Journal Title: The symbolic species

Volume:
Issue:
Month/Year: 1997
Pages: 47-80

Article Author:

Article Title:

Call #: QP399 .D43 1997

Location: evans

Not Wanted Date: 02/19/2006

Status: Faculty
Phone: 845-2503
E-mail: tya@psyc.tamu.edu

Name: Yamauchi, Takashi

Pickup at Evans

Mail Stop 4235
Dept. of Psychology
College Station, TX 77843
CHAPTER TWO

A Loss for Words

If the only tool you have is a hammer, you tend to treat everything as if it were a nail.

—Abraham Maslow

Gymnastics of the Mind

We often find that an apparently simple task is difficult, not because it is complicated but because it is awkward. Sometimes, for instance, you just lack the right tools for the job. No matter how easy it is in principle to tighten a screw, if it has a slotted head (-) and all you have is a Phillips screwdriver (+), forget it. This is a familiar source of difficulty in physical activities. We humans are unprepared to perform a number of tasks that other species perform with ease. Other species can boast such adaptations as streamlining and fins for swimming, large flexible mobile surfaces for gliding and flying, claws for clinging to the trunks of trees or snagging prey, or sharp canines for tearing flesh. Attempt a task that you are poorly
suited for and your performance is at best inelegant and clumsy, and often it just doesn’t succeed.

Such awkwardness is, in essence, the opposite of preadaptation—that lucky chance of evolution when preexisting body parts coincidentally happen to be predisposed for a novel adaptational challenge. In comparison, this kind of built-in ineptness might be called “pre-maladaptation.” Pre-maladaptation accounts for our difficulties trying to sleep standing up, cut meat without a knife, or distinguish friends from foes by their smell, though certain other species do each well. In the same way that we find certain actions or movements to be impossible unaided, or find ourselves maladroit at manipulating certain objects, certain kinds of mental tasks can also be unwieldy for a brain predisposed to different sorts of analyses. In addition, certain mental predispositions that serve well in some domains can get in the way of accomplishing otherwise trivial tasks that require a new perspective or logic. Cognitive pre-maladaptations might also include predispositions to behave inappropriately, in ways that are opposite to how events tend to occur, or preferences that lead us to pay attention to irrelevant or misleading details. This is, of course, the secret to many magicians’ tricks and the basis for animals’ protective coloration. These could make an otherwise simple task difficult by misdirecting attention and interfering with appropriate actions.

Learning is not any one general process. Learning always occurs in a particular context, involving particular senses and types of motor actions, and certain ways of organizing the information involved. The process of learning is also not just committing an association to memory under the influence of reinforcement. Learning involves figuring out what is relevant and then figuring out how the relevant variables are associated. It involves sorting and organizing, and sometimes recoding what we have already learned. When a pigeon learns that pecking at a red-lighted button will cause food to be delivered and that pecking at a blue-lighted button will shut off an aversive noise, it does not just commit these links to memory; it also learns to ignore a great deal that is irrelevant: the talking in the background, the time of day, the temperature of the room, the pecking of neighboring pigeons, odors that periodically waft by, and the other lights and buttons and structures of the cage. Success or failure at learning and problem solving depends on habits of attention, what we find salient and what we don’t tend to notice, and how easily one can override and reprogram these tendencies.

There are also internal distractions, in particular those that arise from past learning experiences. Interference effects of this sort are a common experience. Old associations provide a sort of repository of hypotheses. In many circumstances in the real world, we produce similar outcomes. General associations can often provide access to patterns of trial-and-error learning. The opportunities to fall back on incorrect responses so that the new can become a practice can be difficult because they are counterintuitive: to infer what’s missing, to construe what is not there, to imaginatively complete the site of what we want to demand of the other. The key is self-imposed. The ways we think are precisely the causes of our failures.

Intelligence tests and “thinking out of the box” ask us to think outside of self-imposed constraints, to see the facility to make leaps of intuition for calling something a noninnovation, rather than an innovation. What is it that makes us see it. This is why some of the fundamental ideas are often too acquainted with the world for us to see. The biologist Thomas Huxley is known for his popularization of Darwin’s theory of natural selection, saying, “Darwin is the best we can do! That’s what I think of it. What is the problem with it?”)

The difficulties that a problem presents in complexity and the learner’s past experiences and other factors. Comparative psychologists, for example, have noted that general learning abilities vary across species. As they have become more sophisticated in understanding a species’ capabilities, they have also been finding correlated differences across species. In these patterns for generation.
inelegant and clumsy, and often opposite of preadaptation—that body parts coincidentally happen upon challenge. In comparison, called “pre-maladaptation.” Pretending to sleep standing up, cut from foes by their smell, though one way that we find certain aced, or find ourselves maladroit acts of mental tasks can also be at sorts of analyses. In addition, well in some domains can get in tasks that require a new portions might also include predispositions that are opposite to how events to pay attention to irrelevant or secret to many magicians’ tricks. These could make an other attention and interfering with

Learning always occurs in a paradigm of motor actions, and involved. The process of learning to memory under the influence of what is relevant and are associated. It involves sorting what we have already learned. Lighted button will cause food lighted button will shut off an e links to memory, it also learns talking in the background, the pecking of neighboring pig other lights and buttons and learning and problem solving salient and what we don’t tend to reprogram these tendencies. particular those that arise from facts of this sort are a common of repository of hypotheses. In many circumstances in the real world, events with similar features will produce similar outcomes. Generalizing from resemblances of present to past associations can often provide useful shortcuts that avoid wasting long periods of trial-and-error learning. Sometimes, however, superficial similarities are just that, and the responses they elicit are off track. If the habits are strong, the opportunities to compare different habits are few, or the feedback about incorrect responses is weak, forgetting the old in order to learn the new can become a practical impossibility. As a result, certain problems can be difficult because they require us to think in ways that are not typical to infer what’s missing, to work back from an end, to assume the opposite of what we want to demonstrate, and so on. Trick questions and riddles stump us, not because the question is complicated but because the solution is counterintuitive. Jokes are endlessly fascinating because they provide an insidiously logical punch line that we are unable to anticipate or predict. A successful punch line cannot be just a non sequitur; it must follow inevitably and obviously, if only after the fact. The minor cognitive implosion summed up by “Now I get it” is in part an experience of admiration at being fooled so well, so simply, and yet so logically. As with a magic trick, the key to a good joke is misdirection; and the most effective misdirection is self-imposed. The ways we naturally think about the subject of the joke are precisely the causes of our capacity to be fooled.

Intelligence tests and “brain teasers” often challenge our capability to think outside of self-imposed and overly narrow contexts, and we consider the facility to make leaps of logic a mark of intelligence. Indeed, the criterion for calling something a work of genius is seldom its complexity, but rather how innovative the approach and how many others were unable to see it. This is why some of the great discoveries in many fields have been made by people outside the inner circle of true experts, because the experts are often too acquainted with a problem to see it in a novel way. As the great biologist Thomas Huxley is said to have exclaimed after learning of Darwin’s theory of natural selection, “How stupid of me not to have thought of that!” Sometimes what we do best works against us.

The difficulties that a problem poses are a function of both its intrinsic complexity and the learner’s predispositions to attend to its most relevant aspects. Comparative psychologists have long struggled to untangle questions of general learning ability from specific learning abilities in different species. As they have become more sensitive to the problem of understanding a species’ capabilities in the context of its natural environment, they are finding correlated differences in learning. Animal breeders have noticed these patterns for generations. Some breeds of dogs make better shepherds,
some make better hunters, some make better Seeing Eye dogs. And there are almost always trade-offs: behavioral predispositions that are well suited to one sort of task often conflict with others. This sort of complementarity is also evident in studies of the effects of brain damage on animal learning. It is not uncommon to find that rather than producing impairment across the board, localized brain damage enhances learning of some tasks at the same time that it impairs learning of others. Such paradoxical enhancement shows that learning is not a monolithic brain process. Learning rates are species- and task-specific, and depend on the subject's particular balance between alternative predispositions. The notion that one can be better or worse at all forms of learning ignores the intrinsically competitive nature of different learning paradigms. The best attentional and mnemonic strategies for one type of task may be exactly wrong for another. Just as our past experiences and accumulated knowledge can in some cases pose impediments to solving a novel problem, so can the evolutionary heritage of a species set it up to handle some cognitive problems well and others poorly. To the extent that members of a species are innately biased to attend to irrelevant details and to ignore critical elements of a problem, they are pre-adapted.

This suggests a very different way to approach the language mystery. Could there be something about the way even a simple language must be learned that is just awkward for other species? Could symbolic reference be naturally counterintuitive?

The fact that language is an unprecedented form of naturally evolved communication, compared to other forms, suggests that it likely requires a rather different problem-solving orientation in order to learn it. The point of the “missing languages paradox” is that the difficulty most species have in learning a language does not appear to be sufficiently diminished, even when the language being taught is vastly simplified. Somehow, they don’t “get” what the language problem is about. This suggests that it is not only difficult but also goes against certain other predispositions that are quite strong. Language may require learning in ways that run counter to other more typical learning strategies. Just getting started, getting what it means for a word or sign to represent something—not simply point to it or bring it to mind by association but symbolize it—may require a kind of mental gymnastics for which most nonhuman brains are pre-adapted. Even with considerable human social experience or specific training and support for their vocal limitations, only a select few nonhuman animals have ever come to grasp the symbolic relationships we have tried to teach them. Instead, most exhibit a remarkable capacity for anticipating our responses, mimicking our actions, or communicating symbolically—sophisticated abilities.

**In Other Words**

How could something as awkward for other species be so familiar to us? The explanation lies in the field where explanations of evolution have been galloping for years and thousands of textbooks, the field of the relationship that involves our capacity. To be blunt, we do not know how we do it, or when we do. Or rather, we do not know what mental processes underlie things we know so well. Yet we do not know how we do what we do. Or rather, we do not know what mental processes underlie things we know so well. The fact that they are solved for so many generations of us is technical difficulty stands in the way; the concept, on the other hand, is simple.

Few topics have generated as much excitement as language. Few, if any, think that this question is how humans began to develop human language origins. Yet we are not fully aware even before getting started, even before framing the most difficult aspects of the question
down. Few topics have generated as much excitement as language. Few, if any, think that this question is how humans began to develop human language origins. Yet we are not fully aware even before getting started, even before framing the most difficult aspects of the question.

Dozens of theories purport to explain the relationship between a word and its meaning. Linguists have been arguing for years that we have approached a new interface between computer technology and the human brain. In this context, I find it is important to take into account the moment when the human brain has been in the lack of counterparts to language in the physical world play a significant role in the way in which the enormously powerful computer interacts with the human brain.
Looking our actions, or committing large sets of paired associations to memory—sophisticated abilities all, but not symbolic abilities.

In Other Words

How could something as simple as word meaning be counterintuitive and awkward for other species to grasp? This question has its own special awkwardness. The explanation of the nature of word meaning has challenged thinkers since before recorded philosophy, and it continues to plague every field where explanations of thought processes are important. Thousands of years and thousands of texts later, we still do not fully understand the basis of the relationship that invests words with their meanings and referential capacity. To be blunt, we do not really understand one of our most commonplace experiences. We know how to use a word to mean something and to refer to something. We know how to coin new words and to assign new meanings to them. We know how to create codes and artificial languages. Yet we do not know how we know to do this, nor what we are doing when we do. Or rather, we know on the surface, but we do not know what mental processes underlie these activities, much less what neural processes are involved. The fact that a conceptual problem this basic has gone unresolved for so many generations suggests that something more than mere technical difficulty stands in the way of our understanding. It is not just a difficult puzzle; the concept seems to be as counterintuitive for us to understand as it is simple for us to use.

Few topics have generated as much debate and confusion. So one might think that this question is hardly the place to choose to begin an analysis of human language origins. Don’t we risk getting stuck in the philosophical mire even before getting started? Yes, but there is something to be said for framing the most difficult and critical questions first, in order to avoid following pointless leads.

Dozens of theories purport to explain the many aspects of the relationship between a word and its meaning. Philosophers, psychologists, and linguists have been arguing for centuries over such matters, and the debates have approached a new intensity in recent years because of the rise of computer technology and the hint of a possibility of building “intelligent” machines. In this context, I find it curious that scholars and engineers fail to take into account the anomalous nature of this form of reference. Shouldn’t the lack of counterparts to words and sentences in the rest of the biological world play a significant role in our thinking about the problem? Of all the enormously powerful computing devices that we find in the heads of
birds and mammals, only one uses a symbolic mode of reference. What hints about its nature have we missed because we ignored its rarity?

The main reason that this seems to have been ignored is that it is generally assumed that other species do exhibit counterparts to words and sentences in their natural repertoires. Ambiguous definitions of what constitutes words, sentences, and languages on the one hand, and reference, meaning, and understanding on the other, allow researchers to stretch metaphorical uses of these terms to fit. Isn’t any family dog able to learn to respond to many spoken commands? Doesn’t my dog understand the word “come” if he obeys this command? There doesn’t seem to be anything special about learning to associate a sound and an object, and isn’t that the whole basis of word reference? Just multiply such associations and add syntax to string them together into different combinations and you have a simple language, right? Not exactly. We think we have a pretty good idea of what it means for a dog to “understand” a command like “Stay!” but are a dog’s understanding and a person’s understanding the same? Or is there some fundamental difference between the way my dog and I understand the same spoken sounds? Common sense psychology has provided terms for this difference. We say the dog learned the command “by rote,” whereas we “understand” it. But this is a notoriously difficult difference to specify. Our uncertainty about this not only makes it difficult to figure out what animals are capable of, and what they are not, it also blurs the distinction between animal communication and language.

One reason we have such difficulty is that we don’t know how to talk about communication apart from language. We look for the analogues of words and phrases in animal calls, we inquire about whether gestures have meanings, we consider the presence of combination and sequencing of calls and gestures as indicating primitive syntax. On the surface this might seem to be just an extension of the comparative method: looking for the evolutionary antecedents of these language features. But there is a serious problem with using language as the model for analyzing other species’ communication in hindsight. It leads us to treat every other form of communication as exceptions to a rule based on the one most exceptional and divergent case. No analytic method could be more perverse. Social communication has been around for as long as animals have interacted and reproduced sexually. Vocal communication has been around at least as long as frogs have croaked out their mating calls in the night air. Linguistic communication was an afterthought, so to speak, a very recent and very idiosyncratic deviation from an ancient and well-established mode of communicating. It cannot possibly provide an appropriate model against which to assess other forms of communication. As a rule, and a quite anomalous one at that, penguins distinguish birds’ wings with respect to characteristics of penguins’ wings. Mammals with respect to other mammals. It is an understandable anthropomorphism or porcupines we might see in phallic form. Compared to our own more ancient and diverse syntax, the fuscate is a better match.

This inversion of evolution is partial. Animal communication as such is rarely an exception. We often hear animal behaviorists say that animals without syntax, or names without reference to the early stages of language development, cannot cope. Of crippled language or child’s syntax, we often say that it is not clear that we consider language in the language in the intact 2 or 3-year-old child’s language development. This is not so. To treat the language of intact 2 or 3-year-old children’s language development as normal is to treat the language of rehabilitation for brain damage as normal. “Normality” is an illusion created by the absence of abnormality.

Treating animal calls and gestures as exceptions to the sequence of evolutionary events is a mistake as well. We know that other animals is self-sufficient and we don’t need languages to acquire or interpret it. This is not limited to species such as dolphins, but also to extensive nonverbal communication systems, such as the use of prosody, pointing, gesturing and other symbols that are common in human communication.
which to assess other forms of communication. It is the rare exception, not the rule, and a quite anomalous exception at that. It is a bit like categorizing birds' wings with respect to the extent they possess or lack the characteristics of penguins' wings, or like analyzing the types of hair on different mammals with respect to their degree of resemblance to porcupine quills. It is an understandable anthropocentric bias—perhaps if we were penguins or porcupines we might see more typical wings and hair as primitive stages compared to our own more advanced adaptations—but it does more to obfuscate than clarify. Language is a derived characteristic and so should be analyzed as an exception to a more general rule, not vice versa.

This inversion of evolutionary logic leads to persistent attempts to model animal communication as some version of language minus something. I often hear animal behaviorists or linguists remarking that "animal languages just lack a grammar and syntax." In short, we analyze animal communication systems as though they were truncated or degenerate languages. Words without syntax, or names without semantics. Moreover, we often imagine the early stages of language evolution as though our ancestors spoke a sort of crippled language or child's language. Serious, well-received theories have suggested that we consider the speech of linguistically impaired brain-damaged adults (Broca's aphasics) or the speech of very young infants as models of the early stages of language evolution. Notice, however, that just by treating other species' communication as partial languages, modern language becomes the implicit end point, as adult language is the end point of children's language development and recovered language is the ideal goal of rehabilitation after brain damage. Of course, this apparent "final causality" is an illusion created by a sort of Orwellian rewriting of the evolutionary past in terms of the present.

Treating animal calls and gestures as subsets of language not only reverses the sequence of evolutionary precedence, it also inverts their functional dependence as well. We know that the nonlanguage communication used by other animals is self-sufficient and needs no support from language to help acquire or interpret it. This is true even for human calls like sobbing or gestures like smiling. In contrast, however, language acquisition depends critically on nonlinguistic communication of all sorts, including much that is intimately presupposed as many nonhuman counterparts. Not only that, but extensive nonverbal communication is essential for providing the scaffolding on which most day-to-day language communication is supported. In conversations, demonstrations, and explanations using words we make extensive use of prosody, pointing, gesturing, and interactions with objects and other people to disambiguate our spoken messages. Only with the historical in-

Terrence W. Deacon 53
vention of writing has language enjoyed even partial independence from this nonlinguistic support. In the context of the rest of communication, then, language is a dependent stepchild with very odd features.

But if language evolved subsequent to hundreds of millions of years of success with these nonlinguistic modes of vocal communication, and even now functions only by virtue of their presence, this should warn against inverting the relationship and treating language as the measure of these other forms. Language did not supersede or replace other forms of communication. Language evolved in a parallel, alongside calls and gestures, and dependent on them—indeed, language and many human nonlinguistic forms of communication probably co-evolved (a dynamic that will be discussed in later chapters). This is demonstrated by the fact that innate call and gesture systems, comparable to what are available to other primates, still exist side by side with language in us. Their complementarity with and distinction from language are exemplified by the fact that they are invariably produced by very different brain regions than are involved in speech production and language comprehension (details of these differences will be reviewed in Chapters 8–10). In many ways, these human counterparts to innate calls and gestures in other species offer the best source of intuitions concerning this difference. Who, for example, would even bother to consider the possibility that somehow smiles, grimaces, laughs, sobs, hugs, kisses, and all the rest of our panhuman nonlinguistic communications were words without syntax?

The popular notion that the calls and gestures constituting the communications of other species are like words and sentences can mostly be traced to misconceptions about the concept of reference. The problem of reference has always been a major topic of debate in the study of animal communication (see Figure 2.1). On the one extreme, some animal behaviorists have argued that calls and gestures are merely external correlates of internal states, and so have no external reference; on the other extreme, some cognitive ethologists have argued that many animal calls, grunts, and gestures should be considered the equivalents of words that name specific objects in the world. One study, in particular, played a central role in rekindling debate on the role of reference in animal communication. In the mid-1980s, Robert Seyfarth, Dorothy Cheney, and their colleagues reported that vervet monkeys produced alarm calls that appeared to act something like names for distinct predators. Their observations demonstrated that distinctly different calls were produced to warn other troop members of the presence of either eagles, leopards, or snakes (an even wider range of calls is now recognized). In response to hearing one of these calls, other troop
Honeybee Recruitment Dance

Returning forager dances for others to communicate direction and distance to food source.

Angle to vertical = angle of flight to food source with respect to sun.

Intensity of waggle indicates distance to food source.

Humpback Whale Song

Song is composed of themes that are altered each year, and sung by all males of the group.

Vervet Monkey Alarm Calls

Distinct calls are given for distinct predators & produce distinct escape behaviors.

Figure 2.1 Three often-cited examples of animal communication systems: honeybee recruitment dance, humpback whale songs, and vervet monkey alarm calls (discussed in the text). Each exhibits features that are similar to features thought to be special to language.
members either raced out of the trees (eagle), climbed into the trees (leopard), or just rose up to peer into the bushes around them (snake). Thus, distinct calls referred to distinct classes of predators, not simply to some state of the caller (though they indicated a fearful state of mind as well).

It is not difficult to reconstruct the evolutionary processes that produced such distinctive calls and referential relationships. The key can be found in the behaviors that the calls induce in other troop members. These predators attack using very different approaches, and the appropriate defense behaviors for each turn out to be mutually exclusive. The worst place to be when a leopard is prowling nearby is on the ground. But since leopards can also climb trees, it is best to wait out on the thinner branches of trees. Unfortunately, this is the worst place to be when eagles are threatening. Best to be hiding under a tree on the ground. Imagine, then, the dilemma if this species had only a single type of alarm call. To ascend or descend, that is the question. Just to sit frozen with indecision or stand up and peer around is the worst response of all (unless a snake is the predator, for which there is another call), since it leaves you vulnerable to both. Consequently, predation selects against individuals that have trouble determining which call is which, and it will select against the kin of animals that do not in some way provide distinct information that helps with this choice, such as sound differences. This evolutionary logic is what is generally termed disruptive selection: selection against the intermediate (compromise) value of a trait and favoring the extremes. The referential specificity of these calls evolved over time, then, as the consequences of warning and escape provided selection pressures that changed the calling and responding predisposition of members of the species. Not surprisingly, a similar evolutionary logic has also shaped the alarm calls of many other species, and other design factors are also often involved, such as the localizability of the sounds themselves.4

Cheney and Seyfarth originally suggested that vervet alarm calls were analogous to “names” for these predators, and might be used in the way we shout the warning: “Fire!” This led many to argue that this system of calls was like a very simple language. Some even suggested it was comparable to the way that infants just beginning to learn to speak will use single words like “juice,” to request a drink, or “doggy,” to indicate that they want to pet a dog, and so on. Such human examples of single word sentences lacking overt syntax (though often with characteristic gestural support) have been called holophrastic utterances, though there is considerable debate about how much potential syntax can be read into these. The core of the alarm call argument is that these calls are different from a pain cry or a grimace, in the same way that words are different. In other words, they refer to something other than the animal’s own pain cries and grimaces directly.

This interpretation implies that many, if not all, other species, more of these distinctive to calls. Perhaps objects, maybe even sounds. Dolphins appear to use distinctive calls. Would an elaborate repertoire of language? Are the calls essentially compelling language? Probably not. As the number of species increase, the number of calls increases, and eventually a grammar is developed. Reference being equivalent to a single word complete in some important way. How they are not the same, though.

Reference itself is not universal across animal communication. In many species it is not even a concept other than itself and other than itself. For example, human laughter is a concept we all have. Laughter is an experience we all participate in. When we laugh, we often turn to analyze both the experience of the joke or the awkward social situation. The laughter points to this. But in the case of some of the characteristic laughter, it did not come from sourness. We laugh at a threat, and so on. It categorizes. The laughter tells us about the class of experiences that an event is part of. If I just heard a great joke.” An event that is not the way words do.

Another difference, such as the way we use words...
climbed into the trees (leop-
around them (snake). Thus, dis-
itors, not simply to some state
state of mind as well).
mary processes that produced
ships. The key can be found in
roop members. These preda-
the appropriate defense be-
lusive. The worst place to be
round. But since leopards can
inner branches of trees. Un-
eagles are threatening. Best
gue, then, the dilemma if this
To ascend or descend, that is
or stand up and peer around
the predator, for which there
 to both. Consequently, pre-
bable determining which call
als that do not in some way
this choice, such as sound dif-
fallen or termed disruptive se-
ompromise) value of a trait and
ity of these calls evolved over
md escape provided selection
finition include the predisposition of mem-
ary evolutionary logic has also
, and other design factors are
of the sounds themselves.3
that vervet alarm calls were
might be used in the way we
argue that this system of calls
suggested it was comparable to
so speak will use single words
icate that they want to pet
single word sentences lacking
( gestural support) have been
is considerable debate about
these. The core of the alarm
from a pain cry or a grimace,
her words, they refer to some-
thing other than the animal’s mental and physical state. This of course in-
cludes an unmentioned assumption that other sorts of calls or gestures like
pain cries and grimaces can’t be referential.

This interpretation implicitly invites us to imagine a species with many
more of these distinctive types of calls, some for foods, some for important
objects, maybe even sounds to identify specific individuals (for example, dol-
pins appear to use distinctive “signature whistles” to identify themselves).
Would an elaborate repertoire of this kind constitute a kind of protol-
ague? Are the calls essentially like a vocabulary? This even suggests a com-
pelling language evolution scenario: Individual calls evolved first, they
increased in number and in variety, they were combined in various ways,
and eventually a grammar and syntax evolved to systematize the patterns of
combinations.4 Unfortunately, the entire house of cards is dependent on call
reference being equivalent to word reference, and this resemblance is not
complete in some important ways. Let’s see if we can be more precise about
how they are not the same.

Reference itself is not unique to language, in fact it is ubiquitous through-
out animal communication. Even a symptom can refer to something other
than itself and other than the state of the body that produces it. Take, for
example, human laughter: a symptom of being in a highly amused state of
mind. Laughter is an excellent example of a human innate call (I will re-
turn to analyze both the evolution and physiology of this call in later chaps-
ters). Like other calls it need not be intentionally produced; it often erupts
spontaneously even when we would rather suppress it, even though it can
also be faked (with variable success) if the social context demands. For the
most part we tend to think of it as a way to work off feelings inspired by a
joke or an awkward social situation. But laughter also refers to things as well.
For example, when someone walks into a room laughing, it suggests that
they probably heard or saw something funny just outside, before entering.
The laughter points to this something else that caused it. And it specifies
some of the characteristics that this cause probably exhibits; specifically, it
was not a source of sorrow, not a disgusting or repulsive scene, not a real
threat, and so on. It categorizes the event that induced it by virtue of what
the laughter tells us about the state of the laugh. It points to a definite
class of experiences that are deemed funny. But notice how different the
reference to the same event is when the person stops laughing and says, “I
just heard a great joke.” Alarm calls refer to objects the way laughter does,
not the way words do.

Another difference, subsequently noted by Cheney and Seyfarth,5 has
to do with the way we use words and sentences to transmit information from
one to another. Specifically, they were interested in determining whether these calls were used intentionally to communicate information or just incidentally communicated information, and how this might relate to what individuals know about what others know and don’t know.

Some critical features that distinguish automatic (unintended) forms of communication from intentional communications are characteristic of laughter. Laughter provides others with information about the laughers’ state of mind and recent history, but it also exerts a more direct effect, a sort of compulsion to laugh along. We often acknowledge this by saying that laughter is contagious. Sitting in a room full of laughing people, one finds it difficult not to laugh as well, even though the reason for their laughter may not be entirely clear. Indeed, so strong is this odd compulsion, that artificial laughter produced by a mechanical device (e.g., a laugh box or the laugh track on a TV sitcom) can induce us to laugh, even though we are fully aware that it is not real laughter. To add insult to injury in our sense of self-control, the faked laughter actually induces us to experience things as funnier as a result. This involuntary power of laughter is shared by many other innate social signals as well, including sobbing, smiling, grimacing, etc., and contrasts sharply with the absence of such an echoic tendency in normal language communication. Not only do we seldom parrot what we have just heard from another, such a response is generally oddly annoying, as most children at some point discover to their glee and their siblings’ and parents’ distress. How odd and unnatural it would feel to enter a room where people were echoing each other’s speech in the same way that they tend to echo each other’s laughter? This may be why certain ritual practices that employ such patterns of language use are at the same time both disturbing and powerful, depending on whether one feels included or excluded.

In general, even casual conversation takes a certain degree of conscious effort and monitoring. Following another’s speech takes at least a modicum of attention and intentionally controlled analysis—something that quickly becomes obvious when one conversant’s mind begins to wander. Part of this derives from the fact that what one says is typically influenced to some degree by assumptions about what the other already knows. So a common factor in the use of language is an intention to convey something that the other person presumably doesn’t know. The influential philosopher H. P. Grice has even argued that a sort of reflexive logic of the form “I believe that you believe that I believe x” is an essential component of communicating meaning in language. But in this regard, both laughter and vervet monkey alarm calls ultimately fail the test in part because both are involuntary and contagious. When one vervet monkey produces an alarm call, others in the troop both escape and echo the cause of the general excitement is gone.

Reference is not the difference that refers to things in the world, there is a difference. This difference is understood about the nature of communication. It is a difference in different forms of referencing, essential versus nonreferential, and modes of reference may differ in the ways. We cannot hope to be different forms of human communication, human language and other systems. We may clear up this confusion and create.

The Reference Problem

So what is the difference? the way a vervet monkey alarm call thing else? Word meaning has been so simple and yet so elusive to be no more than a mapping between a sound or conventional setting and an object, process, or state (Figure 2.2). How the thing the signer is thought to differ between words and be the arbitrariness and convention instead probing into these relations more to it.

Precisely identifying this
ted in determining whether the information is just in or out. This might relate to what I don't know.

Atactic (untended) forms of communication are characteristic of ordinary attention about the laughter's effect: a more direct effect, a knowledge this by saying that if laughing people, one finds the reason for their laughter this odd compulsion, that arrises (e.g., a laugh box or the hugh, even though we are fully to injury to our sense of self, to experience things as funnier is shared by many other smiling, grimacing, etc., and echoic tendency in normal human parrot what we have just oddly annoying, as most of their siblings' and parents' to enter a room where people way that they tend to echo ritually practices that employ me both disturbing and pow- ered or excluded.

A certain degree of conscious edge takes at least a modicum insight—something that quickly begins to wander. Part of this is functionally influenced to some degree. So a common fancy something that the other the nal philosopher H. P. Grice the form "I believe that you cut of communicating meaning and vervet monkey alarm both are involuntary and con- alarm call, others in the troop

both escape and echo the call again and again, and all will join in until the general excitement is gone. This redundancy provides a minimum of new predator information for any troop members subsequent to the first presentation (though it may indicate that danger is still present). It almost certainly has more to do with communicating and recruiting arousal than predator information, as is analogously the case with laughter and humor. Knowing that another saw a predator does not appear to inhibit the vervet's tendency to call and knowing that another also got a joke does not apparently inhibit a person's tendency to laugh—indeed, both may be potentiated by this knowledge.

So, in this regard, the example of vervet alarm calls offers a false lead. Reference is not the difference between alarm calls and words. Both can refer to things in the world and both can refer to internal states, but there is a difference. This difference is the source of the most common misunderstanding about the nature of linguistic versus nonlinguistic communication. It is a difference in the kind of reference. We tend to confuse different forms of reference with one another or else dichotomize referential versus nonreferential communication, instead of recognizing that modes of reference may differ and may depend on one another in complicated ways. We cannot hope to make appropriate comparisons between different forms of human communication, much less comparisons between human language and other species' forms of communication, unless we can clear up this confusion and discover what exactly constitutes this difference.

The Reference Problem

So what is the difference between the way a word refers to things and the way a vervet monkey alarm call, a laugh, or a portrait can refer to something else? Word meaning has always fascinated people because it is at once so simple and yet so elusive in the way it works. On the surface it seems to be no more than a mapping or pairing between one thing and another—a sound or conventional set of markings (the "signifier") on the one hand, and an object, process, or state of things (the "signified") on the other (see Figure 2.2). How the thing signified is brought into correspondence with the signifier is thought to distinguish different forms of reference. The difference between words and other means of referring to things appears to be the arbitrariness and conventionality of the linguistic link. But a little further probing into these relationships demonstrates that there must be more to it.

Precisely identifying this difference has probably nowhere been more
Presumably, when he wants to make us laugh. Should we now say that he means "Enter!"

Refering to the interpretation of words as a Procrustean lig.

Though words and sentences have the property of saying that laughter has a meaning, a person's laughter indicates some lack of context. We would only construe that metaphorical looseness of phrase as a "prearranged signal" if we were also told that it means "Enter!"

The late nineteenth-century philosopher Gottlob Frege provided a concise distinction between the referential and denotational aspects of word meaning. He distinguished between "sense" and "reference." Its sense is the idea - the content - that one might associate with a particular word or phrase: Frege calls this the "sense" or "extension" of a word. Reference, by contrast, is the entity or concept to which a word refers. Frege distinguishes between the "meaninglessness" of a word and its "meaning." Frege offers the example of the word "star," which refers to the same object in different senses. In this case, "star" has two different senses in different contexts.

Numerous philosophers have provided insightful perspectives on reference (extension). Frege's distinction between sense and reference (extension) is widely accepted. However, the word "cracker" refers to results from a parrot's behavior. If we ask questions: what do we mean by "cracker"? Indeed, Frege's definitions of words may play some role. But how do we determine reference? Can we list all the possible meanings, or are we left with a vague idea of the referential power of words?
Presumably, when he wants a cracker, he will say so. Is this different? Should we now say that he knows what the words mean?

Referring to the interpretation of a call or gesture as its “meaning” again betrays a Procrustean linguistic bias in our analysis of communication. Though words and sentences have meanings, there is something odd about saying that laughter has a meaning. A more accurate statement is that a person’s laughter indicates something, both about the person and about the context. We would only say it means something (except in a sort of metaphorical looseness of phrase) under rather special circumstances, when it is used as a prearranged signal; e.g., for someone to enter the room. Here too, we would tend to say that it indicates the right moment to enter, not that it means “Enter!”

The late nineteenth-century mathematician-philosopher Gottlob Frege provided a concise distinction between these two often-confused aspects of word meaning. He distinguished between the sense of a term and its reference. Its sense is the idea one has in mind that corresponds with considering a particular word or phrase. This is distinguished from the reference of the same word or phrase, which is something in the world which corresponds with this term and its sense. The logic of this distinction has influenced most subsequent theories and is only slightly redefined by such complementary terms as intension (as distinct from intention) and extension, respectively, in later philosophical discussions. One might summarize this as “sense is something in the head” and “reference is something in the world.” Frege offers the example of “the morning star” and “the evening star,” which refer to the same physical object, the planet Venus, but have different senses. In this case, these different senses trace back to a historical context where this common reference was not recognized.

Numerous philosophers of language and mind have built on this original insight. One classic view argues that sense (intension) is used to determine reference (extension). This fits the common sense insight that what a word refers to results from the idea that it invokes. This invites the obvious questions: what do we mean by idea in this context, and how does this determine reference? Candidate interpretations of “idea” include mental images, stimulus-stimulus associations, something like a dictionary definition or encyclopedia entry, or checklists of “features” or qualities of objects. All may play some role. But how such mental objects pick out external physical objects, or whether they do, remains a thorny philosophical problem.

A number of critics of this classic conception have demonstrated that the referential power of words can be otherwise independent of their sense in
many cases. They point out that there are cases (particularly in the example of names, which have somewhat minimalistic senses) where we discover that the sense of a term does not correspond to the reference we all along accepted. For example, if we were to discover that William Shakespeare was a pseudonymous proxy for Sir Francis Bacon, who actually wrote the famous plays and sonnets (a claim that has been argued by some), it would not change the reference of either name to a particular historical figure, just statements made about the men the names refer to. Other philosophers have employed more exotic examples to demonstrate this same point because they show how radical falsifications of meaning still do not change reference. For example, imagine someday discovering that all mosquitoe were not animals but really stealth devices designed by extraterrestrials for obtaining human DNA samples. This would not undermine the reference of the term mosquito. The proponents of this view explain it by arguing that the reference in such a case was previously determined in a more concrete and causal fashion by the correlation in space and time of the use of the word with the presence of the physical objects to which it refers. The word and the object must have occurred at least at some point in the past, and all modern uses derive their reference by virtue of an unbroken causal-historical link to some such reference-establishing event or events. The ideas we entertain about this link between a name and what it names may thus help perpetuate and refine its use, but they are not its crucial determinants.

It is interesting, however, that a more subtle version of this logic can probably account for the evolution of a link between alarm calls and predators and between laughter and humorous experiences, via biological evolutionary history. This demonstrates that reference in general does not require some conscious concept or meaning to determine it. But whereas a smile may exhibit only this aspect of reference, it seems that words exhibit this and more. In other words, reference may have a sort of hierarchic character. The way we so often need nonverbal gestures to get across the meaning of things or to explain what we have in mind is another reflection of this dependence of word reference on more basic forms of reference.

A more complicated terminology is necessary, then, to begin to differentiate between the way that words, as opposed to laughter and other non-language signs, refer to things. We need terms that cut beneath word reference and from which word reference can be derived as a special case, since that is the way it evolved and the way it develops in each of us. Words are not just sounds, configurations of ink on paper, or light on a computer screen. What endows these otherwise inanimate things with the capacity to refer to other things is an interpretive process, a critical part of which (though not all) occurs "in" the appropriate symbolic interpretant, which only use physical co-occurrence to be relevant to what is being said to get one sort of reference from something more is the result of some mental act of the person. So maybe it does refine the distinction.

Ultimately, reference is a "proof by elimination"; it is created by the process of generating a response; and differences in the form that those responses take will enable one to infer that there are different references for the same stimulus. We can refer to the two Peircean devices following the terminology of philosopher Charles Sanders Peirce, who recognizes three degrees of complexity but recognizes that we have, as well; moreover, he did not put his name on the head. Whatever process is important. The problem is to explain how these different kinds of reference apply to the interpretants required for language.

So, what are some of the apparent mechanisms of reference? One mechanism is the mental image of some kind of a familiar dog for "dog." Though once treated as a symbol of a thing, imagery has in recent years been shown to play a role in clear neural and behavioral associations on an imagined object, or in patterns of response. Other changes, can have a role in the effort required to consider the case. And things to be involved. But a mental symbol is not the most important case. A word also might bring to mind some specific behavior, or it even might
es (particularly in the exam-

istic senses) where we discover-
to the reference we all along
that William Shakespeare was
who actually wrote the famous
quest by some), it would not

cular historical figure, just
refer to. Other philosophers
onstrate this same point be-
meaning still do not change

discovering that all mosquitoes
signed by extraterrestrials for
not undermine the reference
view explain it by arguing that
determined in a more concrete
and time of the use of the word
which it refers. The word and
same point in the past, and all
true of an unbroken causali-

ing an event or events. The ideas
and what it names may thus
not be its crucial determinants.

version of this logic can pro-
en alarm calls and predators
ences, via biological evolution-

e in general does not require
mine it. But whereas a smile
seems that words exhibit this
 provide a sort of hierarchic charac-
tures to get across the mean-
nd is another reflection of this
f forms of reference.

sary, then, to begin to differ-
d to laughter and other non-
tems that cut beneath word
be derived as a special case,
develops in each of us. Words
paper, or light on a computer
ate things with the capacity to

ess, a critical part of which

(though not all) occurs “in the head.” If I am unable to produce an appro-
appropriate symbolic interpretive response to an unfamiliar foreign word, I can
only use physical co-occurrence to guess at what objects or events might be
relevant to what is being said. Clearly, what distinguishes a dog’s ability to
get one sort of reference from a phrase and a person’s to get that and some-
thing more is the result of something additional that is produced in the head
of the person. So maybe taking this approach to the problem can help de-
fine the distinction.

Ultimately, reference is not intrinsic to a word, sound, gesture, or hi-

gerophy; it is created by the nature of some response to it. Reference de-

rives from the process of generating some cognitive action, an interpretive
response; and differences in interpretive responses not only can determine
different references for the same sign, but can determine reference in dif-
f erent ways. We can refer to such interpretive responses as interpretants,
following the terminology of the late nineteenth-century American philoso-
pher Charles Sanders Peirce. In cognitive terms, an interpretant is what-
ever enables one to infer the reference from some sign or signs and their
context. Peirce recognized that interpretants can not only be of different
degrees of complexity but they can also be of categorically different kinds
as well; moreover, he did not confine his definition only to what goes on in
the head. Whatever process determines reference qualifies as an inter-
pretant. The problem is to explain how differences in interpretants produce
different kinds of reference, and specifically what distinguishes the inter-
pretants required for language.

So, what are some of the interpretants of words? Probably the most
common view of word meaning is that a word is interpreted when one gen-

erates a mental image of something that it refers to: for example, an image
of a familiar dog for “dog” or of someone throwing a baseball for “pitch.”

Though once treated as a sort of fairy tale of introspectionism, mental ima-

ger has in recent years become recognized as an experience that has
clear neural and behavioral correlates. The relative “locations” of features
on an imagined object, or its size, shape, and movement in imagination, or
other changes, can have a direct effect on such factors as the time and ef-
fort required to consider these features and what parts of the brain might
be involved. But a mental image (or the neural process that constitutes it)
is only one sort of interpretive response that a word might elicit, and it may
not be the most important one.

A word also might bring to mind something like a dictionary definition,
or another word that has a related meaning, or it might induce us to act out
some behavior, or it even might produce a vague visceral feeling correlated
with past experiences of what is referred to. All of these are interpretants, but the way they bring a particular word-reference relationship into being can be quite diverse, and of course many can be present simultaneously. The kind of interpretive response determines the nature of the reference relationship. The interpretant is the mediator that brings a sign and its referent together. Differences in the form of reference are due to differences in the form of this mediation process.

Though producing a mental image may be an inevitable product of comprehending certain words, it is not the mediator that distinguishes symbolic reference. A mental image may also be the primary interpretive action in numerous nonsymbolic processes of reference. For example, smelling the scent of a skunk while walking in the woods might bring to mind the mental image of a skunk, as it might also do for a dog who previously had a nasty skunk experience. A more experienced dog or dog owner might even find that the odor produces a revulsion response, which, like the mental image, could also contribute to the reference relationship. In this way the scent refers to the animal by virtue of a mental image. For both the human and the dog the scent would have acquired this reference from past skunk experiences. To a naive puppy or child, however, with no past experience to base an interpretation on, it is just a strong odor. Simply reading these words, of course, might also bring to mind a visual image or a skunk "odor image." This emphasizes that it is not the thing that is ultimately brought into reference, nor even the common images that are elicited, that determines the difference in the ways that words and odors can derive their reference; it is rather how these responses are produced. In the case of these words, there is something more, some additional interpretants that are crucial to their symbolic capacity, and a great deal of additional learning has intervened to make their production possible.

The symbolic basis of word meaning is mediated, additionally, by the elicitation of other words (at various levels of awareness). Even if we do not consciously experience the elicitation of other words, evidence that they are activated comes from priming and interference effects that show up in word association tests. Words referring to abstract qualities, such as "justice," "false," and "peculiarity," that don't easily lend themselves to imagery, may produce word association effects that are just as robust as more concrete words. But there are function words for which we seem unable to think up either kind of interpretant. Words such as "that," "which," and "what" function to point to other words and phrases, but not to specific categories of meanings, and don't evoke mental images. Nevertheless, they produce certain expectations about the grammatical structure of what is to follow.

that we recognize when they use these interpretants, "meanings" in that they are functionally equivalent. All of interpreters that are producing the same forms of imagery and yet have clear and interpreters responses were learning. "Wanna cracker!" where not learned at all? What mechanisms produce different patterns of meaning referred to may be the same. To predict the nature of the meaning of its forms of reference, we need to produce the competence to interpret.

It is probably possible to teach complicated arbitrary sign systems that are appropriate to its sensorimotor capacities to individuals both to produce and to respond to in an appropriate behavior, and to comprehend the intentions of the animals, the responses that were in skeletal form, the recipients. As can behaviorist B. F. Skinner, although it has been challenged by the syntax, many still implicitlyally believe that by training in a study involving pigeons who perch on one cage of the other, the Harvard colleagues demonstrated that the animals learned learning tasks for multiple simultaneously enabled communication with limited sharing. In adjacent cages, he set up responses in each so that on one side the was not able to access it directly. Information transmitted to the second by the first and this could make the food acquired them the crucial information the animals could easily imagine compliance signals. Herrnstein offered the language researchers at the time.
All of these are interpretants, or reference relationships, that can be present simultaneously. The nature of the reference mediator that brings a sign and its reference are due to differences that we recognize when they are violated. Though we hesitate to call these interpretants “meanings” in the same sense as for common nouns and verbs, they are functionally equivalent. Finally, consider the complicated mixtures of interpretants that are produced in response to whole phrases, sentences, and larger narratives or arguments. These can be too abstract to elicit clear imagery and yet have clear meanings. So what does it matter if the interpretive responses were learned and reproduced by rote, like the parrot saying, “Wanna cracker!” when it’s hungry, or learned in some other way, or not learned at all? What matters is that these learning differences will produce different patterns of mental action, so to speak, and although what is referred to may be the same, this difference in the interpretive process will dictate the nature of the referential link that results. So, to distinguish forms of reference, we need to understand the learning processes that produce the competence to interpret things differently.

It is probably possible to train almost any intelligent mammal to use a complicated arbitrary sign system, so long as the medium of expression is appropriate to its sensorimotor abilities. All that is necessary is to train individuals both to produce certain behaviors under specific stimulus circumstances and to respond to these same signals produced by others with an appropriate behavior, and so on. Depending on the mnemonic limitations of the animals, the repertoire could become arbitrarily large. This is, in skeletal form, the recipe for learning language that the famous American behaviorist B. F. Skinner imagined more than fifty years ago, and although it has been challenged as inadequate to produce grammar and syntax, many still implicitly conceive of word reference this way. Recently, in a study involving pigeons in separate training cages linked electronically so that the responses of the one pigeon could be registered as signals in the cage of the other, the Harvard psychologist Richard Herrnstein and his colleagues demonstrated that one could set up a pattern of linked associative learning tasks for multiple subjects so that the resultant behaviors resembled communication with learned arbitrary signs. 10 Using pigeons trained in adjacent cages, he set up the relationships between the stimuli and the responses in each so that only one bird got a signal for food availability but was not able to access it directly. Instead, its response would become a signal transmitted to the second bird in the other cage, who by responding to this could make the food accessible to both. In the end, the first bird transmitted the crucial information to the second bird via an arbitrary code. One could easily imagine complicating the system to include more subjects and signals. Herrnstein offered this experiment as a challenge to primate language researchers at the time who had shown similar interindividual lan-
guage-like communication in chimpanzees (though it is not clear to me whether he intended to show that this form of communication wasn’t equivalent to language or rather that linguistic reference was just this simple).

To me, this experiment demonstrates the simplicity and mechanical nature of this form of reference. And how its key features—learned associations, arbitrariness, reference, and transmission of information from one individual to another—are not sufficient to define symbolic reference. Any bright undergraduate could write a short computer program or build a simple mechanical device to stand in for one of the pigeons. Nevertheless, a system of dozens of signals arranged in such interlocking relationships to one another, and with respect to events and objects of interest to all, could be a powerful communication system. Probably a significant fraction of the communications used by many highly social animals are either partially or wholly dependent upon use of signs in this way. It doesn’t matter whether they are learned and arbitrary, as were those used by the pigeons, or innate and physically linked to some state of arousal. Many animal groups in the wild exhibit regionally variable social behaviors and displays for communication, much of which may be learned and passed from individual to individual by mimicry and association. But such a system is not just words without syntax.

There is something mechanical about innate calls as well as behaviors learned by rote. In our own experiences of learning, we have a sense for the difference between what we have learned by rote and what we say we “understand.” At various stages of learning mathematics, we often find ourselves manipulating symbols and numbers according to certain specific instructions, and although we come up with the correct answer if we follow the instructions exactly, in the end we know what we did without knowing what we did. We were unclear on the concept. Actually, this experience is becoming more common for me as a result of using computers and calculators, which have now become my necessary prostheses. I type in a bunch of numbers and select a few computational operations and a string of values and graphs comes back. At one time I knew why certain operations produced the results that I now produce with a few keystrokes—I learned them the hard way, by rote, until I figured out the significance—but much of that support has faded in my memory. I am left with knowing that pushing certain buttons in certain orders does what I need, and have stopped worrying why.

This is the same intuition we have about words. Kids, trying to impress their friends (or scholars trying to impress their colleagues), may repeat a technical phrase they have heard in conversation, without really knowing what they mean. It often will work, although it asks too many questions, and learning about new meanings involves knowing five or ten more than you really need to change the superficial meaning or significance. The problem of remembering a new meaning or significance, a radical change in cognition.

Many have suggested that language has an arbitrary quality. Innate calls are not arbitrary, but learned vocalizations are. There are different external stimuli. For instance, an alarm call or a laugh differs from that which is the sound of a specific emotional state. The alarm call is not necessarily linked to the vuvuzela alarm call repeated by a dancing gorilla to escape and fear responses. It shows that their relationship is just a matter of degree. Both are relations of events. One is an imitation.

But there is a sense in which this is different in all the nonlanguage experiences. An important clue. If the parrot says “Wanna cracker?” or the dog pushes the kitchen knob, eventually both will learn to start going out the door to eat. “I,” I suspect it would be my interpretation. And if the vuvuzela in Africa, then future evolution would appear from their repertoire (hence purpose). All rely on a relative system to refer.

This is not true of words, which failed to correspond in some cases. It would be of little use, but...
what they mean. It often works if the context is right, and so long as no one asks too many questions, but the application is quite limited. One way we learn about new meanings is to figure out what the right contexts are; but knowing five or ten more contexts in which the same phrase works does not really change the superficial nature of the reference. Learning more and more appropriate contexts does not in itself constitute understanding the meaning or significance. Yet, when we know what the phrase means, the problem of remembering all the applicable contexts becomes irrelevant, and innumerable novel contexts can be immediately recognized as appropriate. In between these alternatives there is not just a quantitative increase, but a radical change in cognitive strategy.

Many have suggested that the key to this flexibility of word reference is arbitrariness. Inmate calls and gestures have some features built in from birth, but learned vocalizations and movements can be freely associated with different external stimuli. For this reason, we might argue that what makes an alarm call or a laugh different and limited in its referential ability is the fact that there is a built-in association between the production of these calls and a specific emotional state. However, there is a sense in which even an alarm call isn’t necessarily linked to its referent. As in a learned association, each vervet alarm call repeatedly co-occurred with a distinct class of predator and escape and fear responses during evolution. The apparent inflexibility of their relationship is just a momentary stage in evolution. The difference between this link and one based on an arbitrarily learned behavior is only a matter of degree. Both are, in one sense, internalizations of external correlations of events. One is built in before birth, one after birth.

But there is a sense in which a degree of necessary association is involved in all the non-language examples we have discussed, and this is an important clue. If the parrot stopped getting fed when he squawked the words, “Wanna cracker!” or the dog stopped being let out when he nosed the door-knob, eventually both would probably stop producing these signs. If I started going out the door every time I said to my dog, “Do you want to eat?”, I suspect it wouldn’t take too long for him to reverse his old habits of interpretation. And if the vervet monkeys’ predators disappeared from Africa, then future evolution would see vervet monkey alarm calls disappear from their repertoire (or perhaps become coopted for some other purpose). All rely on a relatively stable correlation with what they refer, in order to refer.

This is not true of words. Or not in the same way. If our use of words failed to correspond in some way with things in the rest of the world, they would be of little use, but there is something rather odd about this corre-
spondence when we compare it to each of the examples above. If my use of the word “skunk” to refer to a certain animal was sustained by this crit-
ter being present, even a small percentage of the times that I used the word (in other words if there had to be a physical correlation), then the associa-
tion would have been extinguished long ago. A learned association will tend
to get weaker and weaker if some significant degree of co-occurrence of
stimuli is not maintained. I very seldom find myself in the company of
members of this species, if I can help it, and yet I read and talk about them
often. Despite this, I don’t have the impression that the strength of the re-
ferential link between the animals and the name is any less strong than that
between the word “finger” and my flesh-and-blood finger, which is always
present. There is some kind of word-object correspondence, but it isn’t
based on a physical correlational relationship.

To understand this difference, then, we need to be able to describe the
difference between the interpretive responses that are capable of sustain-
ing associations between a word and its reference, irrespective of their
being correlated in experience, and those rate associations that are estab-
lished and dissolved as experience dictates. When we interpret the mean-
ing and reference of a word or sentence, we produce something more than
what a parrot produces when it requests a cracker or what a dog produces
when it interprets a command. This “something more” is what constitutes
our symbolic competence.

Alice laughed.

impossible things

“I daresay you sometimes I’ve been
breakfast.”

The Hierarchical Nature of

The assumption that vice versa is the basic
the work of the turn-
sure. In his widely influ-
language), he argued the by-element mapping be-
stituting the plane of the
view, the mapping of v
considered a signer-signifi
d word reference? Although
of the examples above. If my use of animal was sustained by this critique of the times that I used the word (real correlation), then the association. A learned association will tend to a significant degree of co-occurrence of a find myself in the company of and yet I read and talk about them. The measure that the strength of the referential name is any less strong than that sand-blood finger, which is always absent correspondence, but it isn’t ship.

We need to be able to describe the ones that are capable of sustaining reference, irrespective of their rote associations that are established. When we interpret the mean- we produce something more than a cracker or what a dog produces “nothing more” is what constitutes

CHAPTER THREE

Symbols Aren’t Simple

Alice laughed. “There’s no use trying,” she said. “one can’t believe impossible things.”

“I daresay you haven’t had much practice,” said the Queen.

“When I was your age I always did it for half-an-hour a day. Why, sometimes I’ve believed as many as six impossible things before breakfast.”

—Lewis Carroll, Alice Through the Looking-Glass

The Hierarchical Nature of Reference

The assumption that a one-to-one mapping of words onto objects and vice versa is the basis for meaning and reference was made explicit in the work of the turn-of-the-century French linguist Ferdinand de Saussure. In his widely influential work on semiology (his term for the study of language), he argued that word meaning can be modeled by an element-by-element mapping between two “planes” of objects: from elements constituting the plane of the signifiers (e.g., words) to elements on the plane of the signified (the ideas, objects, events, etc., that words refer to). On this view, the mapping of vervet monkey alarm calls onto predators could be considered a signifier-signified relationship. But how accurately does this model word reference? Although it is natural to imagine words as labels for ob-
jects, or mental images, or concepts, we can now see that such correspondences only capture superficial aspects of word meaning. Focusing on correspondence alone collapses a multileveled relationship into a simple mapping relationship. It fails to distinguish between the rote understanding of words that my dog possesses and the semantic understanding of them that a normal human speaker exhibits. We also saw that the correspondence of words to referents is not enough to explain word meaning because the actual frequency of correlations between items on the two planes is extremely low. Instead, what I hope to show is that the relationship is the reverse of what we commonly imagine. The correspondence between words and objects is a secondary relationship, subordinate to a web of associative relationships of a quite different sort, which even allows us reference to impossible things.

In order to be more specific about differences in referential form, philosophers and semioticians have often distinguished between different forms of referential relationships. Probably the most successful classification of representational relationships was, again, provided by the American philosopher Charles Sanders Peirce. As part of a larger scheme of semiotic relationships, he distinguished three categories of referential associations: *icon*, *index*, and *symbol*. These terms were, of course, around before Peirce, and have been used in different ways by others since. Peirce confined the use of these terms to describing the nature of the formal relationship between the characteristics of the sign token and those of the physical object represented. As a first approximation these are as follows: icons are mediated by a similarity between sign and object, indices are mediated by some physical or temporal connection between sign and object, and symbols are mediated by some formal or merely agreed-upon link irrespective of any physical characteristics of either sign or object. These three forms of reference reflect a classic philosophical trichotomy of possible modes of associative relationship: (a) similarity, (b) contiguity or correlation, and (c) law, causality, or convention. The great philosophers of mind, such as John Locke, David Hume, Immanuel Kant, Georg Wilhelm Friedrich Hegel, and many others, had each in one way or another argued that these three modes of relationship describe the fundamental forms by which ideas can come to be associated. Peirce took these insights and rephrased the problem of mind in terms of communication, essentially arguing that all forms of thought (ideas) are essentially communication (transmission of signs), organized by an underlying logic (or semiotic, as he called it) that is not fundamentally different for communication processes inside or outside of brains. If so, it might be possible to investigate the logic of thought processes by studying the sign production and communication.

To get a sense of this logical trichotomy, consider some examples. When we say something like "no" we mean that there is a resemblance. The pictures of all kinds are icons. When we say that something is an "index" we mean that it is causally or associated with it in some way. The temperature of water, a weather report, or disagreeable odor might indicate an animal's physiology. When we say that something is a "symbol," we mean there is an explicit code which establishes a relationship to something else. A wedding ring symbolizes a particular relationship in English, what should be done to this person.

No particular objects are being used to interpret a word. But if we interpret the word as a symbol, we are engaging in a relationship that is designed to be interpreted that way. So to think of a particular sculpture an icon, one brings it to mind, inferring that there is a relationship of similarity. When we are recognizing a similarity. As Peirce pointed out, almost anything depending on the vagueness of the word.

The same point can be made about the different relationships: neither one is more essential relationship: neither one is more fundamental than the other. However, they are different.
now see that such correspondence in referential form, distinguished between different classes of referential associations: icons, indices, and symbols. We also saw that the correspondence between items on the two planes is that the relationship is the result of a web of associative reasoning between words and concepts. This allows us to interpret the meaning of a sentence in terms of its referential content.

Inference: The logic of signs

To get a sense of this logic of signs, let’s begin by considering a few examples. When we say something is “iconic” of something else we usually mean that there is a resemblance between the two. For example, a thermometer indicates the temperature of water, a weather vane indicates the direction of the wind, and a light switch indicates the presence of electricity. Most forms of animal communication have this quality, from pheromonal odors (that indicate an animal’s physiological state or proximity) to alarm calls (that indicate the presence of a predator). Finally, when we say something is a “symbol,” we mean there is a social convention, tacit agreement, or explicit code which establishes the relationship that links one thing to another. A wedding ring symbolizes a marital agreement; the typographical letter “e” symbolizes a particular sound used in words (or sometimes, as in English, what should be done to other sounds); and taken together, the words of this sentence symbolize a particular idea or set of ideas.

No particular objects are intrinsically icons, indices, or symbols. They are interpreted to be so, depending on what is produced in response. In simple terms, the differences between iconic, indexical, and symbolic relationships derive from regarding things either with respect to their form, their correlations with other things, or their involvement in systems of conventional relationships.

When we apply these terms to particular things, for instance, calling a particular sculpture an icon, a speedometer an indicator, or a coat of arms a symbol, we are engaging in a sort of tacit shorthand. What we usually mean is that they were designed to be interpreted that way, or are highly likely to be interpreted that way. So, for example, a striking resemblance does not make one thing an icon of another. Only when considering the features of one brings the other to mind because of this resemblance is the relationship iconic. Similarity does not cause iconicity, nor is iconicity the physical relationship of similarity. It is a kind of inferential process that is based on recognizing a similarity. As critics of the concept of iconicity have often pointed out, almost anything could be considered an icon of anything else, depending on the vagueness of the similarity considered.

The same point can be made for each of the other two modes of referential relationship: neither physical connection nor involvement in some conventional activity dictates that something is indexical or symbolic.
spectively. Only when these are the basis by which one thing invokes another are we justified in calling their relationship indexical or symbolic. Though this might seem an obvious point, confusion about it has been a source of significant misunderstandings. For example, there was at one time considerable debate over whether hand signs in American Sign Language (ASL) are iconic or symbolic. Many signs seemed to resemble pantomime or appeared graphically to "depict" or point to what was represented, and so some researchers suggested that their meaning was "merely iconic" and by implication, not wordlike. It is now abundantly clear, however, that despite such resemblances, ASL is a language and its elements are both symbolic and wordlike in every regard. Being capable of iconic or indexical interpretation in no way diminishes these signs' capacity of being interpreted symbolically as well. These modes of reference aren't mutually exclusive alternatives; though at any one time only one of these modes may be prominent, the same signs can be icons, indices, and symbols depending on the interpretive process. But the relationships between icons, indices, and symbols are not merely a matter of alternative interpretations. They are to some extent internally related to one another.

This is evident when we consider examples where different interpreters are able to interpret the same signs to a greater or lesser extent. Consider, for example, an archeologist who discovers some elaborate markings on clay tablets. It is natural to assume that these inscriptions were used symbolically by the people who made them, perhaps as a kind of primitive writing. But the archeologist, who as yet has no Rosetta Stone with which to decode them, cannot interpret them symbolically. The archeologist simply infers that to someone in the past these may have been symbolically interpretable, because they resemble symbols seen in other contexts. Being unable to interpret them symbolically, he interprets them iconically. Some of the earliest inscription systems from the ancient Middle Eastern civilizations of the Fertile Crescent were in fact recovered in contexts that provided additional clues to their representations. Small clay objects were marked with repeated imprints, then sealed in vessels that accompanied trade goods sent from one place to another. Their physical association with these other artifacts has provided archeologists with indexical evidence to augment their interpretations. Different marks apparently indicated a corresponding number of items shipped, probably used by the recipient of the shipment to be sure that all items were delivered. No longer merely iconic of other generic writinglike marks, they now can be given indexical and tentative symbolic interpretations, because something more than resemblance is provided.

This can also be seen by an inverse example: a descent down a hierar-
the basis by which one thing invokes another's relationship indexical or symbolic. A point, confusion about it has been a running. For example, there was at one-fist hand signs in American Sign Language. Many signs seemed to resemble partly to "depict" or point to what was chers suggested that their meaning was not wordlike. It is now abundantly clear. lances, ASL is a language and its elements every regard. Being capable of iconic or diminishes these signs' capacity of being those modes of reference aren't mutually y one time only one of these modes may be icons, indices, and symbols depend the relationships between icons, indices, of alternative interpretations. They are one another.

examples where different interpreters to a greater or lesser extent. Consider, covers some elaborate markings on clay these inscriptions were used symboli perhaps as a kind of primitive writing, to Rosetta Stone with which to decode it. The archeologist simply infers have been symbolically interpretable, in other contexts. Being unable to interprets them iconically. Some of the ear- in contexts that provided additional clay objects were marked with rests that accompanied trade goods sent ical association with these other arti-indexical evidence to augment their entity indicated a corresponding num- the recipient of the shipment to be longer merely iconic of other generic indexical and tentative symbolic re than resemblance is provided. example: a descent down a hierar-
modes of reference. Though I may fail to grasp the symbolic reference of a sign, I might still be able to interpret it as an index (i.e., as correlated with something else), and if I also fail to recognize any indexical correspondences, I may still be able to interpret it as an icon (i.e., recognize its resemblance to something else). Breakdown of referential competence leads to an ordered descent from symbolic to indexical to iconic, not just from complex icons, indices, or symbols to simpler counterparts. Conversely, increasing the sophistication of interpretive competence reverses the order of this breakdown of reference. For example, as human children become more competent and more experienced with written words, they gradually replace their iconic interpretations of these marks as just more writing with indexical interpretations supported by a recognition of certain regular correspondences to pictures and spoken sounds, and eventually use these as support for learning to interpret their symbolic meanings. In this way they trace a path somewhat like the archaeologist learning to decipher an ancient script.

This suggests that indexical reference depends upon iconic reference, and symbolic reference depends upon indexical reference—a hierarchy diagrammatically depicted in Figure 3.1. It sounds pretty straightforward on the surface. But this simplicity is deceiving, because what we really mean is that the competence to interpret something symbolically depends upon already having the competence to interpret many other subordinate relationships indexically, and so forth. It is one kind of competence that grows out of and depends upon a very different kind of competence. What constitutes competence in this sense is the ability to produce an interpretive response that provides the necessary infrastructure of more basic iconic and/or indexical interpretations. To explain the basis of symbolic communication, then, we must describe what constitutes a symbolic interpretant, but to do this we need first to explain the production of iconic and indexical interpretants and then to explain how these are each recoded in turn to produce the higher-order forms.

So, we need to start the explanation of symbolic competence with an explanation of what is required in order to interpret icons and build upward. Usually, people explain icons in terms of some respect or other in which two things are alike. But the resemblance doesn’t produce the iconicity. Only after we recognize an iconic relationship can we say exactly what we saw in common, and sometimes not even then. The interpretive step that establishes an iconic relationship is essentially prior to this, and it is something negative, something that we don’t do. It is, so to speak, the act of not making a distinction. Let me illustrate this with a very stripped-down example.

Consider camouflage, as in the moth on a tree whose wings resemble the tree and the bird moves, or the bird has been trained to distinguish the appearance of the moth an interpretive process would have taken place which wasn’t just negative—e.g., the bird might even have learned that the wing pattern to bark, at least at the first movement of the hapless moth. Some features were recognized, irrespective of their degree, and this positively supports the interpretation (an inattentive).

Now, it might seem awkward to call this behavior a value judgment, but it could be considered to be an act of discrimination and to clarify the shift in emphasis.
Consider camouflage, as in the case of natural protective coloration. A moth on a tree whose wings resemble the graininess and color of the bark, though not perfectly, can still escape being eaten by a bird if the bird is inattentive and interprets the moth’s wings as just more tree. Admittedly, this is not the way we typically use the term iconic, but I think it illuminates the most basic sense of the concept. If the moth had been a little less matching, or had moved, or the bird had been a little more attentive, then any of the differences between the moth and the tree made evident by those additional differences would have indicated to the bird that there was something else present which wasn’t just more tree. If the bird had been in a contemplative mood, it might even have reflected on the slight resemblance of the wing pattern to bark, at least for the fraction of a second before it gobbled the hapless moth. Some features of the moth’s wings were iconic of the bark, irrespective of their degree of similarity, merely because under some interpretation (an inattentive bird) they were not distinguished from it.

Now, it might seem awkward to explain iconicity with an example that could be considered to be no representation at all, but I think it helps to clarify the shift in emphasis I want to make from the relationship to the
process behind it. What makes the moth wings iconic is an interpretive process produced by the bird, not something about the moth's wings. Their coloration was taken to be an icon because of something that the bird didn't do. What the bird was doing was actively scanning bark, its brain seeing just more of the same (bark, bark, bark . . . ). What it didn't do was alter this process (e.g., bark, bark, not-bark, bark . . . ). It applied the same interpretive perceptual process to the moth as it did to the bark. It didn't distinguish between them, and so confused them with one another. This established the iconic relationship between moth and bark. Iconic reference is the default. Even in an imagined moment of reflective reverie in which the bird ponders on their slight resemblance, it is the part of its responding that does not distinguish wing from bark that determines their relationship to be iconic. Iconic resemblance is not based on some prior ground of physical similarity, but in that aspect of the interpretation process that does not differ from some other interpretive process. Thus, although a respect in which two things are similar may influence the ways they tend to be iconically related, it does not determine their iconicity. Iconism is where the referential buck stops when nothing more is added. And at some level, due either to limitations in abilities to produce distinguishing responses or simply a lack of effort to produce them, the production of new interpreters stops. Whether because of boredom or limitations of a minimal nervous system, there are times when almost anything can be iconic of anything else (stuff, stuff, stuff . . . ).

What does this have to do with pictures, or other likenesses such as busts or caricatures that we more commonly think of as icons? The explanation is essentially no different. That facet or stage of my interpretive recognition process that is the same for a sketch and the face it portrays is what makes it an icon. I might abstractly reflect on what aspects of the sketch caused this response, and might realize that this was the intention of the artist, but a sketch that is never seen is just paper and charcoal. It could also be interpreted as something that soaked up spilled coffee (and the spilled coffee could be seen as a likeness of Abe Lincoln!). Peirce once characterized an icon as something which upon closer inspection can provide further information about the attributes of its subject. Looking at the one is like looking at the other in some respects. Looking at a caricature can, for example, get one to notice for the first time that a well-known politician has a protruding jaw or floppy jowls. The simplification in a diagram or the exaggeration in a cartoon takes advantage of our spontaneous laxness in making distinctions to trick us into making new associations. In this way a caricature resembles a joke, a visual pun, and a diagram can be a source of discovery.

In summary, the interpretive process is none other than what in other recognition, but not necessarily is it all: to "think [about something] present something again. Iconic resemblance of which things can be re-presented by representation are built. It is to iconic resemblance is interpreted, and thus seen analyzed to its component reducibility (due to competitiveness constraints), and thus is ultimately reducible, not necessarily require any effort ceases. It can merely be a matter of consciousness where experience begins.

Interpreting something as an icon requires that physical contiguity (nearness) precede occurrence are the basis for it, but as with the case of icons, temporally associated with anything temporally associated with anything is the interpretive response will be understandable relationship to interpretations, it is necessary for interpretations arises. In common, be attributed to interpretive competence of new interpreters, indexical competence, in fact, icons arise from a failure of things.

Consider the example of smelling smoke, I begin to suspect possibility to treat this smell as an indicator, because I had past experience things that were burning. After association, and the smell of smoke will be near. If we consider more indexical competence, the iconic indexical competence is conscious, icons, and the indexical interpretation assembly of iconic relationship.
In summary, the interpretive process that generates iconic reference is none other than what in other terms we call recognition (mostly perceptual recognition, but not necessarily). Breaking down the term re-cognition says it all: to "think [about something] again." Similarly, representation is to present something again. Iconic relationships are the most basic means by which things can be re-presented. It is the base on which all other forms of representation are built. It is the bottom of the interpretive hierarchy. A sign is interpreted, and thus seen to be a representation, by being reduced (i.e., analyzed to its component representations) to the point of no further reducibility (due to competence or time limitations, or due to pragmatic constraints), and thus is ultimately translated into iconic relationships. This does not necessarily require any effort. It is in many cases where interpretive effort ceases. It can merely be the end of new interpretation, that boundary of consciousnes where experience fades into redundancy.

Interpreting something as an indexical relationship is this and more. Physical contiguity (nearness or connectedness) or just predictable co-occurrence are the basis for interpreting one thing as an index for another, but as with the case of icons, these physical characteristics are not the cause of the indexical relationship. Almost anything could be physically or temporally associated with anything else by virtue of some extension of the experience of nearness in space or time. What makes one an index of another is the interpretive response whereby one seems to "point to" the other. To understand the relationship that indexical interpretations have to iconic interpretations, it is necessary to see how the competence to make indexical interpretations arises. In contrast to iconic interpretations, which can often be attributed to interpretive incompetence or the cessation of production of new interpretants, indexical interpretations require something added. In fact, icons arise from a failure to produce critical indices to distinguish things.

Consider the example of a symptom, like the smell of smoke. When I smell smoke, I begin to suspect that something is burning. How did my ability to treat this smell as an indication of fire arise? It likely arose by learning, because I had past experiences in which similar odors were traced to things that were burning. After a few recurrences it became a familiar association, and the smell of smoke began to indicate to me that a fire might be near. If we consider more closely the learning process that produced the indexical competence, the critical role of icons becomes obvious. The indexical competence is constructed from a set of relationships between icons, and the indexical interpretation is accomplished by bringing this assembly of iconic relationships to bear in the assessment of new stimuli. The
smell of smoke brings to mind past similar experiences (by iconically representing them). Each of these experiences comes to mind because of their similarities to one another and to the present event. But what is more, many of these past experiences also share other similarities. On many of these occasions I also noticed something burning that was the source of the smoke, and in this way those experiences were icons of each other.

There is one important feature added besides all these iconic recognitions. The repeated correlation between the smelling of smoke and the presence of flames in each case adds a third higher-order level of iconicity. This is the key ingredient. Because of this I recognize the more general similarity of the entire present situation to these past ones, not just the smoke and not just the fire but also their co-occurrence, and this is what brings to mind the missing element in the present case: the probability that something is burning. What I am suggesting, then, is that the responses we develop as a result of day-to-day associative learning are the basis for all indexical interpretations, and that this is the result of a special relationship that develops among iconic interpretive processes. It's hierarchic. Prior iconic relationships are necessary for indexical reference, but prior indexical relationships are not in the same way necessary for iconic reference. This hierarchic dependency of indices on icons is graphically depicted in Figure 3.2.

Okay, why have I gone to all this trouble to rename these otherwise common, well-established uses of perception and learning? Could we just substitute the word “perception” for “icon” and “learned” association for index? No. Icons and indices are not merely perception and learning, they refer to the inferential or predictive powers that are implicit in these neural processes. Representational relationships are not just these mechanisms, but a feature of their potential relationship to past, future, distant, or imaginary things. These other things are not physically re-presented but only virtually re-presented by producing perceptual and learned responses like those that would be produced if they were present. In this sense, mental processes are no less representational than external communicative processes, and communicative processes are no less mental in this regard. Mental representation reduces to internal communication.

What, then, is the difference between these uncontroversial cognitive processes underlying icons and indices and the kind of cognitive processes underlying symbols? The same hierarchical logic applies. As indices are constituted by relationships among icons, symbols are constituted by relationships among indices (and therefore also icons). However, what makes this a difficult step is that the added relationship is not mere correlation.

Figure 3.2  A schematic diagram between iconic and indexical references. As iconic of something decreases similarity and decreases, objects have a decreasing capacity to be unobscured. The following past stimulus memories (e.g., neural processes. Thus, any remembered iconic stimulus. Similarly, each referential relationship need not be the same relationship to be produced. Indexical in order to interpret something as iconic of something else recognized. First, the instances (the top iconic relations also correlate (arrows) with adequate to be iconic of one another (the relationships need to be interpreted as an arrangement of arrows). The importance of iconic interpretations, with one treating them as parts of a whole kind of reference provided by any

The Symbolic Threshold

The common sense idea is that we learn to pair a sound or type of similar utterances) and then.
ur experiences (by iconically representing them) comes to mind because of their recent event. But what is more, we other similarities. On many of them, that was the source of the were icons of each other.

Besides all these iconic recognitions, the smell of smoke and the presholar order level of iconicity. This recognize the more general simi-

pares past ones, not just the smoke irrence, and this is what brings to case: the probability that some-

then, is that the responses we de-

ve learning are the basis for all the result of a special relationship processes. It’s hierarchic. Prior
exical reference, but prior index-

v necessary for iconic reference.

Icons is graphically depicted in

able to rename these otherwise
tion and learning? Could we just
n “unlearned” association for
ly perception and learning, they
are implicit in these neural
are not just these mechanisms, but
past, future, distant, or imaginary
ally re-presented but only virtually
and learned responses like those

In this sense, mental processes
communicative processes, and
in this regard. Mental repre-

these uncontroversial cognitive
of the kind of cognitive processes
logic applies. As indices are consigns are constituted by relations-
ons). However, what makes this

 Figure 3.2 A schematic diagram depicting the internal hierarchic relationships between iconic and indexical reference processes. The probability of interpreting something as iconic of something else is depicted by a series of concentric domains of decreasing similarity and decreasing iconic potential among objects. Surrounding objects have a decreasing capacity to serve as icons for the target object as similarities become unobvious. The form of a sign stimulus (S) elicits awareness of a set of past stimulus memories (e.g., mental “images”) by virtue of stimulus generalization processes. Thus, any remembered object (O) can be said to be re-presented by the iconic stimulus. Similarly, each mental image is iconic in the same way; no other referential relationship need necessarily be involved for an iconic referential relationship to be produced. Indexical reference, however, requires iconic reference. In order to interpret something as indexical, at least three iconic relationships must be also recognized. First, the indicating stimulus must be seen as an icon of other similar instances (the top iconic relationships); second, instances of its occurrence must also correlate (arrives) with additional stimuli either in space or time, and these need to be iconic of one another (the bottom iconic relationships); and third, past correlations need to be interpreted as iconic of one another (indicated by the concentric arrangement of arrows). The indexical interpretation is thus the conjunction of three iconic interpretations, with one being a higher-order icon than the other two (i.e., treating them as parts of a whole). As pointed out in the text, this is essentially the kind of reference provided by a conditioned response.

The Symbolic Threshold

The common sense idea is that a symbolic association is formed when we learn to pair a sound or typed string with something else in the world. But in the terms we have been developing, this is what we mean by an indexical association. The word (iconically associated with past occurrences of similar utterances) and the object (iconically associated with similar ob-
jects from past experiences) and their past correlations enable the word to bring the object to mind. In this view, the association between a word and what it represents is not essentially distinguished from the kind of association that is made by an animal in a Skinner box. We might, for example, train a rat to recognize a correlation between hearing the sound of the word “food” and food being dropped into a tray. The conditioned stimulus takes on referential power in this process: it represents something about the state of the apparatus for the animal. It is an index of the availability of food in the Skinner box; a symptom of the state of the box. Words can serve indexical functions as well, and are sometimes used for this purpose almost exclusively, with minimal symbolic content. Consider, for example, the use of function words like “there,” exclamations like “Aha!,” or even proper names like “George Washington.” These derive reference by being uniquely linked to individual contexts, objects, occasions, people, places, and so on, and they defy our efforts to define them as we would typical nouns or verbs.

One indication that someone understands the meaning of a new word is whether they can use it in a new sentence or novel context. If the new word was just learned as a part of an unanalyzed phrase, or mapped to some restricted acquisition context, then we might not expect it to be correctly used out of this context. But the ability to use a word correctly in a variety of contexts, while fair evidence of symbolic understanding, is not necessarily convincing as a proof of understanding. The ability to shift usage to a novel context resembles transference of one learning set; and indeed, searching for the common learning set features among the many contexts in which the same word might be used is a good way to zero in on its meaning. If someone were to learn only this—i.e., that a particular phrase works well in a range of contexts that exhibit similar features or social relationships—they might well be able to fool us into believing that they understood what they said. However, on discovering that they accomplished this by simply mapping similar elements from one context to another, we would conclude that they actually did not understand the word or its role in context in the way we originally imagined. Theirs would be an iconic and indexical understanding only. Being able easily to transfer referential functions from one “set” to another is a characteristic of symbols, but is this the basis for their reference?

Psychologists call transfer of associations from one stimulus to another similar one “stimulus generalization,” and transfer of a pattern of learning from one context to another similar context the transfer of a “learning set.” These more complex forms of indexical association are also often confused with symbolic associations. Transference of learning from stimulus to stim-

ulus or from context to context is a characteristic of learning. These are not really based on iconic projection of one stimuli onto another spontaneously because there are no essential parameters of the stimulus that are a subsequent desired or understood feature of the stimulus from a finite number of common experiences. These can provide a basis for choosing a stimulus. To the extent that we have a familiar set of stimuli used for naming them, these other potential and the psychological models of this process have learned rules for identifying symbols. This is based on an iconic relationship between the symbol and what it is learned; only a failure to discriminate will be excluded by the training.

Words for kinds of things are similar objects, such as could be used for qualities and properties of things. It is often the basis for stimulus projection that produce the same sign when presented a group of familiar objects or animals, or any other objects in common, even subtle ones like the vervet monkeys’ eagle alarm cry, or any other predators if they were introduced into a new habitat. If these referents are not by symbol, they apply our own symbolic criteria for a label. The basis for their common indexical use may contribute to the symbol, but often first learned by its members. It is only one subordinate component that determines their reference.

This same logic applies to the question of the human’s ability to choose the more similar to each other than any other kind of oddity—perhaps as a way to define an associated response on the available options. The recognition of iconicity between symbols is a hierarchy of symbols that is referred to a novel stimulus.
past correlations enable the word to
the association between a word and
tinguished from the kind of associ-a-
ter box. We might, for example, train
sen hearing the sound of the word
The conditioned stimulus takes
reprents something about the state
n index of the availability of food in
of the box. Words can serve indexical
used for this purpose almost exclu-
Consider, for example, the use of
is like "Aha!", or even proper names
reference by being uniquely linked
s, people, places, and so on, and they
would typical nouns or verbs.
stands the meaning of a new word is
c or novel context. If the new word
ized phrase, or mapped to some rea-
ght not expect it to be correctly used
word correctly in a variety of con-
derstanding, is not necessarily con-
he ability to shift usage to a novel
learning set; and indeed, searching
mong the many contexts in which
way to zero in on its meaning. If
that a particular phrase works well
features or social relationships—
believing that they understood what
t accomplished this by simply
ct to another, we would conclude
r word or its role in context in the
ld be an iconic and indexical un-
fer referential functions from one
bolts, but is this the basis for their
ons from one stimulus to another
id transfer of a pattern of learning
ext the transfer of a "learning set." 
association are also often confused
of learning from stimulus to stim-
ulus or from context to context occurs as an incidental consequence of
learning. These are not really separate forms of learning. Both are based
on iconic projection of one stimulus condition onto another. Each arises
spontaneously because there is always some ambiguity as to what are the
essential parameters of the stimulus that a subject learns to associate with
a subsequent desired or undesired result: learning is always an extrapola-
tion from a finite number of examples to future examples, and these sel-
don provide a basis for choosing between all possible variations of a
stimulus. To the extent that new stimuli exhibit features shared by the fa-
miliar set of stimuli used for training, and none that are inconsistent with
them, these other potential stimuli are also incidentally learned. Often,
psychological models of this process are presented as though the subject
has learned rules for identifying associative relationships. However, since
this is based on an iconic relationship, there is no implicit list of criteria that
is learned; only a failure to distinguish that which hasn't been explicitly ex-
cluded by the training.

Words for kinds of things appear to refer to whole groups of loosely sim-
ilar objects, such as could be linked by stimulus generalization, and words
for qualities and properties of objects refer to the sorts of features that are
often the basis for stimulus generalization. Animals can be trained to pro-
duce the same sign when presented with different kinds of foods, or trees,
or familiar animals, or any other class of objects that share physical attrib-
utes in common, even subtle ones (e.g., all hoofed mammals). Similarly, the
vervet monkeys' eagle alarm calls might become generalized to other aeri-
al predators if they were introduced into their environment. The grouping
of these referents is not by symbolic criteria (though from outside we might
apply our own symbolic criteria), but by iconic overlap that serves as the
basis for their common indexical reference. Stimulus generalization may
contribute essential structure to the realms to which words refer, but it is
only one subordinate component of the relationship and not what deter-
mines their reference.

This same logic applies to the transference of learning sets. For exam-
ple, learning to choose the odd-shaped object out of three, where two are
more similar to each other than the third, might aid in learning a subse-
quent oddity-discrimination task involving sounds. Rather than just trans-
ferring an associated response on the basis of stimulus similarities, the subject rec-
ognizes an iconicity between the two learning tasks as wholes. Though this
is a hierarchically more sophisticated association than stimulus generaliza-
tion—learning a learning pattern—it is still an indexical association trans-
ferred to a novel stimulus via an iconic interpretation. Here the structure

Terrence W. Deacon
of the new training context is seen as iconic of a previous one, allowing the subject to map corresponding elements from the one to the other. This is not often an easy association to make, and most species (including humans) will fail to discover the underlying iconicity when the environment, the training stimuli, the specific responses required, and the reinforcers are all quite different from one context to the next.

There are two things that are critically different about the relationships between a word and its reference when compared to transference of word use to new contexts. First, for an indexical relationship to hold, there must be a correlation in time and place of the word and its object. If the correlation breaks down (for example, the rat no longer gets food by pushing a lever when the sound “food” is played), then the association is eventually forgotten (“extinguished”), and the indexical power of that word to refer is lost. This is true for indices in general. If a smokelike smell becomes common in the absence of anything burning, it will begin to lose its indicative power in that context. For the Boy Who Cried Wolf, in the fable of the same name, the indexical function of his use of the word “wolf” fails because of its lack of association with real wolves, even though the symbolic reference remains. Thus, symbolic reference remains stable nearly independent of any such correlations. In fact, the physical association between a word and an appropriate object of reference can be quite rare, or even an impossibility, as with angels, unicorns, and quarks. With so little correlation, an indexical association would not survive.

Second, even if an animal subject is trained to associate a number of words with different foods or states of the box, each of these associations will have little effect upon the others. They are essentially independent. If one of these associations is extinguished or is paired with something new, it will likely make little difference to the other associations, unless there is some slight transference via stimulus generalization. But this is not the case with words. Words also represent other words. In fact, they are incorporated into quite specific individual relationships to all other words of a language. Think of the way a dictionary or thesaurus works. They each map one word onto other words. If this shared mapping breaks down between users (as sometimes happens when words are radically reused in slang, such as “bad” for “very good” or “plastered” for “intoxicated”), the reference also will fail.

This second difference is what ultimately explains the first. We do not lose the indexical associations of words, despite a lack of correlation with physical referents, because the possibility of this link is maintained implicitly in the stable associations between words. It is by virtue of this sort of dual reference, to objects and to other words (or at least to other semantic alternatives), that a word can become objects of reference. This dual association indicates other words—foods, places, persons, and words point to other words, rather than refer them. Thus, the reference, not vice versa.

This referential relationship, indicating other words—foods, persons, places, and objects themselves. But this is also true of the indexical words, in phrases and sentences. Their indexical power is dependent on the correlation between words. Symbolic reference involves possibilities and impossibilities, and these are never discoverable (during learning). Thus the imagined version of a word is dependent on its referent, but lacking the relevant context is ultimately a contradiction.

Even without struggling with this relationship, we can immediately see that the problem associated with the lack of correlation is that what determines the referent of a single word, object or event is not their absolute or complex function or relationship. This is a separate but linked problem from the third, higher-order unlearned framework of the probability of correlation and the behavioral level. Past correlations must be subordinated to reduce the troublesome shift of emphasis from symbol-object relationships treated as more than characteristic of the word. And these relationships are not quite the reverse. Words that can be used alternatively and not (or as complementary) referential function in sentences. Worst of all, few relationships and the frequency with which they are also extremely low. Hardly
ic of a previous one, allowing the
from the one to the other. This is
most species (including humans)
when the environment, the train-
and the reinforcers are all quite
different about the relationships
compared to transference of word
relationship to hold, there must
word and its object. If the corre-
no longer gets food by pushing a
then the association is eventually
it will begin to lose its indicative
ried Wolf, in the fable of the same
the word “wolf” fails because of
though the symbolic referen-
ces stable nearly independent of any
sociation between a word and an
ite rare, or even an impossibility,
so little correlation, an indexical
trained to associate a number of
box, each of these associations
are essentially independent. If
or is paired with something new,
other associations, unless there is
ralization. But this is not the case
ords. In fact, they are incorporat-
e to all other words of a language.
s works. They each map one word
breaks down between users (as
ally reused in slang, such as “bad”
ated”), the reference also will fail.
ately explains the first. We do not
despite a lack of correlation with
of this link is maintained implica-
ords (or at least to other semantic
alternatives), that a word conveys the information necessary to pick out ob-
jects of reference. This duality of reference is captured in the classic dis-
tinction between sense and reference. Words point to objects (reference)
and words point to other words (sense), but we use the sense to pick out
the reference, not vice versa.

This referential relationship between the words—words systematically
indicating other words—forms a system of higher-order relationships that
allows words to be about indexical relationships, and not just indices in
themselves. But this is also why words need to be in context with other
words, in phrases and sentences, in order to have any determinate refer-
ence. Their indexical power is distributed, so to speak, in the relationships
between words. Symbolic reference derives from combinatorial possibili-
ties and impossibilities, and we therefore depend on combinations both to
discover it (during learning) and to make use of it (during communication).
Thus the imagined version of a nonhuman animal language that is made up
of isolated words, but lacking regularities that govern possible combinations,
is ultimately a contradiction in terms.

Even without struggling with the philosophical subtleties of this rela-
tionship, we can immediately see the significance for learning. The learning
problem associated with symbolic reference is a consequence of the fact
that what determines the pairing between a symbol (like a word) and some
object or event is not their probability of co-occurrence, but rather some
complex function of the relationship that the symbol has to other symbols.
This is a separate but linked learning problem, and worse yet, it creates a
third, higher-order unlearning problem. Learning is, at its base, a function
of the probability of correlations between things, from the synaptic level to
the behavioral level. Past correlations tend to be predictive of future cor-
relations. This, as we’ve seen, is the basis for indexical reference. In order
to comprehend a symbolic relationship, however, such indexical associations
must be subordinate to relationships between different symbols. This is a
troublesome shift of emphasis. To learn symbols we begin by learning
symbol-object correlations, but once learned, these associations must be
treated as no more than clues for determining the more crucial rela-
thions. And these relationships are not highly correlated; in fact, often just
the reverse. Words that carry similar referential function are more often
used alternatively and not together, and words with very different (com-
plementary) referential functions tend to to be adjacent to one another in
sentences. Worst of all, few sentences or phrases are ever repeated exactly,
and the frequency with which specific word combinations are repeated is
also extremely low. Hardly a recipe for easy indexical learning.

Terrence W. Deacon
One of the most insightful demonstrations of the learning difficulties associated with the shift from conditioned associations to symbolic associations comes not from a human example, but from a set of experiments that attempted to train chimpanzees to use simple symbols. This study was directed by Sue Savage-Rumbaugh and Duane Rumbaugh,¹ now at the Language Research Center of Georgia State University, and included four chimpanzees, two of which, Sherman and Austin, showed particular facility with the symbols. It is far from the “last word” on how far other species can go in their understanding of languagelike communication, and further studies of another chimpanzee (from a different subspecies) that show more developed abilities will be described subsequently (see Chapter 4),³ but this work has the virtue of exposing much of what is often hidden in children’s comparatively easy entry into symbolic communication, and so provides an accessible step-by-step account of what we usually take for granted in the process. In what follows I will outline these experiments briefly. Only the most relevant highlights will be described and other aspects will be simplified for the sake of my purpose here. Of course, my attempts to “get inside the chimps’ heads” during this process are fantasy. Though I will use somewhat different terminology from the experimenters to describe this transition from indexical to symbolic communication, I am reasonably confident that my interpretation is not at odds with theirs. However, the interested reader should refer to the excellent account of these experiments and their significance in Savage-Rumbaugh’s book describing them.

The chimps in this study were taught to use a special computer keyboard made up of lexigrams—simple abstract shapes (lacking any apparent iconism to their intended referents) on large illuminated keys on a keyboard mounted in their cage. Duane Rumbaugh’s previous experiments (with a chimpanzee named Lana)⁶ had shown that chimps have the ability to learn a large number of paired associations between lexigrams (and in fact other kinds of symbol tokens) and objects or activities. But in order to respond to critics and more fully test other features of this ability, Duane and Sue began a new series of experiments with a group of chimps to test both chimps’ ability to use lexigrams in combinations (e.g., syntactic relationships). Not surprisingly, the chimps exhibited some interesting difficulties when they were required to use lexigrams in combinations, but they eventually solved their learning problems and exhibited a use of the lexigrams that was clearly symbolic. In so doing they have provided us with a remarkably explicit record of the process that leads from index to symbol.

In order to test Sherman and Austin’s symbolic understanding of the lex-
of the learning difficulties as- 
asociations to symbolic associ-
ations from a set of experiments that 
as single symbols. This study was di-
ne Rumbaugh, now at the Lan-
1 University, and included four 
as showed particular facility with 
as on how far other species can go 
munication, and further studies 
species) that show more de-
sently (see Chapter 4), but this 
as that is often hidden in children's 
munication, and so provides an 
as usually take for granted in the 
s experiments briefly. Only the 
as and other aspects will be sim-
urse, my attempts to "get inside 
ity. Though I will use some-
emers to describe this trans-
lation, I am reasonably confident 
as theirs. However, the interest-
... of these experiments and their 
s describing them.

use a special computer keyboard 
apes (lacking any apparent icon-
illuminated keys on a keyboard 
as previous experiments (with a 
as have the ability to learn a large 
exigrams (and in fact other kinds 
as But in order to respond to crit-
as ability, Duane and Sue began 
p of chimp to test both chimps 
s to use lexigrams in combinations 
sly, the chimp exhibited some 
s required to use lexigrams in com-
leaning problems and exhibited 
s. In so doing they have pro-
s of the process that leads from 
symbolic understanding of the lex-
igrams, the chimp were trained to chain lexigram pairs in a simple verb-
noun relationship (a sequence glossed as meaning "give," which caused a 
dispenser to deliver a solid food, and "banana" to get a banana). Initially 
there were only 2 "verb" lexigrams and 4 food or drink lexigrams to choose 
and each pair had to be separately taught. But after successful training 
of each pairing, the chimp were presented with all the options they had 
learned independently, and were required to choose which combination was 
most appropriate on the basis of food availability or preference. Curiously, 
the solution to this task was not implicit in their previous training. This was 
evident in the fact that some chimps tended stereotypically to repeat only 
the most recent single learned combination, whereas others chained to-
gether all options, irrespective of the intended meanings and what they 
knew about the situation. Thus they had learned the individual associations 
but failed to learn the system of relationships of which these correlations 
were a part. Although the logic of the combinatorial relationships between 
lexigrams was explicit in the particular combinations that the chimps 
learned, the converse exclusive relationships had not been learned. For 
example, they were not explicitly trained to avoid any number of inap-
ropriate combinations such as "banana juice give." Though these errors are 
explicit for us, who treat them symbolically from the start, the combina-
torial rules that allow pairing in some but not other cases was vastly under-
determined by the training experience (as it is also in a child's experience 
of others' word use).

It is not immediately obvious exactly how much exclusionary informa-
tion is implicit, but it turns out to be quite a lot. Think about it from the 
a naive chimpanzee perspective for a moment. Even with this ultra-simple 
symbol system of six lexigrams and a two-lexigram combinatorial grammar, 
the chimpanzee is faced with the possibility of sorting among 720 possible 
ordered sequences (6*5*4*3*2*1) or 64 possible ordered pairs. The train-
ing has offered only four prototype examples, in isolation. Though each 
chimp may begin with many guesses about what works, these are unlikely 
to be in the form of rules about classes of allowed and disallowed combina-
tions, but rather about possible numbers of lexigrams that must be 
pressed, their positions on the board, their colors or shape cues that might e associated with a reward object, and so on. Recognizing this limitation, 
the experimenters embarked on a rather interesting course of training. 
They set out explicitly to train the chimp on which cues were not relevant 
and which combinations were not meaningful. This poses an interesting 
problem that every pet trainer has faced. You can't train what not to do un-
less the animal first produces the disallowed behavior. Only then can it be
immediately punished or at least explicitly not rewarded (the correlation problem again). So the chimps were first trained to produce incorrect associations (e.g., mistaking keyboard position as the relevant variable) and then these errors were explicitly not rewarded, whereas the remaining appropriate responses were. By a complex hierarchic training design, involving thousands of trials, it was possible to teach them to exclude systematically all inappropriate associative and combinatorial possibilities among the small handful of lexigrams. At the end of this process, the animals were able to produce the correct lexigram strings every time.

Had training out the errors worked? To test this, the researchers introduced a few new food items and corresponding new lexigrams. If the chimps had learned the liquid/solid rule, and got the idea that a new lexigram was for a new item, they might learn more quickly. Indeed they did. Sherman and Austin were able to respond correctly the first time, or with only a few errors, instead of taking hundreds of trials as before. What had happened to produce this difference? What the animals had learned was not only a set of specific associations between lexigrams and objects or events. They had also learned a set of logical relationships between the lexigrams, relationships of exclusion and inclusion. More importantly, these lexigram-lexigram relationships formed a complete system in which each allowable or forbidden co-occurrence of lexigrams in the same string (and therefore each allowable or forbidden substitution of one lexigram for another) was defined. They had discovered that the relationship that a lexigram has to an object is a function of the relationship it has to other lexigrams, not just a function of the correlated appearance of both lexigram and object. This is the essence of a symbolic relationship.

The subordination of the indexical relationships between lexigrams (symbol tokens) and foods (referents or objects) to the system of indexical relationships between lexigrams is schematically depicted in three stages of development in Figure 3.3. Individual indexical associations are shown as single vertical arrows, mapping each token to a kind of object, because each of these relationships is independent of the others. In contrast, the token-token interrelationships (e.g., between lexigrams or words), shown as horizontal arrows interconnecting symbols, form a closed logical group of combinatorial possibilities. Every combination and exclusion relationship is unambiguously and categorically determined. The indexical reference of each symbol token to an object after symbolic reference is achieved is depicted with arrows reversed to indicate that these are now subordinate to the token-token associations.

In the minimalistic symbol system first learned by Sherman and Austin,
iety not rewarded (the correlation trained to produce incorrect as- ision as the relevant variable) and varded, whereas the remaining ap- hierarchic training design, involv- each them to exclude systematically torial possibilities among the small process, the animals were able to try time. To test this, the researchers intro- ducing new lexigrams. If the chimp- the idea that a new lexigram was quickly. Indeed they did. Sherman s the first time, or with only a few als as before. What had happened animals had learned was not only a grams and objects or events. They ships between the lexigrams, rela- more importantly, these lexigram- e system in which each allowabl s in the same string (and therefore of one lexigram for another) was relationship that a lexigram has to it has to other lexigrams, not just e of both lexigram and object. This p- relationships between lexigrams (sym- to the system of indexical rela- tically depicted in three stages of indexical associations are shown as ten to a kind of object, because each e others. In contrast, the token- lexigrams or words), shown as hor- ls, form a closed logical group of ution and exclusion relationship trimmed. The indexical reference to symbolic reference is achieved is de- that these are now subordinate to rst learned by Sherman and Austin.

![Figure 3.3](image-url) Figure 3.3 A schematic depiction of the construction of symbolic referential relationships from indexical relationships. This figure builds on the logic depicted in Figure 3.2, but in this case the iconic relationships are only implied and the indexical relationships are condensed into single arrows. Three stages in the construction of symbolic relationships are shown from bottom to top. First, a collection of different indices are individually learned (varying strength indicated by darkness of arrows). Second, systematic relationships between index tokens (indexical stimuli) are recognized and learned as additional indices (gray arrows linking indices). Third, a shift (reversal of indexical arrows) in mnemonic strategy to rely on relationships between tokens (darker arrows above) to pick out objects indirectly via relationships between objects (corresponding lower arrow system). Individual indices can stand on their own in isolation, but symbols must be part of a closed group of transformations that links them in order to refer, otherwise they revert to indices.
reference to objects is a collective function of relative position within this token-token reference system. No individual lexigram determines its own reference. Reference emerges from the hierarchic relationship between these two levels of indexicality, and by virtue of recognizing an abstract correspondence between the system of relationships between objects and the system of relationships between the lexigrams. In a sense, it is the recognition of an iconic relationship between the two systems of indices. Although indexical reference of tokens to objects is maintained in the transition to symbolic reference, it is no longer determined by or dependent on any physical correlation between token and object.

This makes a new kind of generalization possible: logical or categorical generalization, as opposed to stimulus generalization or learning set generalization. It is responsible for Sherman and Austin's ability to acquire new lexigrams and know their reference implicitly, without any trial-and-error learning. The system of lexigram-lexigram interrelationships is a source of implicit knowledge about how novel lexigrams must be incorporated into the system. Adding a new food lexigram, then, does not require the chimps to learn the correlative association of lexigram to object from scratch each time. The referential relationship is no longer solely (or mainly) a function of lexigram-food co-occurrence, but has become a function of the relationship that this new lexigram shares with the existing system of other lexigrams, and these offer a quite limited set of ways to integrate new items. The chimps succeed easily because they have shifted their search for associations from relationships among stimuli to relationships among lexigrams. A new food or drink lexigram must fit into a predetermined slot in this system of relationships. There are not more than a few possible alternatives to sample, and none requires assessing the probability of paired lexigram-food occurrence because lexigrams need no longer be treated as indices of food availability. Like words, the probability of co-occurrences may be quite low. The food lexigrams are in a real sense "nouns," and are defined by their potential combinatorial roles. Testing the chimps' ability to extrapolate to new lexigram-food relationships is a way of demonstrating whether or not they have learned this logical-categorical generalization, which is a crucial defining feature of symbolic reference.

At some point toward the end of the training, the whole set of explicitly presented indexical associations that the chimps had acquired was "recoded" in their minds with respect to an implicit pattern of associations whose evidence was distributed across the whole set of trials. Did this recoding happen as soon as they had learned the full set of combination/exclusion relationships among their lexigram set? I suspect not. Try to imagine yourself in their situation for a moment. You are not making errorless trials and struggling to remember who got what, especially at the level of one-by-one pairing of objects and memory all the details. What would you know, because there are a lot of them? The furry, support you discover that there is a regularity that begins to appear. It is the reward regularities: the reward patterns are far more regular than the individual associations that you acquired, yet weren't apparent previously. The relationship begins to emerge. But the chimps were using added mnemonics to help them with the errorless trials over and over. From the initial details well, they also became familiar with each other, and they would, otherwise, that there was a systemic pattern of information, information about the context, greatly the mnemonic load of the tasks. They could now afford to focus on the context and could keep track of them via context.

What I am suggesting here is that the shift to symbolic predictions is in nostalgic way. It is a way of offloading some of the cognitive processing by recognizing a higher-order representation that can accomplish the same task. Unfortunately, nature seldom provides us with strategies, so not much selection is needed to create artificial systems that do accomplish the same point is that when such a system allows a shift in mnemonic storage, it adds the mode of representation structure onto another way. It gets recoded and you know them also in a different way. From the bottom up, indexically, then, because this recoding is based on the full set of hierarchical details, it often vastly augments the representational power of the system.
on of relative position within this dual lexigram determines its own
hierarchic relationship between
the process of recognizing an abstract cor-
relations between objects and the
grams. In a sense, it is the recog-
ize two systems of indices. Although
is maintained in the transition to
ed or dependent on any phys-
ontological views: logical or categorical
eralization or learning set gen-
and Austin's ability to acquire new
icity, without any trial-and-error in
m. Interrelationships is a source of
grams must be incorporated into
then, does not require the chimp
gram to object from scratch each
nger solely (or mainly) a function
become a function of the rela-
the existing system of other lexi-
t of ways to integrate new items.
their search for asso-
to relationships among lexigrams.
 a predetermined slot in this sys-
than a few possible alternatives to
probability of paired lexigram-
no longer be treated as indices of
ty of co-occurrences may be quite
"nouns," and are defined by their
chimps' ability to extrapolate to
of demonstrating whether or not
generalization, which is a crucial

raining, the whole set of explicitly
chimps had acquired was "re-
 an implicit pattern of associations
the whole set of trials. Did this re-
ed the full set of combination/ex-
set? I suspect not. Try to imagine

yourself in their situation for a moment. You have just come to the point
where you are not making errors. What is your strategy? Probably, you are
struggling to remember what specific things worked and did not work, still
at the level of one-by-one associations. The problem is, it is hard to re-
member all the details. What you need are aids to help organize what you
know, because there are a lot of possibilities. But in the internal search for
supports you discover that there is another source of redundancy and reg-
ularity that begins to appear, besides just the individual stimulus-response-
reward regularities: the relationships between lexigrams! And these
redundant patterns are far fewer than the messy set of dozens of individual
associations that you are trying to keep track of. These regularities
weren't apparent previously, because errors had obscured any underlying
systematic relationship. But now that they are apparent, why not use them
as added mnemonics to help simplify the memory load? Forced to repeat
errorless trials over and over, Sherman and Austin didn't just learn the de-
tails well, they also became aware of something they couldn't have noticed
otherwise, that there was a system behind it all. And they could use this new
information, in the form of what they had already learned, to simplify
greatly the mnemonic load created by the many individual rote associations.
They could now afford to forget about individual correlations so long as they
could keep track of them via the lexigram-lexigram rules.

What I am suggesting here is that the shift from associative predictions
to symbolic predictions is initially a change in mnemonic strategy, a recod-
ing. It is a way of offloading redundant details from working memory, by
recognizing a higher-order regularity in the mess of associations, a trick that
can accomplish the same task without having to hold all the details in mind.
Unfortunately, nature seldom offers such nice neat logical systems that can
help organize our associations. There are not many chances to use such
strategies, so not much selection for this sort of process. We are forced to
create artificial systems that have the appropriate properties. The crucial
point is that when such a systematic set of tokens becomes available, it al-
ows a shift in mnemonic strategy that results in a radical transformation in
the mode of representation. What one knows in one way gets recoded in
another way. It gets re-represented. We know the same associations, but we
know them also in a different way. You might say we know them both from
the bottom up, indexically, and from the top down, symbolically. And be-
cause this recoding is based on higher-order relationships, not the indi-
idual details, it often vastly simplifies the mnemonic problem and vastly
augments the representational possibilities. Equally important is the vast
amount of implicit knowledge it provides. Because the combinatorial rules
encode not objects but ways in which objects can be related, new symbols can immediately be incorporated and combined with others based on independent knowledge about what they symbolize.

The experimenters working with Sherman and Austin provided a further, and in some ways even more definitive, demonstration of the difference between indexical reference of lexigram-object correlations and symbolic reference in a subsequent experiment that compared the performance of the two symboling apes (Sherman and Austin) to a previous subject (Lana), who had been trained with the same lexigram system but not in the same systematic fashion. Lana had learned a much larger corpus of lexigram-object associations, though by simple paired associations. In this new experiment (see Figure 3.4), all three chimps were first tested on their ability to learn to sort food items together in one pan and tool items together in another (Lana learned in far fewer trials than Sherman and Austin). When all three chimps had learned this task, they were presented with new foods or tools to sort and were able to generalize from their prior behavior to sort these new items appropriately as well. This is essentially a test of stimulus generalization, and it is based on some rather abstract qualities of the test items (e.g., edibility). It shows that chimps have a sophisticated ability to conceptualize such abstract relationships irrespective of symbols. Of course, chimpanzees (as well as most other animal species) must be able to distinguish edible from inedible objects and treat each differently. Learning to sort them accordingly takes advantage of this preexisting categorical discrimination in a novel context. In this sense, then, what might be called an indexical concept of food and nonfood precedes the training. Each bin is eventually treated as indexical of this qualitative sensory and behavioral distinction, and so the ability to extend this association to new food and nonfood items involved stimulus generalization (though of an indirectly recognizable stimulus parameter).

This sorting task was followed by a second task in which the chimps were required to associate each of the previously distinguished food items with the same lexigram (glossed as “food” by the experimenters) and each of the tool items with another lexigram (“tool”). Initially, this task simply required the chimps to extend their prior associations with bins to two additional stimuli, the two lexigrams. Although all three chimps learned this task in a similar way, taking many hundreds of trials to make the transference, Sherman and Austin later spontaneously recoded this information in a way that Lana did not. This was demonstrated when, as in the prior task, novel food and novel tool items were introduced. Sherman and Austin found this to be a trivial addition and easily guessed without any additional learning which

Figure 3.4 Summary of part of a symbolic learning of lexigram referential trials and the panels on the right show the indexical learning of lexigram referential with the panels on the right show the indexical learning of lexigram referential. The second task required identifying the food and tool objects with the appropriate general lexigram was appropriate. Lana, given the new items, the novelty and stereotyped behavior that caused her to fail the test. Though on the surface the conflicting results demonstrate the inadequacy of the rote learning strategy and the recombining used by Sherman and Austin, it’s due to the fact that the sorting task used object and object, whereas the lexigram corresponded. Lana appeared more
etcs can be related, new symbols aligned with others based on symbolize.

In and Austin provided a further, non-trivial demonstration of the difference between correlations and symbolic refinement, the performance of the tool subjects (Lana), who system but not in the same system, was larger corpus of lexigram-object associations. In this new experiment, they tested on their ability to learn tool items together in another man and Austin. When all three resented with new foods or tools their prior behavior to sort these essentially a test of stimulus generalization. The distinct qualities of the test items is a sophisticated ability to conceive of symbols. Of course, a species must be able to distinguish each differently. Learning to this preexisting categorical divisive, then, what might be called an recedes the training. Each bin is native sensory and behavioral dissociation to new food and notation (though of an indirectly task in which the chimps were able, distinguished food items with the experimenters) and each of the initially, this task simply required a with bins to two additional stimuli. The chimps learned this task in a simple way: the transference, Sherman as information in a way that Lana in the prior task, novel food and man and Austin found this to be but any additional learning which

Figure 3.4 Summary of part of a 1980 test of lexigram reference in chimpanzees by E. Sue Savage-Rumbaugh and her colleagues. This compares three levels of symbolic learning of lexigram reference by the chimps Sherman and Austin to the inductive learning of lexigram reference by another chimpanzee, Lana, who is unable to complete tasks requiring symbolic reference. The panels on the left depict training trials and the panels on the right depict items added in test trials. Test trials introduced new lexigrams and tested to determine generalization to items for which there was no previous experience. The top task was merely a sorting task to determine that all animals understood the distinction between foods and tools (nonfood). The second task required identification with one of two lexigrams (“food,” “tool”). Though all three learned it, only Sherman and Austin made the shift to symbolic categorization of reference and were able to generalize to new items (because of past symbol-learning experience). Lana was excluded from the remaining two procedures (not shown), where Sherman and Austin learned first to associate lexigrams to pictures of the foods and tools, and then to associate individual food and tool lexigrams with the appropriate general lexigram for food or tool.

lexigram was appropriate. Lana not only failed to extend her categorization to the new items, the novelty and errors appeared to produce a kind of counterevidence that caused her to abandon her prior training in a subsequent test. Though on the surface this task resembles the sorting task, these conflicting results demonstrate that there is a critical difference that undermined the rote learning strategy used by Lana and favored the symbolic recoding used by Sherman and Austin. The difference is probably related to the fact that the sorting task involved a physical-spatial association of sign and object, whereas the lexigram “labeling” involved only temporal correspondence. Lana appeared not to be using these underlying qualities to
solve the task. For her, each lexigram object association was an independent datum, and so provided no information about other associations.

In contrast Sherman and Austin, as a result of their experience with a previous symbol system, recoded these new lexigram-object associations into two new symbolic categories that superseded the individual associations. It took them hundreds or thousands of trials to learn the first simple one-to-many associations. This was because they began with no systemic relationship in their still small lexigram repertoire for a general reference to “food” or “tool.” They had to learn them the hard way, so to speak, indexically. But as soon as they did learn these associations, they were primed to look for another higher-order logic, and once it was discovered, they were able to use this logic to generalize to new associations. Instead of hundreds or even thousands of trials, the availability of a symbolic coding allowed them to bypass further trials altogether, an incredible increase in learning efficiency. The chimps essentially knew something that they had never explicitly learned. They had gained a kind of implicit knowledge as a spontaneous byproduct of symbolic recoding.

I have chosen to recount this ape language study not because it portrays any particularly advanced abilities in chimpanzees, or because I think it is somehow representative. In fact (as noted earlier), more recent studies by these same experimenters, with a pygmy chimpanzee (or bonobo) named Kanzi, have demonstrated far more effortless and sophisticated symbolic abilities. Rather, I have focused on this earlier study because of the clarity with which it portrays the special nature of symbol learning, and because it clearly exemplifies the hierarchic relationship between symbolic and indexical reference. The reductio ad absurdum training ploy is particularly instructive, not because it is an essential element but because it provides an explicit constructive demonstration of the index-by-index basis of the eventual symbolic relationship. It also demonstrates how normal associative learning strategies can interfere with symbol learning. Indexical associations are necessary stepping stones to symbolic reference, but they must ultimately be superseded for symbolic reference to work.

**Unlearning an Insight**

The problem with symbol systems, then, is that there is both a lot of learning and unlearning that must take place before even a single symbolic relationship is available. Symbols cannot be acquired one at a time, the way other learned associations can, except after a reference symbol system is established. A logically complete system of relationships among the set of symbol tokens must be learned, and this occurs prior to recognizing that it emerges from a system; it is an emergent. For this reason, it’s hard for relationship requires holding a number of possible combinations work and which can not be covered or perceived, in some way. In other words, it is an implicit relationship between the indexical and the relationship of something symbolically interdependent associations. The matrix of symbol-symbol relationships of them to refer symbolically on recognizing their symbolic association not individual indexical referential referential reference, and symbolic association is a rest. Symbolic associations are suddenly seen with respect to one another. This means to press one set of associative ones on them. Discovering the superimposed learning step, it is just implicit and object relationships that have we might call a symbolic inference strategy and gradual memory searches.

What I have described are symbolic reference would clearly although my analysis suggest
object association was an independence about other associations.
A result of their experience with new lexigram-object associations superseded the individual associations of trials to learn the first simple use they began with no systemic repertoire for a general reference to the hard way, so to speak, indexical associations, they were primed to once it was discovered, they were associations. Instead of hundreds of a symbolic coding allowed them an incredible increase in learning efficiency that they had never explicit knowledge as a spontaneous language study not because it portrays impanzaes, or because I think it is earlier, more recent studies by chimpanzee (or bonobo) named for less and sophisticated symbolic an earlier study because of the clarity of symbol learning, and because the relationship between symbolic and individual training ploy is particularly element but because it provides one of the index-by-index basis of the demonstrates how normal associational symbol learning. Indexical association, a symbolic reference, but they must reference to work.

en, is that there is both a lot of learning before even a single symbolic relationship acquired one at a time, the way after a reference symbol system is essential relationships among the set of symbolic tokens must be learned before the symbolic association between any one symbol token and an object can even be determined. The learning step occurs prior to recognizing the symbolic function, and this function only emerges from a system; it is not vested in any individual sign-object pairing. For this reason, it's hard to get started. To learn a first symbolic relationship requires holding a lot of associations in mind at once while at the same time mentally sampling the potential combinatorial patterns hidden in their higher-order relationships. Even with a very small set of symbols the number of possible combinations is immense, and so sorting out which combinations work and which don't requires sampling and remembering a large number of possibilities.

One of the most interesting features of the shift in learning strategy that symbolic reference depends upon is that it essentially takes no time; or rather, no more time than the process of perceptual recognition. Although the prior associations that will eventually be recoded into a symbolic system may take considerable time and effort to learn, the symbolic recoding of these relationships is not learned in the same way; it must instead be discovered or perceived, in some sense, by reflecting on what is already known. In other words, it is an implicit pattern that must be recognized in the relationships between the indexical associations. Recognition means linking the relationship of something new to something already known. The many interdependent associations that will ultimately provide the nodes in a matrix of symbol-symbol relationships must be in place in order for any one of them to refer symbolically, so they must each be learned prior to recognizing their symbolic associative functions. They must be learned as individual indexical referential relationships. The process of discovering the new symbolic association is a restructuring event, in which the previously learned associations are suddenly seen in a new light and must be reorganized with respect to one another. This reorganization requires mental effort to suppress one set of associative responses in favor of another derived from them. Discovering the superordinate symbolic relationship is not some added learning step, it is just noticing the system-level correspondences that are implicitly present between the token-token relationships and the object-object relationships that have been juxtaposed by indexical learning. What we might call a symbolic insight takes place the moment we let go of one associative strategy and grab hold of another higher-order one to guide our memory searches.

What I have described as the necessary cognitive steps to create symbolic reference would clearly be considered a species of “insight learning,” though my analysis suggests that the phrase is in one sense an oxymoron.

Terrence W. Deacon 493
Psychologists and philosophers have long considered the ability to learn by insight to be an important characteristic of human intelligence. Animal behaviorists have also been fascinated with the question, Can other animals learn by insight? The famous Gestalt psychologist Wolfgang Köhler described experiments with chimpanzees in which to reach a fruit they had to “see” the problem in a new way. Köhler set his chimp the problem of retrieving a banana suspended from the roof of the cage and out of reach, given only a couple of wooden boxes that when stacked one upon the other could allow the banana to be reached. He found that these solutions were not intuitively obvious for a chimpanzee, which would often become frustrated and give up for a long period. During this time she would play with the boxes, often piling them up, climbing on them, and then knocking them down. At some point, however, the chimp eventually appeared to have recognized how this fit with the goal of getting at the banana, and would then quite purposefully maneuver the boxes into place and retrieve the prize. Once learned, the trick was remembered. Because of the role played by physical objects as mnemonic place-holders and the random undirected exploration of them, this is not perhaps the sort of insight that appears in cartoons as the turning on of a light bulb, nor is it what is popularly imagined to take place in the mind of an artist or scientist. On the other hand, what goes on “inside the head” during moments of human insight may simply be a more rapid covert version of the same, largely aimless object play. We recognize these as examples of insight solely because they involve a recoding of previously available but unlinked bits of information.

Most insight problems do not involve symbolic recoding, merely sensory recoding: “visualizing” the parts of a relationship in a new way. Transference of a learning set from one context to another is in this way also a kind of insight. Nevertheless, a propensity to search out new “perspectives” might be a significant advantage for discovering symbolic relationships. The shift in mnemonic strategy from indexical to symbolic use of food and food-delivery lexigrams required the chimps both to use the regularities of symbol-token combinations as the solution to correct performance, and to discover that features of the food objects and delivery events correspond to these lexigram combination regularities. In other words, they had to use these combination relationships to separate the abstract features of liquid and solid from their context of indexical associations with the food-delivery events. The symbolic reference that resulted depended on digging into these aspects of the interrelationships between things, as opposed to just mapping lexigrams to things themselves. By virtue of this, even the specific combinations of tokens cannot be seen as indexical, so that it is not just that the ability to combine tokens in a way that using two digits instead of one is to use numerical values. Which tokens, and cannot substitute for one another, to what linguists call “semantics” the response of some property like “solid” or “to grow.” New elements can be brought in to add semantical features that the system integrates with those that somehow can be associated as symbol groups, independently integrated with each other. Once a structure is recognized, there can be an enormous number of new configurations.

The insight-recoding processual recoding steps become much more frequent. In this way, a child’s initial direct recoding of language is only the beginning of the learning/unlearning process. Each new relationship (i.e., more about the world) is, in a sense, a process anew. This produces a number of more or less discrete stages. New relationships that must be sampled in any new acquisition increases geometrically with the growing size of the lexigram vocabulary. It is always necessary to confine the symbolic recoding becomes apparent as the seemingly endless rote learning of some mathematical operations at the crux of the psychologist Jean Piaget’s pattern of symbolic concept acquisition, the pattern of symbolization, information processing and new learning.

The ability of Sherman and his associates for “food” and “tool” to be perceived as different in potential conflict between the previously acquired and new concepts and the further testing of their occurrence, then re-pairing the
the ability to combine tokens vastly multiplies referential possibilities, in the
way that using two digits instead of one makes it possible to represent more
numerical values. Which tokens can and cannot be combined and which can
and cannot substitute for one another determines a new level of mapping
to what linguists call “semantic features,” such as the presence or absence
of some property like “solidity.” This is what allows a system of symbols to
grow. New elements can be added, either by sharing reference with semi-
scopic features that the system already defines, or by identifying new fea-
tures that somehow can be integrated with existing ones. Even separate
symbol groups, independently constructed, can in this way become inte-
grated with each other. Once the relationship between their semantic fea-
ture sets is recognized, their unification can in one insight create an
enormous number of new combinatorial possibilities.

The insight-reencoding problem becomes increasingly difficult as addi-
tional recoding steps become involved in establishing an association. For
this reason, a child’s initial discovery of the symbolic relationships underly-
ing language is only the beginning of the demand on this type of learn-
ing/unlearning process. Each new level of symbols coding for other symbolic
relationships (i.e., more abstract concepts) requires that we engage this
process anew. This produces a pattern of learning that tends to exhibit
more or less discrete stages. Since the number of combinatorial possibili-
ties that must be sampled in order to discover the underlying symbolic logic
increases geometrically with each additional level of recoding, it is almost
always necessary to confine rote learning to one level at a time until the sym-
broic recoding becomes apparent before moving on to the next. This limita-
tion is frustratingly familiar to every student who is forced to engage in
seemingly endless rote learning before “getting” the underlying logic of
some mathematical operation or scientific concept. It may also contribute
to the crudely stagelike pattern of children’s cognitive development, which
the psychologist Jean Piaget initially noticed.10 However, this punctuated
pattern of symbolic conceptual development is a reflection of symbolic in-
formation processing and not an intrinsic feature of developing brains and
minds.

The ability of Sherman and Austin to discover the abstract symbolic ref-
erences for “food” and “tool” provides an additional perspective on the dif-
ference between indexical associations and symbolic associations. Consider
the potential conflict between the lexigram-object relationships they had
previously acquired and this new set of associations. If their prior associa-
tions were supported only by the correlations in lexigram-object-reward oc-
currence, then re-pairing the same objects with a new lexigram would be

Terrence W. Deacon 495
expected to partially if not totally extinguish the prior association. Although it would be possible to provide additional contextual cues to enable the chimps to decide which of two competing associative strategies to use (e.g., simply run trials without the alternatives available) and thus learn and retain both, there would still be interference effects (i.e., their prior associations might interfere both with relearning the new associations and with shifting between them in different contexts). Unfortunately, data to assess this are not available, but we can infer from Sherman and Austin’s learning shifts, and their subsequent maintenance of the prior symbolic associations, that neither extinction nor interference was a significant problem. Though it was not tested explicitly in this series of experiments, we should expect that this should also distinguish Sherman and Austin from Lana. Certainly Lana’s rapid decline in performance when new items were added points to such effects.

This ability to remember large numbers of potentially competing associations is an additional power of symbolic reference that derives from the shift in mnemonic strategy to token-token relationships. Competition effects grow with increasing numbers of overlapping associative categories in typical indexical reference relationships. Not only would the choice among alternatives in any use become a source of confusion, but because they were competing for reinforcement, each would weaken the association of the others. Though some of the interference effects also attend symbol use, and often are a cause of word retrieval errors and analysis delays, in terms of associative strength there is an opposite effect. Competing sets of overlapping associative relationships on the indexical level translate into mutually supportive higher-order semantic categories on the symbolic level. These become sources of associative redundancy, each reinforcing the mnemonic trace of the other. So, rather than weaken the strength of the association, they actually reinforce it.

This helps to explain where the additional associative glue between words and their referents comes from. Though token-object correlations are not consistently available to the symbol user, indeed are rare, this loss of associative support is more than compensated by the large number of other associations that are available through symbolically mediated token-token relationships. Individually, these are comparatively weak associations, with a low correlated occurrence of any two tokens in the same context; but they are not just one-to-one associations. They are one-to-many and many-to-one associations that weave symbol tokens together into a systematic network of association relationships, and the pattern has a certain coded isomorphism with relationships between objects and events in the world.

As a result of sharing many indexical associations, the word “animal” gains mnemonic strength because they are multiply connected. The associative strengths make them an easier external correlate of their associations. Thus, not only is “animal” so is its mnemonic support. And behind the bits of information that are so useful to recall. How else do we use every day be retrieved? Speaking or listening?

Numerous neuropsychological studies have demonstrated this for word meanings as a source of priming effects between words in overlapping categories. A word prime later memory tasks in interesting is the fact that this allows these words as well. Receipt of the word “cat” would cause the physiological correlates of stress response (e.g., skin response) upon hearing the word “feline” response will also be present, even though there had never been a response to a rhyming word like “tense” response. All of these distantly related word relationships to one another and to a conditioned word, the shared trace, do not decay with shock. The extent to which this appears to correlate with the stimulus generalization, it is some stimulus parameters that distinguish produce a similar priming. That there is an independent transfer of association. Rhyme associations are a true and some transference of physiology.

This analogy between effect of shared semantic features shows...
ish the prior association. Although nontextual cues to enable the g associative strategies to use (e.g., s available) and thus learn and recall effects (i.e., their prior associating the new associations and with external cues). Unfortunately, data to assess om Sherman and Austin’s learning ice of the prior symbolic associative was a significant problem. s series of experiments, we should erman and Austin from Lana. Cerence when new items were added

ers of potentially competing asso-
ic reference that derives from the xon relationships. Competition ef-overlapping associative categories in . Not only would the choice among confusion, but because they were weaken the association of the others also attend symbol use, and s and analysis delays, in terms of as-
fect. Competing sets of overlapping al level translate into mutually sup-
port on the symbolic level. These be-
y, each reinforcing the mnemonic ten the strength of the association,

as a result of sharing many weak interpenetrating indexical links, each indexical association gains mnemonic support from a large number of others because they are multiply coded in memory. Together, their combined associative strengths make them far more resistant to extinction due to diminished external correlations with objects than are individual indexical associations. Thus, not only is symbolic reference a distributed relationship, so is its mnemonic support. This is why learning the symbolic reasons behind the bits of information we acquire by rote learning offers such a powerful aid to recall. How else could the many thousands of different words we use every day be retrieved so rapidly and effortlessly during the act of speaking or listening?

Numerous neuropsychological probes of semantic field effects demonstrate this for word meaning. Hearing, memorizing, or using a word can be a source of priming effects for subsequent recall or identification of other words in overlapping categories. For example, hearing the word “cat” might prime later memory tasks involving “dog” or “animal.” Even more interesting is the fact that this also transfers to indexical associations involving these words as well. Receiving a mild electric shock every time you hear the word “cat” would cause you to learn to spontaneously produce physiological correlates of stress response (such as change in heart rate or galvanic skin response) upon hearing that word repeated. But a similar but less intense response will also be produced whenever you hear a word like “dog,” even though there had never been shocks associated with these sounds. A lesser response will also be produced whenever you hear a word like “meow” or “animal,” demonstrating lexical (word-word) associations, and in response to a rhyming word like “mat,” demonstrating stimulus generalization effects. All of these distinct associative relationships are brought into relationship to one another by the symbolic relationship. Because each arouses an associative network that overlaps with that of the shock-conditioned word, the shared activation raises an arousal level also associated with shock. The extent of both the symbolic and indexical overlap appears to correlate with the extent of the transference. Though analogous to stimulus generalization, it is clearly different. There are no shared stimulus parameters that distinguish “dog” and “cat” from “car,” which does not produce a similar priming. The difference is also reflected in the fact that there is an independent transference to words that rhyme, like “flat” or “sat.” Rhyme associations are true stimulus generalization effects and also show some transference of physiological responses.

This analogy between effects involving shared stimulus features and shared semantic features shows that the brain stores and retrieves both sym-

Terrence W. Deacon ≤97
bolic and nonsymbolic associations as though they were the same sort of thing. Just as the contingencies of co-occurrence and exclusion in the same context determine the strengths of stimulus associations, so too do these statistics in language affect the strengths of word associations.

With each shift of referential control to a token-token system of relationships, it became possible for Sherman and Austin to add new lexical items to their growing symbol system with a minimum of associative learning, often without any trial-and-error testing. This produces a kind of threshold effect whereby prior associative learning strategies, characterized by an incremental narrowing of stimulus response features, are replaced by categorical guesses among a few alternatives. The result is a qualitative shift in performance. The probabilistic nature of the earlier stage is superseded by alternative testing that has a sort of all-or-none character. This change in behavior can thus be an indication of the subject’s shift in mnemonic strategy, and hence the transition from indexical to symbolic reference. The simplest indicator of this shift is probably the rate of acquisition of new lexical items, since this should be highly sensitive to the hundred- to thousand-fold reduction in trial-and-error learning required to reach 100 percent performance.

In young children’s learning of language, apparent threshold effects have long been noticed in vocabulary growth and sentence length. Vocabulary and utterance length are of course linked variables in two regards. First, the more words a child knows, the more there are to string together. But this does not simply translate into larger sentences. Creating a larger sentence in a human language cannot just be accomplished by stringing together more and more words. It requires the use of hierarchic grammatical relationships, as well as syntactic tricks for condensing and embedding kernel sentences in one another. Thus, not only does vocabulary need to grow, but the types of words must diversify. In other words, the regular discovery of new grammatical classes must be followed by a rapid filling of these classes with new alternative lexical items.

Each time a new logical group is discovered among a set of tokens, it essentially opens up one or more types of positional slots that can be filled from an open class of symbols. Each slot determines both a semantic and a grammatical category. Recall that although Sherman and Austin could add new food items to their lexigram “vocabulary” with little difficulty, when they had to learn to recode food items in terms of the higher-order semantic category “food,” they essentially had to start over. Their prior knowledge of the symbolic designations of distinct foods with respect to food-delivery modes was of no help. It may even have been a source of interference, since the same foods were now being recoded into this new symbolic association, usually involving no error.

In the small symbol system, semantic features that were available might be specified by the delivery (of food). Discovering or uncovering these semantic features provided the basis for adding correlations. All that was new was represented with respect to the preceding one, in order to know implicitly correlation. Beginning with any interrogative, repeated stages. Each stage results in new building and a large system of tokens.

In summary, then, symbolic reference is the collection of tokens that map onto a new class of objects or features. It’s not just tokens or symbols; tokens are yet another layer of abstraction. Symbols in the sentence refer to them by virtue of referring to different sorts of entities with referents that are moving determinate positions in the sentence and occupy different positions in the sentence. The structure of the whole refers to the ways symbols can be combined into different combinations. Because symbolic reference, no collection of tokens or symbols refers only because the entire collection conforms to a symbolic reference emerging from the process of emergent systems. Symbolic reference emerges only as the collection of tokens is organized as to form a logical and systematic open system. This determinate collection of tokens has to supplant the entire collection in a collective consciousness of that system. This collective consciousness involves those symbols in a new way of knowing power. The structure implicit in such a system is in the system before symbolic reference, but...
same foods were now being linked with different lexigrams. But again, once this new symbolic association was established, adding new items proved trivial, usually involving no errors.

In the small symbol system initially learned by Sherman and Austin, the semantic features that were implicit in the few combinatorial possibilities available might be specified in terms of solid versus liquid and food versus delivery (of food). Discovering the combinatorial rules was the key to discovering these semantic features, and, conversely, these semantic features provided the basis for adding new symbols without needing to relearn new correlations. All that was necessary was prior knowledge of the object to be represented with respect to one or more of the relevant semantic features in order to know implicitly a token's combinatorial possibilities and reference. Beginning with any initial core, the system can grow rapidly in repeated stages. Each stage represents a further symbolic transition that must begin with incremental indexical learning. But past experience at symbol building and a large system of features can progressively accelerate this process.

In summary, then, symbols cannot be understood as an unstructured collection of tokens that map to a collection of referents because symbols don't just represent things in the world, they also represent each other. Because symbols do not directly refer to things in the world, but indirectly refer to them by virtue of referring to other symbols, they are implicitly combinatorial entities whose referential powers are derived by virtue of occupying determine positions in an organized system of other symbols. Both their initial acquisition and their later use requires a combinatorial analysis. The structure of the whole system has a definite semantic topology that determines the ways symbols modify each other's referential functions in different combinations. Because of this systematic relational basis of symbolic reference, no collection of signs can function symbolically unless the entire collection conforms to certain overall principles of organization. Symbolic reference emerges from a ground of nonsymbolic referential processes only because the indexical relationships between symbols are organized so as to form a logically closed group of mappings from symbol to symbol. This determinate character allows the higher-order system of associations to supplant the individual (indexical) referential support previously invested in each component symbol. This system of relationships between symbols determines a definite and distinctive topology that all operations involving those symbols must respect in order to retain referential power. The structure implicit in the symbol-symbol mapping is not present before symbolic reference, but comes into being and affects symbol com-

Terrence W. Deacon
binations from the moment it is first constructed. The rules of combination that are implicit in this structure are discovered as novel combinations are progressively sampled. As a result, new rules may be discovered to be emergent requirements of encountering novel combinatorial problems, in much the same way as new mathematical laws are discovered to be implicit in novel manipulations of known operations.

Symbols do not, then, get accumulated into unstructured collections that can be arbitrarily shuffled into different combinations. The system of representational relationships, which develops between symbols as symbol systems grow, comprises an even more complex matrix. In abstract terms, this is a kind of tangled hierarchic network of nodes and connections that defines a vast and constantly changing semantic space. Though semioticists and semiotic theorists have proposed various analogies to explain these underlying topological principles of semantic organization (such as plus/minus feature lists, dictionary analogies, encyclopedia analogies), we are far from a satisfactory account. Whatever the logic of this network of symbol-symbol relationships, it is inevitable that it will be reflected in the patterns of symbol-symbol combinations in communication.

Abstract theories of language, couched in terms of possible rules for combining unspecified tokens into strings, often implicitly assume that there is no constraint on theoretically possible combinatorial rule systems. Arbitrary strings of uninterpreted tokens have no reference and thus are unconstrained. But the symbolic use of tokens is constrained both by each token's use and by the use of other tokens with respect to which it is defined. Strings of symbols used to communicate and to accomplish certain ends must inherit both the intrinsic constraints of symbol-symbol reference and the constraints imposed by external reference.

Some sort of regimented combinatorial organization is a logical necessity for any system of symbolic reference. Without an explicit syntactic framework and an implicit interpretive mapping, it is possible neither to produce unambiguous symbolic information nor to acquire symbols in the first place. Because symbolic reference is inherently systemic, there can be no symbolization without systematic relationships. Thus syntactic structure is an integral feature of symbolic reference, not something added and separate. It is the higher-order combinatorial logic, grammar, that maintains and regulates symbolic reference; but how a specific grammar is organized is not strongly restricted by this requirement. There need to be precise combinatorial rules, yet a vast number are possible that do not ever appear in natural languages. Many other factors must be taken into account in order to understand why only certain types of syntactic systems are actually em-

ployed in natural human languages.

So, before turning to the question about human brains that many think is far more difficult for us than for the chimpanzee or the other nonhuman species among whom we see evidence of this view of symbol systems in action, we only do this analysis suggest a way of freeing dependent facets of language from a semiotic theory—it also forces one to reconsider the nature of grammatical knowl-
structured. The rules of combination are often discovered as novel combinations are rules, may be discovered to be emergent combinatorial problems, in much as symbols are discovered to be implicit in...

divided into unstructured collections of different combinations. The system of development between symbols as symbol-complex matrix. In abstract terms, work of nodes and connections that semantic space. Though semanticists various analogies to explain these untied organization (such as +/− features, analogies), we are far from a satisfactorily network of symbol-symbol relations to be reflected in the patterns of communication. If understood in terms of possible rules for combination, we often implicitly assume that there is a combinatorial rule system. Arbitrary no reference and thus are unconstrained both by each token's with respect to which it is defined, rate and to accomplish certain ends hints of symbol-symbol reference and reference.

Hierarchical organization is a logical necessity. Without an explicit syntactic mapping, it is possible neither to produce nor to acquire symbols in the first inherently systemic, there can be no relationships. Thus syntactic structure is necessary, not something added and separate, logic, grammar, that maintains and a specific grammar is organized in, that. There need to be precise context that do not only appear in, must be taken into account in order of syntactic systems are actually em-
ployed in natural human languages and how we are able to learn the incredibly complicated rule systems that result.

So, before turning to the difficult problem of determining what it is about human brains that makes the symbolic recoding step so much easier for us than for the chimpanzees Sherman and Austin (and members of all other nonhuman species as well), it is instructive to reflect on the significance of this view of symbolization for theories of grammar and syntax. Not only does this analysis suggest that syntax and semantics are deeply interdependent facets of language—a view at odds with much current linguistic theory—it also forces us entirely to rethink current ideas about the nature of grammatical knowledge and how it comes to be acquired.