, very few children are studied, those who are studied are often the children of academic linguists (who have the motivation to keep the necessary detailed records, but who may be subject to certain observational biases), and many studies sample a very small fraction of the child's total language output. Even worse, the context in which the child is speaking often goes unrecorded, as does the overall ambient language that the specific child is exposed to. This problem is particularly acute when the critical question concerns how suddenly a child begins to apply a rule. On that account, the child should go from having a mixture of regular and irregular forms, because production will be based on repeating memorized input, to essentially 100% regular forms once the rule kicks in, after which there should be a gradual decline in the overall proportion of regular forms as the irregular forms once again assert themselves in the child's output. When real children's output is examined, they appear to start out marking only some verbs using the regular past tense, and only gradually increase their use of the regular past tense (McClelland and Patterson, 2002).

Probabilistic learning advocates have also noted that, rather than being applied across the board, regularization of past-tense verbs occurs more often in some semantic contexts than others, and more often in some phonological contexts than others. All of these phenomena pose problems for the words-and-rules approach. Nonetheless, we might still favor this approach if there were no plausible alternatives. However, as with the development of word category knowledge, proponents of the probabilistic learning framework have developed a number of connectionist models of English past-tense morphology (Joanisse and Seidenberg, 1999; McClelland and Patterson, 2002; McClelland and Rumelhart, 1985). These models have the advantage of capturing the u-shaped pattern of acquisition that normally developing children exhibit. They have the additional advantage of applying over a wider scope than the words-and-rules formulation. For example, implemented connectionist models can be used to predict children's responses to novel verbs in different phonological and semantic contexts as well as explaining morphological errors in developmental language disorder.

Acquisition of phrase structure knowledge

An acquisition theory that faces occasional counterexamples is better than no acquisition theory at all.
STEVEN PINKER

Phrases exist so that components of events can be tied to linguistic elements that convey who did what to whom. Different languages combine words in different ways, and they use different forms to mark agents, patients, instruments, and other role-players in events. A child learning a language must discover the way phrases are organized in order to convey thoughts relating to events.

Eve Clark (Clark, 2009, p. 158) summarizes the challenge infants face in this way:

> *Children who wish to talk about events need to be able to analyze what they observe to decompose scenes into constituent parts relevant to linguistic expressions in the language they happen to be learning. They have to work out … how to talk about agent versus patient, location versus instrument, or beneficiary versus recipient. They must find out how to mark grammatical relations such as subject and object. And they must also learn how to indicate that the elements in a constituent (a noun phrase or predicate for instance) belong together, through agreement, adjacency, or both, depending on the language.*

As with other aspects of grammar learning, nativist and probabilistic learning approaches make competing claims about how and why children acquire the skills they need to organize words into phrases that conform to the standard imposed by the adult language.

The nativist approach to phrase structure learning argues that the basic knowledge children need to combine words into phrases is present in latent form at birth in the form of *parameters*. Basic word order—whether you produce your subjects before your verbs, or your verbs before your subjects—varies across languages. Thus, there is no pre-existing parameter that says "put your subjects at the beginning of your sentences," but there *is* a pre-existing parameter that says "subjects either come before verbs or after." It is up to the individual infant, armed with this knowledge, to pay attention to the ambient language and figure out how their own particular language relates its subjects to its verbs.

Other language characteristics are also argued to be governed by parameters. For example, English requires subjects to be expressed in the verbatim form of a sentence (e.g. *He ate bananas*), even if the subject is a semantically null placeholder (as in *It rained*). Other languages, such as Italian, allow sentences to have an implicit subject. *Ate bananas* would be OK in Italian, as long as context makes it clear who the eater is, but that sentence is not OK in English. The infant learner of Italian sets the "pro-drop" parameter one way; the infant learner of English sets it a different way. According to the nativist theoretician, languages differ from one another because the overall set of parameters takes on different settings for different languages (in addition to arbitrary differences in vocabulary).

One prominent nativist account (Pinker, 1996) proposes that children have considerable knowledge of phrase structure formation very early on, even when they are just beginning to produce utterances that have more than one word in them. Note that the phrase structure rules that children are claimed to possess are essentially identical to the rules that adults have. This claim of equivalence between the child and adult grammars goes by the name the *continuity hypothesis*, which emphasizes the hypothetical similarity of child and adult linguistic knowledge. Examples of this grammatical knowledge are that sentences have subject noun phrases and verbs, that verb phrases consist of a verb and an object noun, and that noun phrases can consist of possessives, adjectives, and quantifying modifiers (e.g. *my shoe, big shoe*, and *some shoes*, respectively; other phrase structure rules are also argued to be in place). A grammar that contains rules such as these will be capable of producing the kinds of utterances that are seen early on in the two-word stage, such as *Mommy fix* (sentence = NP + VP), *mama dress* (NP = adjective + noun), and *more milk* (NP = quantifier + noun).

One complication for such an account is that, while many utterances produced by young children are grammatical according to the adult grammar, some are not. For instance, young children do not always include the grammatical morpheme *to* when they produce sentences that require it. So, they say **I want hold Postman Pat* instead of *I want to hold Postman Pat* (Kirjavainen et al., 2009). Thus, it is possible that the child version of the relevant phrase structure rule is that V (infinitival) = infinitival verb + noun, rather than VP = *to* + infinitival verb + noun. Pinker argues that this explanation of the child's production patterns is unsatisfactory because the child will never encounter evidence that this rule is incorrect, and so there should be adults running around saying things like *I want go Denny's*, or *I need talk my lawyer get the charges dropped*.

In contrast with the nativist account of phrase structure knowledge, probabilistic learning theorists contend that learning phrase structure rules, like learning other aspects of language, results from children analyzing the input to which they are exposed. This conclusion is supported by observations that children acquire knowledge of some phrase structure types gradually (parameter setting predicts sudden onset of phrase structure knowledge) and that children's spontaneous language production mirrors the frequency with which sequences of words occur in language addressed to children (Kirjavainen et al., 2009; Marchman et al., 1991). For example, verbs that require the grammatical morpheme *to* when they appear in verb-to-verb sequences (e.g. *want to dance*) can often appear in constructions without the word *to* (e.g. *want ice cream*). If children do not have the adult phrase structure rule (infinitival verb = verb + *to* + verb), but instead construct *schemas* by paying attention to what precedes and follows a specific verb, such as *want*, then they will notice that the verb *want* sometimes is followed by to, but sometimes is not.

The idea that children learn about phrase structure on a verb-by-verb basis predicts that their production patterns should correlate with the way their caregivers talk. In fact, children who make more production errors (e.g. saying *I want hold Postman Pat*) are those children who frequently hear verbs like *want* in sentences that do not have the word *to* in them (e.g. *I want ice cream, Polly wants a cracker*). The ease with which children learn to use new verbs and new syntactic structures also corresponds to the frequency with which particular phrase structures appear in the input (Abbot-Smith and Tomasello, 2010; Casenhiser and Goldberg, 2005).

Findings that errors correlate with input have motivated some theorists to favor the *usage-based grammar* account of phrase structure learning (Kidd et al., 2010; Lany et al., 2007; McClure et al., 2006; Tomasello, 2000). According to this account, phrase structure acquisition is closely tied to the acquisition of individual verbs. Rather than developing a general phrase structure rule, such as "a verb phrase is a verb plus a noun phrase," young children first learn how individual verbs behave, and only gradually form larger abstract classes of verbs by noticing that different verbs behave in similar ways.[27] Once these larger classes of verbs have been formed, children can then develop the idea that some phrase structure patterns occur repeatedly in the language, and therefore develop a more abstract notion of what phrase structures can look like in the language as a whole.

The usage-based account makes a number of predictions that can be evaluated by observing how young children speak. For instance, it predicts that, because children are paying attention to how individual verbs behave, they will be conservative in the way they use newly acquired verbs. That is, they will be unlikely to use a verb to express a particular phrase structure unless they have heard somebody else use that verb with that phrase structure. If children hear the sentence *Mommy drank*, which has a subject argument only, they will be unlikely to say *Mommy drank the milk*, which has both a subject and object argument. Lacking a phrase structure rule that says "a verb phrase can be a verb plus a noun phrase," young children are reluctant to add an object noun phrase to *drank* until they have positive evidence that the language allows that to happen. In fact, young children between 2 and 3 years old are conservative in just this way (Lieven et al., 1997). Children in this age range are even willing to violate the general word order that the adult grammar dictates (e.g. subject–verb–object), but only for low-frequency verbs or verbs that they have not been exposed to previously, if an adult speaker models the strange word order (Chan et al., 2010). Children are also more likely to correct ungrammatical phrase structures when the ungrammatical phrase structure contains a familiar verb (Matthews. et al., 2007). Even complicated phrase structures, such as the embedded sentence structure in *I think Mommy drank the milk* (sentence = sentence + embedded sentence, embedded sentence = noun phrase + verb phrase, verb phrase = verb + noun phrase) depend on the acquisition of specific verbs. In this instance, children's use of the embedded sentence phrase structure is driven almost entirely by their acquisition of the verb *think* (Kidd et al., 2010), which appears very frequently accompanied by an embedded sentence in the language that young children hear. These findings, and others, suggest that young children's knowledge of phrase structure is intimately connected to individual verbs that the child knows, rather than being fully abstract. As with acquisition of word category and morphological knowledge, usage-based theorists have developed mathematical models that acquire phrase structure knowledge as the result of exposure to samples of child-directed speech (Bannard et al., 2009).

Summary and Conclusions

Language learners face difficult tasks as they try to acquire a first language. They have to take a stimulus that comes at them in huge blocks and break it down into manageable chunks. They must learn to associate each bite-sized chunk with some sort of meaning. Neither of these tasks is trivial. Both involve the child deploying a lot of mental firepower to overcome substantial obstacles, such as the segmentation and poverty of the stimulus problems. Fortunately, the child gets an early start, as learning about prosodic features of the native language begins in the third trimester, well before the baby is born. Infants also appear to be endowed with perceptual and representational skills that enable them to tell the difference between different speech sounds from the moment they are born (or at most, within the first 24–48 hours). Knowledge of the prosodic characteristics of utterances represents at least a plausible mechanism for infants to break into the speech stream and start identifying words. Children as young as 2 months old can tell the difference between phonetically identical utterances that have different prosodic qualities. While it may take infants some time to figure out all the details of the prosodic system, they are capable of using prosodic cues to segment (and remember) words out of fluent utterances by 7 months old. Characteristics of IDS appear to be particularly well suited to help infants make use of prosodic cues to word boundaries. Words that are spoken in isolation also appear early on in infants' vocabularies, and so infants may rely to some degree on caregivers doing the segmenting for them some of the time.

Once infants have begun to amass a vocabulary of familiar words, they can use the boundaries of those familiar words to mark out the edges of unfamiliar words. In fact infants as young as 6 months old appear to use such a "top-down" strategy. Infants also appear to use statistical likelihood information. Children as young as 6 months old can use the likelihood of one syllable following another to segment word-like units out of a continuous speech stream. We know for sure that infants have a set of tools that is up to the task of segmenting speech; whether language scientists have identified all the tools, or the right set of tools, remains an open question.

Infants use another set of tools to assign meanings to words. They need these tools because the environment does not uniquely specify which meanings go with which words. Infants and young children, however, bring to the task a powerful set of perceptual abilities (for example, object recognition comes to the language-learning domain essentially for free) as well as a powerful set of social-cognitive abilities. Point-and-say leverages infants' object recognition abilities. When you point at a rabbit in the grass, you can be certain that the infant will appreciate that the rabbit is a coherent object, separate from its surroundings. Combine the infant's object recognition skills with a bias to interpret new words as whole-object labels, and *gavagai* becomes *rabbit*. But there is much more to word learning than point and say. Children appear to understand without being taught that other people have different knowledge and perspectives on events, and that private knowledge and perspectives will affect the way speakers behave. Thus, children can figure out where adults are focusing their attention, and they can flavor their interpretations of utterances accordingly (in Bloom's terms, infants and young children are pretty good mind readers). But, as with the point-and-say game, children are not slaves to joint attention. Young children (3–4 years old) can view the world from a speaker's perspective, and use inferences about that perspective to assign meanings to novel words. Once children have acquired some basic knowledge about the grammar and syntax of their native language, they can add this knowledge to the toolkit and use it to infer the meanings of novel verbs. As with segmentation, we know that children solve the poverty of the stimulus problem for word meanings. We have some good hypotheses about what those tools are, and research continues to further refine and develop these hypotheses.

Young children begin to string words together into multi-word utterances starting about age 2. To create multi-word utterances that are consistent with adult language standards, children must identify word categories, the morphological markings that go along with different semantic and syntactic functions, and the patterns that govern how words can be put together into phrases and clauses. The nativist approach argues that much of this knowledge is innate. In particular, nativists propose that word category and phrase structure knowledge are in place prior to the onset of language learning. This approach has the advantage of offering a straightforward answer to the poverty of the stimulus problem. Probabilistic learning advocates explicitly deny the existence of innate word category, morphological, and phrase structure knowledge. To support their position, they have presented modeling data showing that word category and morphological knowledge can be acquired as a by-product of unsupervised learning processes. They have also challenged predictions made by the nativist position, specifically that young children will apply phrase structure rules suddenly and broadly once the appropriate parameters are set. When actual children are observed, they appear to master morphological and phrase structure knowledge more gradually than had been previously assumed. In addition, knowledge of phrase structures appears to be closely tied to individual words, especially verbs. Phrase structure knowledge also appears to differ in strength across different verbs, with strength being closely associated with patterns that occur in the language that the child hears.

TEST YOURSELF

1. Explain how prenatal infants can acquire knowledge about language. What aspects of language do they learn and why? How might this knowledge pave the way for postnatal development?
2. Describe a typical HAS experiment. What do the results indicate about infant language skills?
3. What do we know about newborn infants' phonological perception abilities? What role does innate knowledge play? What evidence favors or challenges the idea that infant phonological perception depends on species-specific biological mechanisms?
4. How does an infant's ability to perceive phonological contrasts change as the infant matures? What accounts for these changes?
5. Describe the segmentation problem and explain how children solve it. What role does prosody play? What role does statistical learning play?
6. Describe IDS and explain how it affects the acquisition of language skills. What happens when an infant's caregiver is depressed? What happens in cultures where adults do not produce IDS?
7. How do infants and young children learn the meanings of words? What role does "point-and-say" play? How do children overcome the poverty of the stimulus problem? What role do categorization biases play? What role does (non-linguistic) conceptual knowledge play?
8. Describe two competing accounts explaining the acquisition of grammar and the evidence that supports each. Who has the better case, the nativists or the probabilistic learning theorists? Why?

THINK ABOUT IT

1. How is it possible for blind children to acquire vocabulary? How do you think their vocabulary acquisition process compares to sighted children? How do you think it compares to deaf children?
2. Design an experiment to see whether your friends can learn a new language via statistical learning. (Hint: You could use Saffran's fake syllable method.) Can your friends identify the "words" in your fake language if you expose them to those "words" the way Saffran exposed babies to new "words"?

Notes

1 Although in some experiments, the baby is rewarded for sucking less frequently than baseline (e.g. DeCasper and Fifer, 1980).

2 Steve Pinker has suggested that humans and chinchillas develop categorical perception of speech sounds for different reasons and via different perceptual mechanisms. Human infants demonstrate categorical perception after one or two training trials. Chinchillas and quail can require hundreds or thousands of training trials before they demonstrate similar ability. While this objection may apply to categorical perception in non-human animals, it does not apply to findings of categorical perception in humans of nonspeech sounds.

3 To learn about aspiration, hold your hand an inch or so in front of your mouth and say the words *pill* and *spill*. When you say the word *pill*, the /p/ sound is aspirated. You can feel the burst of air right after you start saying *pill*. When you say the word *spill*, there is no aspiration, and you should feel much less air moving after the /p/ sound in *spill*.

4 See Pardo and Remez (2006) for arguments against this proposal.

5 7½-month-olds can also do this trick in reverse. If they are trained on short sentences and later tested on isolated words, they can recognize individual words that were presented as part of the fluent-speech training stimulus (see Jusczyk and Aslin, 1995, Experiment 4).

6 The metaphor is that of "pulling yourself up by your own bootstraps," as babies have to learn how to segment the speech stream all by themselves. No one can teach them how to do it because you can't take verbal instruction before you know how to identify words. Ignore the fact that pulling yourself up by your own bootstraps is, in fact, physically impossible.

7 Some may object that it is a huge leap to assume that infants could develop such a hypothesis spontaneously (i.e. that babies are dumb). But given the current state of research on infant cognition, we have to at least consider the possibility that babies are smart.

8 This experiment takes advantage of the ERP component called a *mismatch negativity*. In experiments where people are repeatedly exposed to one kind of stimulus (e.g. iambic word) called the *standard*, and are *in*frequently exposed to a different stimulus (e.g. trochaic word) called the *deviant*, the ERP signal shows a negative-going change in voltage measured at the scalp about 200 ms after the onset of the deviant stimulus. Because the ERP wave occurs in response to a stimulus that is different than, or does not match, the more commonly occurring stimulus, researchers call this a mismatch negativity.

9 They *might* want to segment it as *Thisp lace*, though, because some English words do end in /sp/ (as in *clasp*, *grasp*, and *wasp*), and some end in /isp/ (as in *crisp*). So, in this case, phonotactics, by itself, does not lead to one unique, correct segmentation.

10 That is, their output becomes more accurate faster.

11 At least in the industrialized world. Adults in some cultures do not react to babies in this way. Still, the equivalent of infant-direct speech has been observed in China (Mandarin), France, Germany, Italy, Japan, the United Kingdom, and the United States (Cooper and Aslin, 1990).

12 Males also speak motherese to infants, but the name sticks because females still do the bulk of child rearing in places where motherese is spoken.

13 One more reason why mothers with symptoms of postnatal depression should seek professional help.

14 Similar differences in transitional probabilities between phonemes occur within and between syllables. The transitional probability of a /gp/ sequence within a syllable is essentially zero in English. *igp and *gpi are both blocked by phonotactic rules. But the sequence /gp/ can occur in a word, so long as the /g/ sound ends one syllable and the /p/ sound starts the next (as in *pigpen*). In this case, the transitional probability of a phoneme sequence is *higher* between units (syllables) than within units. Thus, the absolute value of a transitional probability need not determine whether people use that information to identify important units of speech. The learner could use higher probability sequences of syllables to group them together, but use higher probability sequences of phonemes to split syllables apart (Seidenberg and McClelland, 1989).

15 The Achilles heel of this approach is what I call the *Superhero* effect. Consider Batman/Bruce Wayne, Superman/Clark Kent, Wonder Woman/Lynda Carter, and Catwoman/Beyonce Knowles. The two halves of each pair are never seen together, but we know they are the same person (because we saw the movie or read the comic books). A similar situation occurs in language. The /s/ morpheme (as in *cats*) and the /z/ morpheme (as in *dogs*) are in *complementary distribution*: they occur in completely different contexts. But we know that they are the outward manifestations of the same underlying linguistic construct: plural marker. We recognize the similarity (*cats* = more than one cat; *dogs* = more than one dog), even though a straightforward statistical analysis would show that the two forms never co-occur, and therefore should be treated as conceptually separate.

16 A transitional probability of 1.0 means that the first syllable is always followed by the same second syllable. A transitional probability of 0.33 means that the first syllable is followed by a particular syllable a third of the time.

17 The anthropologists inform us that not all cultures play this game. Apparently, the iKung San do not.

18 There are exceptions to the rule. We individuate familiar people and animals. So, if you see a picture of your friend, you don't say *person*, you say *Shelley*. Similarly, if the picture was your pet dog, you wouldn't say, *dog*, you would say *Fido*. This constraint also applies in reverse. If you see a picture of a person and a chair and someone says, *That's Linda*, you assume that the label goes with the person, not the chair. That's why it's funny when George Carlin names his vibrating chair *Linda* in *Scary Movie 3*.

19 Shape bias can be overridden in special circumstances. Children group together objects that have similar functions (are used to accomplish the same goal), whether the two objects are physically similar or not (Kemler-Nelson, 1999).

20 Some instances of apparent overextension may not really be errors. If a child looks at a dog and says *kitty*, it does not mean that the child is mistaking the meaning of the word *kitty* or misidentifying the dog as a cat. The child might simply want us to notice the similarity between dogs and cats, but does not yet have the right vocabulary to express that thought. Babies are smart, so this could really happen.

21 This is somewhat more complicated than it seems. Children develop concepts from directly perceiving the world, in addition to having some innate ideas about how the world works. But languages divide up conceptual space in different ways, and assign words to concepts in different patterns. In English, the word *fit* describes a broad range of containment relations. If one thing goes inside another, we say that they *fit*, regardless of whether it's an arm in a sleeve, a key in a lock, or a peanut in a mason jar. But Korean uses different terms to indicate a loose fit and a tight fit. So, English-speaking and Korean-speaking infants are faced with different perceptual problems. The Korean language learner, but not the English one, has to pay attention to how tightly two objects join to select the right word (Hespos and Spelke, 2004). Further, the Korean speaker may notice relationships between objects that the English speaker does not perceive as a result of needing to select between competing versions of *fit* (see Chapter 1).

22 That does not mean that children *consciously* engage in logical deduction to infer word meanings. But it does mean that the thought process that underlies vocabulary acquisition factors in a variety of sources of information, including those that are made available by the child's theory of mind.

23 This cue cannot be sufficient by itself, however, as many nouns also end in -ing, and many words are ambiguous between noun and verb meanings, as in *spring* (noun vs. verb; *coiled metal* vs. *jump*) and *stinging* (verb vs. adjective).

24 This could happen because babies are smart.

25 Prosodic cues may also play a role in bootstrapping syntactic knowledge. Prosodic cues such as pauses and particular tone patterns often appear between important syntactic components, such as phrases and clauses (Jusczyk, 1997; Speer and Blodgett, 2006). Although the correspondences are not perfect, they are consistent enough to provide cues to important syntactic boundaries. Research shows that infants are sensitive to these cues, as they will listen longer to fluent speech that has pauses and other prosodic cues inserted between syntactic constituents than to speech where the prosodic cues are misaligned with the syntax (Jusczyk et al., 1992; Jusczyk and Kemler-Nelson, 1996). IDS may enable infants to pick up on syntax-prosody correspondences at an earlier age. For example, 9-month-olds are only sensitive to prosody-syntax correspondences in utterances that have IDS prosody, and do not respond to manipulations of prosody-syntax correspondences in adult-directed speech.

26 The innate component of linguistic knowledge includes more than just word categories. "The child is assumed to know, prior to acquiring a language, the overall structure of the grammar, the formal nature of the different sorts of rules it contains, and the primitives from which those rules may be composed" (Pinker, 1996, p. 31).

27 Artificial grammar learning experiments also support the idea that children start with narrow assumptions about what the language allows, and switch to broader generalizations when positive evidence for such generalizations appears in the input (Gerken, 2006, 2010; Gerken and Bollt, 2008).

References

Abbot-Smith, K., and Tomasello, M. (2010). The influence of frequency and semantic similarity on how children learn grammar. *First Language*, *30*, 79–101.

Aguiar, A., and Baillargeon, R. (1999). 2.5-month-old infants' reasoning about when objects should and should not be occluded. *Cognitive Psychology*, *39*, 116–157.

Bagou, O., and Frauenfelder, U.H. (2018). Lexical segmentation in artificial word learning: The effects of converging sublexical cues. *Language and Speech*, *61*(1), 3–30.

Baldwin, D.A., Markman, E.M., Bill, B., et al. (1996). Infants' reliance on a social criterion for establishing word-object relations. *Child Development*, *67*, 3135–3153.

Bannard, C., Lieven, E.V.M., and Tomasello, M. (2009). Modeling children's early grammatical knowledge. *Proceedings of the National Academy of Sciences of the United States, 106,* 17284–17289.

Barajas, M.C.O., Guevara, R., and Gervain, J. (2021). The origins and development of speech envelope tracking during the first months of life. *Developmental Cognitive Neuroscience, 48,* 100915.

Berko, J. (1958). The child's learning of English morphology. *Word, 14,* 150–177.

Birch, S.A.J., and Bloom, P. (2002). Preschoolers are sensitive to the speaker's knowledge when learning proper names. *Child Development, 73,* 434–444.

Birch, S.A.J., Vauthier, S.A., and Bloom, P. (2008). Three- and four-year-olds spontaneously use others' past performance to guide their learning. *Cognition, 107,* 1018–1034.

Bloom, P. (2000). *How Children Learn the Meanings of Words.* MIT Press.

Bornstein, M.H., Cote, L.R., Maital, S., et al. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child Development, 75,* 1115–1139.

Bortfeld, H., Morgan, J.L., Golinkoff, R.M., et al. (2005). Mommy and me: Familiar names help launch babies into speech-stream segmentation. *Psychological Science, 16,* 298–304.

Brent, M.R., and Siskind, J.M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition, 81,* B33–B44.

Brown, R.W. (1957). Linguistic determinism and the part of speech. *Journal of Abnormal and Social Psychology, 55,* 1–5.

Caselli, C., Casadio, P., and Bates, E. (1999). A comparison of the transition from first words to grammar in English and Italian. *Journal of Child Language, 26,* 69–111.

Casenhiser, D., and Goldberg, A. E. (2005). Fast mapping between a phrasal form and meaning. *Developmental Science, 8,* 500–508.

Chambers, K. E., Onishi, K. H., and Fisher, C. (2003). Infants learn phonotactic regularities from brief auditory experience. *Cognition, 87,* B69–B77.

Chan, A., Meints, K., Lieven, E.V.M., et al. (2010). Young children's comprehension of English SVO word order revisited: Testing the same children in act-out and intermodal preferential looking tasks. *Cognitive Development, 25,* 30–45.

Chomsky, N. (1965). *Aspects of the Theory of Syntax.* MIT Press.

Chomsky, N. (1981). *Lectures on Government and Binding.* Mouton de Gruyter.

Christiansen, M.H., Allen, J., and Seidenberg, M.S. (1998). Learning to segment speech using multiple cues: A connectionist model. *Language and Cognitive Processes, 13,* 221–268.

Clark, E. (2009). *First Language Acquisition* (2nd ed.). Cambridge University Press.

Cooper, R.P., and Aslin, R.N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development, 61,* 1584–155.

Cooper, R.P., and Aslin, R.N. (1994). Developmental differences in infant attention to the spectral properties of infant-directed speech. *Child Development, 65,* 1663–1677.

Crain, S., Goro, T., and Thornton, R. (2006). Language acquisition is language change. *Journal of Psycholinguistic Research, 35,* 31–49.

Cutler, A. (1996). Prosody and the word boundary problem. In J.L. Morgan and K. Demuth (Eds.), *Signal to Syntax* (pp. 87–100). Erlbaum.

Cutler, A., and Norris, D.G. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 113–121.

Cutting, J.E., and Rosner, B.S. (1974). Categories and boundaries in speech and music. *Perception and Psychophysics, 16,* 564–571.

DeCasper, A.J., and Fifer, W. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science, 208,* 1174–1176.

DeCasper, A.J., Lecanuet, J., Busnel, M., et al. (1994). Fetal reactions to recurrent maternal speech. *Infant Behavior and Development, 17,* 159–164.

DeCasper, A.J., and Spence, M.J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development, 9,* 133–150.

Dehaene-Lambertz, G., and Dehaene, S. (1994). Speed and cerebral correlates of syllable discrimination in infants. *Nature, 370,* 292–295.

Dittmar, M., Abbot-Smith, K., Lieven, E., and Tomasello, M. (2008). Young children's early syntactic competence: A preferential looking study. *Developmental Science*, *11*, 575–582.

Eimas, P.D., Siqueland, E.R., Jusczyk, P., et al. (1971). Speech perception in infants. *Science*, *171*, 303–306.

Elman, J.L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, *48*, 71–99.

Evans, N., and Levinson, S.C. (2009). With diversity in mind: Freeing the language sciences from Universal Grammar. *Behavioral and Brain Sciences*, *32*, 472–484.

Everett, D.L. (2016). An evaluation of Universal Grammar and the phonological mind. *Frontiers in Psychology*, *7*, 15. https://www.frontiersin.org/articles/10.3389/fpsyg.2016.00015/full.

Fernald, A., and Hurtado, N. (2006). Names in frames: Infants interpret words in sentence frames faster than words in isolation. *Developmental Science*, *9*, F33–F40.

Fernald, A., and Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology*, *27*, 209–221.

Fernald, A., and Morikawa, H. (1993). Common themes and cultural variations in Japanese and American mothers' speech to infants. *Child Development*, *64*, 637–656.

Fernald, A., Perfors, A., and Marchman, V.A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, *42*, 98–116.

Fernald, A., Swingley, D., and Pinto, J.P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development*, *72*, 1003–1015.

Fisher, C. (1996). Structural limits on verb mapping: The role of analogy in children's interpretations of sentences. *Cognitive Psychology*, *31*, 41–81.

Fisher, C. (2002). Structural limits on verb mapping: The role of abstract structure in 2.5-year-olds' interpretations of novel verbs. *Developmental Science*, *5*, 55–64.

Fisher, C., Gleitman, H., and Gleitman, L.R. (1991). On the semantic content of subcategorization frames. *Cognitive Psychology*, *23*, 331–392.

Fisher, C., Hunt, C., and Chambers, K. (2001). Abstraction and specificity in preschoolers' representations of novel spoken words. *Journal of Memory and Language*, *45*, 665–687.

Frost, R., Armstrong, B.C., and Christiansen, M.H. (2019). Statistical learning research: A critical review and possible new directions. *Psychological Bulletin*, *145*(12), 1128.

Frost, R.L., Dunn, K., Christiansen, M.H., et al. (2020). Exploring the "anchor word" effect in infants: Segmentation and categorisation of speech with and without high frequency words. *PLoS One*, *15*(12), e0243436.

Gerken, L. (2006). Decisions, decisions: Infant language learning when multiple generalizations are possible. *Cognition*, *98*, B67–B74.

Gerken, L. (2010). Infants use rational decision criteria for choosing among models of their input. *Cognition*, *115*, 362–366.

Gerken, L., and Aslin, R.N. (2005). Thirty years of research on infant speech perception: The legacy of Peter W. Jusczyk. *Language Learning and Language Development*, *1*, 5–21.

Gerken, L., and Bollt, A. (2008). Three exemplars allow at least some linguistic generalizations: Implications for generalization mechanisms and constraints. *Language Learning and Development*, *4*, 228–248.

Gerken, L., Wilson, R., and Lewis, W. (2005). Infants can use distributional cues to form syntactic categories. *Journal of Child Language*, *32*, 249–268.

Gleitman, L.R. (1990). The structural sources of verb meanings. *Language Acquisition*, *1*, 3–55.

Gleitman, L.R., and Gleitman, H. (1997). What is language made out of? *Lingua*, *100*, 29–55.

Goksun, T., Kiintay, A.C., and Naigles, L.R. (2008). Turkish children use morphosyntactic bootstrapping in interpreting verb meaning. *Journal of Child Language*, *35*, 291–323.

Graf-Estes, K., Evans, J.L., Alibaldi, M.W., et al. (2007). Can infants map meaning to newly segmented words? *Psychological Science*, *18*, 254–260.

Hamrick, P., Lum, J.A., and Ullman, M.T. (2018). Child first language and adult second language are both tied to general-purpose learning systems. *Proceedings of the National Academy of Sciences*, *115*(7), 1487–1492.

Hespos, S.J., Ferry, A.L., and Rips, L.J. (2009). Five-month-old infants have different expectations for solids and liquids. *Psychological Science*, *20*, 603–611.

Hespos, S.J., and Spelke, E.S. (2004). Conceptual precursors to language. *Nature*, *430*, 453–456.

Hohne, E., and Jusczyk, P.W. (1994). Two-month-old infants' sensitivity to allophonic differences. *Perception and Psychophysics*, *56*, 613–623.

Houston, D.M., Jusczyk, P.W., Kuijpers, C., et al. (2000). Cross-language word segmentation by 9-month-olds. *Psychonomic Bulletin and Review*, *7*, 504–509.

Hurtado, N., Marchman, V.A., and Fernald, A. (2007). Spoken word recognition by Latino children learning Spanish as their first language. *Journal of Child Language*, *33*, 227–249.

Huttenlocher, J., Vasilyeva, M., Waterfall, H.R., et al. (2007). The varieties of speech to young children. *Developmental Psychology*, *43*, 1062–1083.

Izard, V., Sann, C., Spelke, E.S., et al. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences*, *106*, 10382–10385.

Joanisse, M.F., and Seidenberg, M.S. (1999). Impairments in verb morphology following brain injury: A connectionist model. *Proceedings of the National Academy of Sciences*, *96*, 7592–7597.

Jusczyk, P.W. (1997). *The Discovery of Spoken Language*. MIT Press.

Jusczyk, P.W., and Aslin, R.N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, *29*, 1–23.

Jusczyk, P.W., and Bertoncini, J. (1988). Viewing the development of speech perception as an innately guided learning process. *Language and Speech*, *31*, 217–238.

Jusczyk, P.W., Friederici, A.D., Wessels, J.M., et al. (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language*, *32*, 402–420.

Jusczyk, P.W., Hirsch-Pasek, K., Kemler-Nelson, D.G., et al. (1992). Perception of the acoustic correlates of major phrasal units by young infants. *Cognitive Psychology*, *24*, 252–293.

Jusczyk, P.W., and Hohne, E.A. (1997). Infants' memory for spoken words. *Science*, *277*, 1984–1986.

Jusczyk, P.W., Houston, D.M., and Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, *39*, 159–207.

Jusczyk, P.W., and Kemler-Nelson, D.G. (1996). Syntactic units, prosody, and psychological reality during infancy. In J.L. Morgan and K. Demuth (Eds.), *Signal to Syntax* (pp. 389–410). Erlbaum.

Jusczyk, P.W., Luce, P.A., and Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*, *33*, 630–645.

Jusczyk, P.W., Pisoni, D.B., Walley, A.C., et al. (1980). Discrimination of the relative onset time of two-component tones by infants. *Journal of the Acoustical Society of America*, *67*, 262–270.

Jusczyk, P.W., Rosner, B.S., Cutting, J.E., et al. (1977). Categorical perception of nonspeech sounds by 2-month-old infants. *Perception and Psychophysics*, *21*, 50–54.

Kalashnikova, M., Peter, V., Di Liberto, G.M. et al. (2018). Infant-directed speech facilitates seven-month-old infants' cortical tracking of speech. *Scientific Reports*, *8*(1), 1–8.

Kaplan, P.S., Bachorowski, J., Smoski, M.J., et al. (2002). Infants of depressed mothers, although competent learners, fail to learn in response to their own mothers' infant-directed speech. *Psychological Science*, *13*, 268–271.

Kemler-Nelson, D.G. (1999). Attention to functional properties in toddler's naming and problem-solving. *Cognitive Development*, *14*, 77–100.

Kidd, E., Lieven, E.V.M., and Tomasello, M. (2010). Lexical frequency and exemplar-based learning effects in language acquisition: Evidence from sentential complements. *Language Sciences*, *32*, 132–142.

Kirjavainen, M., Theakston, A., Lieven, E., et al. (2009). "I want hold Postman Pat": An investigation into the acquisition of the infinitival maker "to." *First Language*, *29*, 313–339.

Kluender, K., and Kiefte, M. (2006). Speech perception within a biologically realistic information-theoretic framework. In M.J. Traxler and M.A. Gernsbacher (Eds.), *The Handbook of Psycholinguistics* (2nd ed., pp. 153–200). Elsevier.

Kolodny, O., Lotem, A., and Edelman, S. (2015). Learning a generative probabilistic grammar of experience: A process-level model of language acquisition. *Cognitive Science*, *39*(2), 227–267.

Krentz, U.C., and Corina, D.P. (2008). Preference for language in early infancy: The human language bias is not speech specific. *Developmental Science*, *11*, 1–9.

Krueger, C., and Garvan, C. (2019). Cardiac orienting to auditory stimulation in the fetus. *SAGE Open Nursing*, *5*, doi: 2377960819861486.

Krueger, C., Holditch-Davis, D., Quint, S., et al. (2004). Recurring auditory experience in the 28- to 34-week-old fetus. *Infant Behavior and Development*, *27*, 537–543.

Kuhl, P.K., Stevens, E., Hayashi, A., et al. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, *9*, F13–F21.

Kuhl, P.K., Williams, K.A., Lacerda, F., et al. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, *255*, 606–608.

Landau, B., and Gleitman, L. (1985). *Language and Experience: Evidence from the Blind Child.* Harvard University Press.

Lany, J., Gomez, R.L., and Gerken, L. (2007). The role of prior experience in language acquisition. *Cognitive Science*, *31*, 481–507.

Lecanuet, J., Granier-Deferre, C., and Busnel, M. (1988). Fetal cardiac and motor responses to octave-band noises as a function of central frequency, intensity, and heartrate variability. *Early Human Development*, *18*, 81–93.

Lee, J.N., and Naigles, L.R. (2008). Mandarin learners use syntactic bootstrapping in verb acquisition. *Cognition*, *106*, 1028–1037.

Lenneberg, E.H. (1967). *Biological Foundations of Language*. John Wiley & Sons.

Liberman, A.M., Cooper, F.S., Shankweiler, D.P., et al. (1967). Perception of the speech code. *Psychological Review*, *74*, 421–461.

Lieven, E.V.M., Pine, J.J., and Baldwin, G. (1997). Lexically-based learning and early grammatical development. *Journal of Child Language*, *24*, 187–219.

Liu, H., Kuhl, P.K., and Tsao, F. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Psychology*, *6*, F1–F10.

MacWinney, B. (1998). Models of the emergence of language. *Annual Review of Psychology*, *49*, 199–227.

Marchman, V.A., Bates, E., Burkhardt, A., et al. (1991). Functional constraints of the acquisition of the passive: Toward a model of the competence to perform. *First Language*, *11*, 65–92.

Marchman, V.A., and Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, *11*, F9–F16.

Markman, E.M., and Hutchinson, J.E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology*, *16*, 1–27.

Markman, E.M., and Wachtel, G.F. (1988). Children's use of mutual exclusivity to constrain the meaning of words. *Cognitive Psychology*, *20*, 121–157.

Markson, L., and Bloom, P. (1997). Evidence against a dedicated system for word learning in children. *Nature*, *385*, 813–815.

Markson, L., Diesendruck, G., and Bloom, P. (2008). The shape of thought. *Developmental Science*, *11*, 204–208.

Masur, E.F. (1997). Maternal labeling of novel and familiar objects: Implications for children's development of lexical constraints. *Journal of Child Language*, *24*, 427–439.

Matthews., D., Lieven, E.V.M., Theakston, A.L., et al. (2007). French children's use and correction of weird word orders: A constructivist account. *Journal of Child Language*, *34*, 381–409.

McClelland, J.L., and Patterson, K. (2002). Rules or connections in past-tense inflections: What does the evidence rule out? *Trends in Cognitive Sciences*, *6*, 465–472.

McClelland, J.L., and Rumelhart, D.E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General*, *114*, 159–188.

McClure, K., Pine, J.M., and Lieven, E.V.M. (2006). Investigating the abstractness of children's early knowledge of argument structure. *Journal of Child Language*, *33*, 693–720.

McMurray, B. (2007). Defusing the childhood vocabulary explosion. *Science*, *317*, 631.

McMurray, B., and Hollich, G. (2009). Special section: Core computational principles of language acquisition: Can statistical learning do the job? *Developmental Science*, *12*, 365–368.

Mehler, J., Jusczyk, P., Labertz, G., et al. (1988). A precursor of language acquisition in young infants. *Cognition*, *29*, 143–178.

Mills, D.L., Prat, C., Zangl, R., et al. (2004). Language experience and the organization of brain activity to phonetically similar words: ERP evidence from 14- and 20-month-olds. *Journal of Cognitive Neuroscience*, *16*, 1452–1464.

Mirman, D., Magnuson, J.S., Graf-Estes, K., et al. (2008). The link between statistical segmentation and word learning in adults. *Cognition*, *108*, 271–280.

Monroy, C.D., Meyer, M., Schröer, L., et al. (2019). The infant motor system predicts actions based on visual statistical learning. *Neuroimage*, *185*, 947–954.

Morgan, J.L., and Demuth, K. (1996). Signal to syntax: An overview. In J.L. Morgan and K. Demuth (Eds.), *Signal to Syntax* (pp. 1–24). Erlbaum.

Novack, M.A., Brentari, D., Goldin-Meadow, S., et al. (2021). Sign language, like spoken language, promotes object categorization in young hearing infants. *Cognition*, *215*, 104845.

Nurmsoo, E., and Bloom, P. (2008). Preschooler's perspective taking in word learning. *Psychological Science*, *19*, 211–215.

Onnis, L., Truzzi, A., and Ma, X. (2018). Language development and disorders: Possible genes and environment interactions. *Research in Developmental Disabilities*, *82*, 132–146.

Onnis, L., Waterfall, H.R., and Edelman, S. (2008). Learn locally, act globally: Learning language from variation set cues. *Cognition*, *109*, 423–430.

Pan, B.A., Rowe, M.L., Singer, J.D., et al. (2005). Maternal correlates of growth in toddler vocabulary production in low-income families. *Child Development*, *76*, 763–782.

Pardo, J.S., and Remez, R.E. (2006). The perception of speech. In M.J. Traxler and M.A. Gernsbacher (Eds.), *The Handbook of Psycholinguistics* (2nd ed., pp. 201–248). Elsevier.

Pelucchi, B., Hay, J.F., and Saffran, J.R. (2009). Statistical learning in a natural language by 8-month-old infants. *Child Development*, *80*, 674–685.

Pinker, S. (1984). *Language Learnability and Language Development*. MIT Press.

Pinker, S. (1994a). How could a child use verb syntax to learn verb semantics? *Lingua*, *92*, 377–410.

Pinker, S. (1994b). *The Language Instinct*. Harper.

Pinker, S. (1996). *Language Learnability and Language Development (2nd ed.)*. Harvard University Press.

Pinker, S. (2000). *Words and Rules*. Harper.

Pye, C. (1983). Mayan telegraphese: Intonational determinants of inflectional development in Quiche Mayan. *Language*, *59*, 583–604.

Quine, V.O. (1960). *Word and Object*. MIT Press.

Redington, M., and Chater, N. (1997). Probabilistic and distributional approaches to language acquisition. *Trends in Cognitive Sciences*, *1*, 273–281.

Reh, R.K., Hensch, T.K., and Werker, J.F. (2021). Distributional learning of speech sound categories is gated by sensitive periods. *Cognition*, *213*, 104653.

Rosch, E., and Mervis, C.B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*, 573–605.

Sabbagh, M.A., and Baldwin, D. (2001). Learning words from knowledgeable versus ignorant speakers: Links between preschoolers' theory of mind and semantic development. *Child Development*, *72*, 1054–1070.

Saffran, J.R. (2001). Words in a sea of sounds: The output of infant statistical learning. *Cognition*, *81*, 149–169.

Saffran, J.R. (2002). Constraints on statistical language learning. *Journal of Memory and Language*, *47*, 172–196.

Saffran, J.R. (2003). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science*, *12*, 110–114.

Saffran, J.R., Aslin, R.N., and Newport, E.L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*.

Saffran, J.R., Pollak, S.D., Seibel, R.L., et al. (2007). Dog is a dog is a dog: Infant rule learning is not specific to language. *Cognition*, *105*, 669–680.

Saylor, M.M., Sabbagh, M.A., and Baldwin, D.A. (2002). Children use whole-part juxtaposition as a pragmatic cue to word meaning. *Developmental Psychology*, *38*, 993–1003.

Seidenberg, M.S., and Joanisse, M.F. (2003). Show us the model. *Trends in Cognitive Sciences*, *7*, 106–107.

Seidenberg, M.S., and McClelland, J.L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*, 523–568.

Setoh, P., Cheng, M., Bornstein, M.H., et al. (2021). Contrasting lexical biases in bilingual English–Mandarin speech: Verb-biased mothers, but noun-biased toddlers. *Journal of Child Language, 48*(6), 1185–1208.

Skinner, B.F. (1957). *Verbal Behavior*. Appleton.

Smith, L.B., Jayaraman, S., Clerkin, E., et al. (2018). The developing infant creates a curriculum for statistical learning. *Trends in Cognitive Sciences, 22*(4), 325–336.

Song, H., and Fisher, C. (2005). Who's "she?" Discourse prominence influences preschoolers' comprehension of pronouns. *Journal of Memory and Language, 52*, 29–57.

Speer, S., and Blodgett, A. (2006). Prosody. In M.J. Traxler and M.A. Gernsbacher (Eds.), *The Handbook of Psycholinguistics* (2nd ed., pp. 505–538). Elsevier.

Stager, C.L., and Werker, J.F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature, 388*, 381–382.

Stoll, S., Abott-Smith, K., and Lieven, E. (2009). Lexically restricted utterances in Russian, German, and English child-directed speech. *Cognitive Science, 33*, 75–103.

Stone, A., and Bosworth, R.G. (2019). Exploring Infant Sensitivity to Visual Language using Eye Tracking and the Preferential Looking Paradigm. *JoVE (Journal of Visualized Experiments)*, (147), e59581.

Streeter, L.A. (1976). Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. *Nature, 259*, 39–41.

Swingley, D. (2020, December). Conditional probability and word discovery: A corpus analysis of speech to infants. In M. Hahn and S.C. Stoness (Eds.), *Proceedings of the Twenty First Annual Conference of the Cognitive Science Society* (pp. 724–729). Psychology Press.

Swingley, D., and Fernald, A. (2002). Recognition of words referring to present and absent objects by 24-month-olds. *Journal of Memory and Language, 46*, 39–56.

Swingley, D., Pinto, J.P., and Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition, 71*, 73–108.

Tardif, T., Gelman, S.A., and Xu, F. (1999). Putting the "noun bias" in context: A comparison of English and Mandarin. *Child Development, 70*, 620–635.

Tardif, T., Shatz, M., and Naigles, L. (1997). Caregiver speech and children's use of nouns versus verbs: A comparison of English, Italian, and Mandarin. *Journal of Child Language, 24*, 535–565.

Theakston, A.L., Lieven, E.V.M., Pine, J.M., et al. (2002). Going, going, gone: The acquisition of the verb "go." *Journal of Child Language, 29*, 783–811.

Thiessen, E.D., Hill, E.A., and Saffran, J.R. (2005). Infant-directed speech facilitates word segmentation. *Infancy, 7*, 53–71.

Thiessen, E.D., and Saffran, J.R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology, 39*, 706–716.

Tincoff, R., and Jusczyk, P.W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science, 10*, 172–175.

Tomasello, M. (2000). The item-based nature of children's early syntactic development. *Trends in Cognitive Sciences, 4*, 156–163.

Tsao, F., Liu, H., and Kuhl, P.K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development, 75*, 1067–1084.

Waterfall, H.R., Sandbank, B., Onnis, L., et al. (2010). An empirical generative framework for computational modeling of language acquisition. *Journal of Child Language, 37*, 671–703.

Weber, C., Hahne, A., Friedrich, M., et al. (2004). Discrimination of word stress in early infant perception: Electrophysiological evidence. *Cognitive Brain Research, 18*, 149–161.

Werker, J.F., Gilbert, J.H.V., Humphrey, K., et al. (1981). Developmental aspects of cross-language speech perception. *Child Development, 52*, 349–355.

Werker, J.F., and McLeod, P.J. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology, 43*, 230–246.

Werker, J.F., and Tees, R.C. (1983). Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology, 37*, 278–286.

Werker, J.F., and Tees, R.C. (2002). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *25*, 121–133.

Westermann, G., and Jones, S. (2021). Origins of dissociations in the English past tense: A synthetic brain imaging model. *Frontiers in Psychology*, *12*, 2452.

Westermann, G., Ruh, N., and Plunkett, K. (2009). Connectionist approaches to language learning. *Linguistics*, *47*, 413–452.

Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua*, *106*, 23–79.

Yang, C. (2004). Universal grammar, statistics, or both. *Trends in Cognitive Sciences*, *8*, 451–456.

Yuan, S., and Fisher, C. (2009). "Really? She blicked the baby?" Two-year-olds learn combinatorial facts about verbs by listening. *Psychological Science*, *20*, 619–626.