Chapter 1

Introduction

The reader is likely to come to this book with the question, what do these authors mean by an *empiricist view of language* and of language acquisition? Empiricism is a term, after all, with a good deal of history to it, and it carries with it the bruises and scars, and more simply the reminders, of disputes among groups of philosophers, psychologists, and linguists over a period that can be measured not only in generations but in centuries. We will begin with a very brief guided tour through some of the history of the term *empiricism* and explain which parts of that history we identify with and which parts we do not.

1.1 Empiricism: some history

As a movement, empiricism began in the 17th and 18th centuries as the cousin, on the British Isles, of the rationalist movement on the Continent. Both were engaged in the larger enterprise of establishing experimental science as the most reliable source of knowledge about the world, challenging the Church as the ultimate source of reliable belief.

This early moment of classical empiricism—the empiricism of Locke, Berkeley, and later Hume—laid great emphasis on the source of human knowledge being the senses. This was, after all, one of the great themes of the founders of modern science, a theme shared by thinkers as otherwise diverse as Galileo, Newton, and Bacon: the scientist must learn to read from the book of Nature, not just the books of ancient authorities. So strong was this emphasis that it would not be uncharitable to pronounce the central tenet of classical empiricism to be this: all knowledge comes through the senses.

The rationalists on the Continent, such as Descartes and Leibniz, were not sympathetic to this general epistemological perspective, and it was not an accident that Descartes and Leibniz were brilliant and important mathematicians. They argued that our knowledge of mathematical truths is far more certain than any knowledge that merely arises from the senses. Our certainty regarding mathematical truths does not derive from many encounters with sensory experiences that supported them; it derives from an understanding of the foundations of geometry, of algebra, and of mathematical reasoning in general. And these rationalists could point out, in their support, that the second pillar of modern science was that the language of Nature is mathematics: we not only observe Nature, we also speak its language, the language of mathematics.

This then was where the impasse was situated between the classical empiricists and rationalists: when it came to firm and reliable generalizations, one had to choose between rationalism, with its knowledge that does not come through the senses, or empiricism, which held that there were no grounds for any of these strong convictions.

In the late 18th century, the great Prussian philosopher Immanuel Kant tried to formulate a synthesis that would satisfy both the empiricists and the rationalists. Not all knowledge comes through the senses, he said, but what does not come through the senses is of a different sort than what does. Indeed, the knowledge that is logically *prior* to all experience is necessary to even *have* an experience. There are conditions on knowing and experiencing, and these could not possibly come *from* experience itself. Our notions of space, time, and causality do not come from experience: they are what make experience possible. These elements constitute the box outside of which we cannot think, for the simple reason that thinking is constructed from these elements.

Kant's notion was that one of the ways in which we humans understand the world is through specific intuitions: space and time are intuitions of our sensibility, and causation is an intuition of our understanding. These intuitions structure the way we can think about the world. Kant's term was *Anschauung*, which is translated into English as *intuition*, but Kant's intuition bears little resemblance to our everyday sense of intuition, that is, a weak belief for which we can't give a satisfactory account. These Kantian intuitions comprise the scaffold that make thought and perception possible, not something presented to the mind from without.

Now, Kant's account was enormously influential, but for many it was not very satisfying. His account was neither historical nor social, and it still failed to answer all sorts of questions about how people learn from experience. Worse yet, the utter certainty of some of Kant's a priori knowledge began to show some real shakiness: mathematicians began to talk about alternatives to classical space, and it seemed that ideas that were once certain would have to move to being not quite so certain. Thus, some of the concepts that Kant had assumed to be the very elements of thought, and to define the boundaries of what we can think, started to come under scrutiny; and even to be challenged and modified. It was not clear how such apparent mutability could be compatible with the rationalist view that such concepts are built into the very fabric of thought.

The tension between these two poles of thought, the empiricist and the rationalist, has not diminished in the more than two centuries since this classical period, although the specific claims that have separated the views have shifted over time. In almost every case, the views have changed because of developments in what philosophers once called "the special sciences"—what we today would simply call *science* (though we must remember to include in that not only the physical sciences, but the social sciences and the development of modern views on the foundations of mathematics and of computer science). We focus here on two important cases, both of which cast many once-accepted certainties into doubt: one concerns the development of the theory of evolution, and the other concerns the foundations of mathematics. There is a third case to consider, too: the emergence of a notion of computation, which offered a new way through the suddenly uncertain and shifting landscape; indeed, this notion forms one of the bases on which much of this book builds. But first, let us follow the 19th century philosophers and scientists into the nest of uncertainty caused by the Darwinian revolution and new developments in mathematics.

1.2 Two important developments and their consequences

1.2.1 The emergence of the evolutionary framework

One of the great moments in intellectual history, which fundamentally affected the debate between rationalists and empiricists, was an important realization due to Charles Darwin and Alfred Russel Wallace—namely, that from a biological perspective, there was no sharp cleavage between human beings and the rest of the biological world. The idea that humans had evolved by a process of natural selection from common ancestors with apes, other mammals, and ultimately all living creatures, implies that an account of human knowledge must somehow be consistent with the descent of humans from speechless animals who know nothing of mathematics or science. As we shall see shortly, this rise of evolutionary thinking was one of the important factors leading to the rise of modern psychology.

But what precisely are the implications of rooting human thought and behavior in biology? On the one hand, it might seem natural to assume that most complex animal behavior is instinctual and (in modern terms) encoded in the genes; and hence, to assume that, for example, human linguistic behavior must, despite its superficial variety, be genetically encoded in a similar way. On the other hand, we might stress the observation that while many complex behaviors, including language, are uniquely human, the human brain appears to be highly similar to that of our closest relatives such as chimps and gorillas—so that language might naturally be viewed not as the product of a genetic innovation specific to language but as emerging from a general increase in neural complexity. Either perspective seems reasonable. Thus, while a biological perspective does not immediately resolve the debate between nativist and empiricist views of language acquisition, it radically changes the *ground* of the debate.

1.2.2 The shifting foundations of mathematics

The second great event in recent intellectual history that left its mark on the debate between rationalism and empiricism was a fundamental shift in the conception of mathematical truth. A number of mathematical assertions that had once appeared to be unassailable candidates for certain knowledge began to lose their self-evident character. Not only could they be doubted; this doubt actually became the catalyst for spectacular mathematical developments. The first challenge was to Euclidean geometry. Mathematicians came to the realization that while flat Euclidean geometry might be the natural way for people to imagine shapes and space, it is not the only way to explore geometry. Indeed, physical reality might not play by Euclid's rules: space might have a negative or a positive curvature, if observed closely enough. The second challenge was to even deeper foundations of mathematics: the more closely mathematicians looked at how we must formulate mathematical statements to ensure that they attain the degree of explicitness and clarity required to achieve certainty, the more they realized that such expectations could not always be met. Mathematicians fell into disputes over which abstract objects were well defined and what kinds of logical steps were reasonable to take in a proof.

One set of disputes concerned the proper interpretation of the calculus; indeed, even the specification of a paradox-free notion of the real line proved astonishingly difficult to achieve. The idea that intuition provides a reliable guide to knowledge and is a solid foundation upon which inference can be carried out received its most severe blow, however, over the notion of a set. Frege [1893] sought to construct the machinery sufficient for reconstructing arithmetic and, ultimately, the rest of mathematics, by axiomatizing intuitions about sets—and deriving the rest of mathematics as logically valid inferences from these axioms. Yet Frege's apparently mild and intuitively compelling axiomatization of set theory, designed to be a firm base upon which mathematics might be built, turned out to be inconsistent. Russell's paradox [Russell, 1903], which uncovered the inconsistencies concerning the pathological "set of all sets that are not members of themselves", turned out to be remarkably difficult to evade.

The implication for the nature of human knowledge was harsh and inescapable: intuitive notions—upon which so much of mathematics and other a priori truths were thought to be based—may not be reliable after all. Moreover, intuitions are problematic not merely because they lead to paradox, but also because they may turn out not to lead to a single vision of the truth. A consistent feature of modern mathematics is the observation that apparently unitary notions, such as the concept of a pair of parallel lines, or the real numbers, or, indeed, the notion of a set or of elementary arithmetic, turn out to fragment into many possible notions—as described by many possible geometries, theories of real analysis, set theories [Cohen, 1963], or theories of arithmetic.

1.2.3 Resolving these challenges: how might knowledge come from within?

These two challenges to rationalism, arising from biology and mathematics, were viewed during the 19th century through the philosophical spectacles that Kant had provided.

Kant's idea of intuition had been offered in the first place as an explanation which might bridge the chasm between the empiricists and the rationalists: what we know by intuition is not learned through the senses, and at the same time it is not a reliable roadmap of an external, self-standing reality. But in the light of these 19th century crises, the possibility loomed that there might be faculties of mind whose validity we might need to be downright skeptical about. Even enthusiasts of non-Euclidean geometry had a hard time believing that anyone could think about non-Euclidean geometry as easily and naturally as they could about Euclidean geometry: the conclusion seemed to emerge that some of the intuitions generated by our built-in cognitive mechanisms could be systematically misleading. But this means that our intuitions, however compelling, cannot automatically be treated as a firm guide to *truth*. And once the possibility of doubt, even concerning our firmest intuitions, arises, then all intuitions seem potentially suspect: How can we draw a line in the sea of intuitions, dividing the reliable from the doubtful?

There are several lines of development that have arisen as efforts to provide an answer to this question, and we will sketch several of them, with the goal of placing different trends in context, trends which have influenced each other (and us, as well). But before moving on, note how easy it is for these great moments in the development of modern thought to sound rather catastrophic! Perhaps it would be better to say that these great 19th century advances—the Darwinian revolution, and the mathematical revolutions in geometry and in set theory—set in motion great anxiety with regard to the basis of human knowledge. Yes, we know more now, we have new theorems, we see farther, and we see smaller; but we face increasing difficulties in finding firm foundations for knowledge, of whatever kind, that can withstand serious criticism.

Now we must pick up another strand in this story. We saw that the classical rationalists were motivated by dissatisfaction with the classical empiricist's suggestion that all knowledge comes through the senses. Rationalists were dissatisfied with how little could be said to come through the senses, once we take that notion seriously. Indeed, Hume, the philosopher who pursued empiricism most relentlessly, emerged from his contemplations more than a little depressed with how little of our apparent knowledge of the external world, or even our inner mental lives, could really be justified through the senses alone. Hume concluded that much of our apparent knowledge, and the concepts, such as causality, with which we conceive the world should be viewed with skepticism, from an empiricist standpoint.

What could the rationalists provide as an alternative? What can we know that does not come through the senses? The influential precursor of rationalism, Plato, had provided one answer, which he called *anamnesis*: we know things in this world that we remember from our experiences in another earlier world, where we had lived once upon a time. We today might charge this with being empiricism wrapped in sheep's clothing: the source of the knowledge in question is, if not the senses in the usual sense of the term, at least in experience of one sort or other (prenatal, in this case, or before conception). In the 17th century, the early rationalists were steeped in scholasticism and were content with the notion that God might offer ideas to man or that man's mind could see through the light of *lumen naturalis*: a natural light of reason.

To many, though, these answers begged the question—which is to say, these answers assumed what they should be accounting for. For many, the Darwinian revolution of the 19th century provided a whole new family of answers to the question of how a person could know something without learning it through the senses: the knowledge might find its source in the effects of evolution, and the properties of mind might be accounted for in just the ways that the anatomy of a reptile, mammal, or monocotyledon might be—by seeing how it evolved over time, with natural selection (and not divine intervention) being the critical factor in nudging the organism in a direction that allowed it to best survive and reproduce in its natural environment.

To some, this refinement of rationalism was not appealing at all, and for a simple reason: this Darwinian picture offered no reason to believe that the mental structures that were being bred into humans in this way were in any interesting sense *true* or *justified*. Mental structures that are innate need not be sure guides to truth if their only reason for being is that they allowed their bearers to live to maturity and to procreate [Plantinga, 1993].

One important response to this criticism was *pragmatism*, in its wide variety of forms and guises: according to pragmatism, the notion of *truth*, properly understood, is nothing more than what the Darwinian view could offer. From a pragmatist's perspective, truth should be interpreted as that which works successfully in our world, in the broadest possible sense. Even today, much of the everyday work of pragmatist philosophers consists of efforts to convince skeptics (who are dissatisfied with the apparently slim pickings that come out of pragmatism) that they are being unreasonable in asking for more. Pragmatism is the brand of epistemology that takes Darwinian evolution, and more broadly, a scientific conception of the human mind, seriously; it gives us an account of what gives us a grounding for our beliefs in ideas and theories in all aspects of our lives, from the most mundane to the most theoretical, in terms of practical usefulness.

Of course, it is important to draw a distinction between philosophical conclusions about what one can conclude from science on the one hand, and the character of the models we develop of human mind and behavior on the other. The first involves epistemology, broadly construed, and the second involves the construction of models in the special sciences like psychology and linguistics. In particular, these involve constructing theories of the developmental processes through which the child comes to understand her physical and social world, including her language. The question of truth may thus have different implications in the case of language or psychology than it does in the case of, say, intuitive physics or biology. We can imagine that a false, but useful, theory of physics built in to our perceptual and motor systems might be favored by natural selection because the question of how successfully these principles of "folk physics" might work in practice is separable, in principle at least, from our ideas about physical truth.

In the case of language, it is especially unclear whether there are external linguistic *facts* to which the cognitive system might only approximate. After all, language is itself a product of our cognitive system, rather than a pre-existing and mind-independent phenomenon. One reaction of this observation is that, here at least, pragmatism is unnecessary: truth is manifestly attainable, because intuition and reality are intimately entwined [Katz, 1981]. An alternative, and opposite, reaction is that pragmatism is the only option, because there is no mind-independent truth about how language works to which a theory of language could correspond. In this context, the question of whether the native speaker's ability to use a language should be thought of as *knowledge* at all comes into question. Knowledge after all is at least true belief, whatever other ingredients may be necessary, and there is no need to think of the ability to speak a language as consisting of a collection of propositions that are true of some external object.

In these fields, great battles have been fought over what it means to acknowledge the truth of Darwinian evolution and still try to develop a science of human mind, thought, and behavior. These battles have had an enormous effect on shaping the nature of the then-emerging new science of psychology, to which we now turn.

1.3 The development of psychology and the emergence of behaviorism

It is often said today that psychology as we know it today began in the late 19th century, and there is much truth to that: Wilhelm Wundt did indeed establish the first psychology laboratory in 1879 in Leipzig. However, psychologists at the time saw themselves, quite rightly, as part of a long intellectual tradition with taproots in two areas: first, in speculative philosophy, such as the work of John Locke, and second, in more recent laboratory work in physiology and medicine. The latter was bent on discovering the physical and chemical properties of the nervous system and on formulating quantitative relationships linking the physical and the psychic world (such as the Weber–Fechner Law, that the subjective ability to discriminate between physical stimuli, as measured by, for example, by the Just Noticeable Difference, is proportional to the magnitude of those stimuli). Darwin's revolutionary principle—that we humans are an integral part of the natural biological world and have become who we are as the result of a series of gradual changes shaped by natural selection—forced a renewed interest in the study of behavior, most especially intelligent behavior, in species other than *Homo sapiens*.

One of the first great American psychologists, G. Stanley Hall, wrote the following early in his career, in 1885:

> Experimental psychology... seeks a more exact expression for a more limited field of the philosophy of mind (while widening its sphere to include the physical, emotional, and volitional as well as the intellectual nature of man), to which its fundamental and, in the future, conditionary relation is not all unlike that of physical geography to history [Hall, 1885].

But the simple desire to create a discipline of psychology that could embed what we know about mind inside a larger view of mankind's evolutionary origin was not enough to do the trick; psychologists have been dealing with the challenges inherent in doing this over the course of the last 150 years. In Hall's day (as in ours!), one of the most important concerns was to understand the relationship between the kinds of behaviors described as instinctual in nonhuman species and those we see in ourselves and other humans. In the same paper, he cited a long series of detailed studies of the behavior of animal species, and emphasized the importance of this work for general and comparative psychology:

> [S]uch studies shed light on the nature, and often on the psychic genesis, of what is a priori and innate in man. Not only his automatic nature generally, with impulses, desires, and appetites, but conscience and the movement and rest of attention, are, in a sense, instinctive; so that so far from being inversely as reason, as is often said, much that makes the human soul really great and good rests on and finds its explanation in animal instinct [Hall, 1885].

We see, thus, that the goal of understanding the nature of what is known a priori and innately in man has been a central question in psychology since its beginning. The one apparent exception was the period of disciplinary dominance of behaviorism in the United States, extending from the second decade of the 20th century through the early postwar years. Behaviorism emerged in response to the German-inspired brands of psychology that grafted laboratory methods on top of introspectionist models that had grown from out-dated philosophy. The first strong statement of the principles underlying behaviorism came from a theoretical paper called "Psychology as the behaviourist views it," written by the American psychologist John B. Watson in 1913. Behaviorism rejected the reliance on introspection to obtain data, on the grounds that it was unreliable and unscientific; the goal of behaviorism was to convert psychology into an objective experimental branch of natural science that did not rely on subjective measurements or introspective reports.

Harking back to our earlier distinction between the ways of doing science on one hand, and the nature of the inferences we make about the human mind on the other, we can see that this version of behaviorism focused more on the former than the latter; introspection was rejected because it was not thought to be a sufficiently objective foundation on which to build a science. Theorists differed concerning how far this viewpoint had strong implications for the nature of the mind—but were in agreement that behavior was the domain of what could scientifically be *studied*. With the advent of radical behaviorism, whose chief advocate was B. F. Skinner, this changed. Skinner argued that *everything* an organism does—including having internal states like thoughts and feelings—constitutes behavior; therefore, in order to have a noncircular theory, thoughts and feelings should be included among the things-to-be-explained (explananda), not among the possible explanations (explanans). As a result, he concluded, environmental factors are the proper cause of human behavior, and learning (generally achieved through a slow process of operant conditioning) can have a profound effect on the nature of the resulting organism. Although Skinner did accept that nature places certain limits on what can be acquired through the process of conditioning, his emphasis on the importance of environmental factors led many to consider him to be advocating an extreme blank-slate position.

1.4 Logical empiricism

A parallel, and influential, movement in the first half of the 20th century called itself *logical empiricism*, whose goal was to find a synthesis of the empiricist thinkers of the 19th century, such as John Stuart Mill and Ernst Mach, and the revolutionary work on the foundations of logic, mathematics, and language developed by Bertrand Russell, Gottlob Frege, Ludwig Wittgenstein, and others. This movement, like any philosophical movement, had many variants and flavors, but one important theme that they all shared was an effort to locate certainties in *language* (typically, suitably regimented by translation from natural language into formal, logical languages, thus aiming to reveal the underlying logical form of natural language statements) rather than in innate ideas or in Kantian categories. If we are utterly certain of something, so certain that no counter-evidence could shake our belief, then that certainty must derive from some rule of the language system, not from experience. So, from this point of view, certainly does not arise because of the in-built structure of our minds but by linguistic convention. We are certain that, say, *dogs bark or dogs do not bark*, or that two plus two equals four, in the same way that we are certain that bishops move only along diagonals in chess. However, many times a person may violate such a rule, the rule still holds good—the person is simply making a mistake. And the rule holds good simply because it is true by convention—that is the way that the rules of the game, or the rules of language, are set up.

This line of thought led early versions of logical empiricism to make the blanket claim that all statements could be sorted into three types: those that were strictly empirical, and whose truth could therefore only be learned through the sense; those that were about language and its use; and those that were meaningless.

The logical empiricists of the 20th century also differed from earlier empiricists by being committed to establishing an explicit system of rationality, based essentially on logic and probability and focused on how empirical data could support the general laws or principles [Carnap, 1945a,b]. Basic observations are, we might assume, simply true or false; but most of the things we want to say, particularly in science, involve generalizations, typically going far beyond what has been observed. Logical empiricists realized that they needed to develop an explicit and quantitative account of how observation provides rational support for generalizations.

The truth-by-convention element of logical empiricism proved to be unsustainable. Logical empiricists hoped to translate theoretical claims, whether about subatomic particles, gravitational fields, or linguistic regularities, by logical analysis into claims about experience (e.g., as direct claims about the input to the senses, or at least as claims about readings obtained from scientific instruments). Such a translation of theoretical terms into a so-called "observation language" was required to avoid theoretical terms, and the scientific generalizations defined over them, being consigned to the realms of the meaningless. But such translations, and indeed, the very distinction between theoretical and observational terms, turn out to be unworkable. For one thing, there seems to be no direct relationship between individual theoretical claims and specific empirical observations; rather, entire "theories face the tribunal of experience as a whole" [Quine, 1951].

Yet the project of building a formal theory of learning, which the logical empiricists initiated, has proved to be enormously important, and is central to much debate in the foundations of the linguistics, and to the argument of this book. We shall see that one line of thinking has it that the logical empiricists project of learning general propositions from experience is, at least in the case of learning the grammatical structure of language from observed linguistic data, simply infeasible. If this conclusion is right, then it would seem that our knowledge of language must have some other source. On the other hand though, other theorists have maintained that the empiricist approach to learning is viable in the case of language and thus that linguistic knowledge does come, ultimately, from the senses. These are key themes, to which we shall return repeatedly below.

1.5 Modern cognitive science, linguistics, and the generative program

Behaviorism faded away in the 1950s for many reasons. It had aimed to root out any talk about things that were mental because it saw no way to deal with such talk in a scientific fashion, and it tried to persuade itself that it had no need to, either. But cognitivism came to psychology and to linguistics in the 1950s with a radically new understanding of what we might mean when we talk about mental actions or states: these are no longer based on introspection but on models that made sense to a new generation of scientist who understood both computers in the concrete and computation in the abstract.

1.5.1 The notion of computation

The idea of computation in the abstract has strong roots in work of the 19th century: at about the same time that evolutionary theory was being developed and the paradoxes lying at the core of mathematics were being discovered, scientists were beginning to study and formalize the notion of computation. Although there is some truth to the idea that computation has become important to us recently because of the ubiquity of inexpensive computers and the internet, this is a small part of a larger story. The nature of computation was a question that lay at the heart of the concerns of the earliest rationalists and empiricists. Some computations are logical in their nature, such as the steps that inevitably lead from a set of axioms and postulates to a proven theorem, while others are numerical, such as the calculations that predict the date of the next solar eclipse or transit of Venus.

Intellectual leaders of both the classical rationalists and empiricists believed that the notion that computation lay close to the essence of thought, and they said so in words that have remained famous. In *The Art of Discovery*, Leibniz [1685] wrote

> The only way to rectify our reasonings is to make them as tangible as those of the Mathematicians, so that we can find our error at a glance, and when there are disputes among persons, we can simply say: Let us calculate [calculemus], without further ado, to see who is right.

and Hobbes [1655] wrote

> By reasoning, I understand computation. And to compute is to collect the sum of many things added together at the same time, or to know the remainder when one thing has been taken from another. To reason therefore is the same as to add or to subtract.

The key proposal of these authors was that argument—and hence, thought—might be reconstructed by the application of rules that could unambiguously yield a conclusion, independent of the preferences or prejudices of the person applying the rules—just as in the case with arithmetic calculation. It is a short, but momentous, step to note that these rules might by applied not by a person, but by a machine—and therefore that such a machine would potentially be able serve as a model for human thought. The creation of modern logic, computability theory, and computer science in the twentieth century showed concretely how such a mechanical model of thought might operate.

The most famous of these developments was Alan Turing's notion of what we today call a Turing machine. With the help of this abstract—indeed, imaginary—machine, logicians and mathematicians got a much stronger hold on what it means to calculate, to define, and to prove. Turing machines, and the broader theory of computability of which they form a part, will prove important in some of the discussions in the body of this book. For example, a rigorous notion of computation allows the formulation of a rigorous notion of the complexity of an object, based on the theory of Kolmogorov complexity. This, in turn, provides the basis for a theory of learning and inference that works by finding the simplest explanation of the available data.

The development of the Turing machine, in conjunction with parallel work by John von Neumann on computer architectures and Claude Shannon in information theory, occurred at the same time as the death of behaviorism and the arrival of cognitivism in psychology. Indeed, young leaders in psychology and linguistics like George Miller and Noam Chomsky were strongly influenced by these developments in computational theory. In part because of these historical roots, the notion of computation is central to the project of modern cognitive science and the framework of cognitivism.

1.5.2 **Cognitivism**

Cognitivism is the proposal (or rather an expansive family of proposals) that the mind should be understood in terms of computational explanations of how information is encoded, processed, evaluated, and generalized by humans and animals. Behaviorists attempted to avoid explaining behavior in terms of internal states such as beliefs, desires, and inferences because—they argued—such accounts do not provide an *explanation* in the sense that they thought acceptable. Cognitivism aims instead to explain these and other notions in computational terms and to show that solid, substantial, and important sorts of scientific explanation are possible in such terms and probably only in such terms.

All psychologists and linguists alive today know that data (and in particular data that arrives through the senses) is entirely inert without principles or one sort or another to organize and animate it. Even just putting data into memory is a dynamic and active process; so too is retrieving it from memory, and so is comparing it, generalizing it, compressing it, and so on. Where theorists largely differ is in terms of the nature of the principles that organize the data, and where those principles come from.

Following much discussion by Noam Chomsky, the willingness to posit complex, sophisticated, and specialized computational machinery to the models developed by cognitivists has come to be known as rationalism, though the emphasis on the view that language is learned through an autonomous module has no more roots in classical rationalism than it does in classical empiricism.

A theory of human mind, thought, and behavior must have room both for sensory impression and information, and for the organization of that information; that organization does not come from the impression and information itself, and so, as the classical rationalists said, not everything in the mind comes from or through the senses.

1.5.3 The development of the generative framework

Classical generative grammar, initiated by Chomsky, began with the promise of a new kind of linguistic theory, one that could explain why a particular grammar was the right one, given a particular set of data. It may seem like this would be something close to a theory of learning—albeit an abstract theory of learning—and hence a theory that would be well in line with the empiricist framework. And yet the generative revolution in linguistics was accompanied by a metatheory which strongly rejected the empiricist standpoint, both methodologically and developmentally. This remarkable about-face stemmed from an initial focus centered more on questions of representation rather than questions of acquisition; the original goal was simply to provide an accurate formal characterization of the properties of language in the abstract. Determining the nature of the grammars acquired was taken to be logically prior to determining the process by which such grammars were in fact acquired.

By degrees, this evolved into the study of the universal characteristics of human language, and the belief that these universal characteristics would highlight properties of language that each language learner knew without ever learning them. These universal characteristics were thus assumed to be embodied in a Universal Grammar, encoded in a dedicated "language organ" [Chomsky, 1980] or "language acquisition device" [Chomsky, 1965]. Perhaps individual languages might turn out to be trivial variants of each other, with the common features and mechanisms across languages more significant than the differences. Indeed, Chomsky argued that language acquisition was more akin to growth than to learning—that is, that languages are not really *learned* at all:

> Language learning is not really something that the child does; it is something that happens to the child placed in an appropriate environment, much as the child's body grows and matures in a predetermined way when provided with appropriate nutrition and environmental stimulation [Chomsky, 1988].

This revolution has been so thorough-going that within many areas of linguistics and language development, the nativist framework has come to seem as axiomatic—both as a methodological starting point and as an account of language development—as the empiricist assumptions that once had been taken for granted. One of the goals of this book is to consider whether this revolution may have been premature. We argue for a return to the more commonsensical notion that the study of language is a straightforwardly empirical enterprise, like biology, and that language acquisition is primarily a matter of learning from experience, rather than the unfolding of a genetic program [Chomsky, 1980; Fodor, 1983] or the operation of an instinct [Pinker, 1994].

More recently, with the advent of the Minimalist Program [Chomsky, 1995], metatheoretic issues within the generative tradition have been thrown into some confusion, as we shall mention briefly in the final chapter. Nonetheless, it remains true that a strong nativist perspective is still dominant within linguistics and some areas of language acquisition research, and the assumption that there is an instinct, organ, or special-purpose acquisition device for language has been taken as a paradigm case for a broader emphasis on innately specified domain-specific modules across a broad range of cognitive domains [Hirschfeld and Gelman, 1994], a view which has been become central to some strands of evolutionary psychology [Pinker, 1997].

In this book, we aim to offer an alternative perspective, one which does not start from the assumption that the child begins the process of learning a language with a rich endowment of innately specified, language-specific knowledge. The child is not, of course, a blank slate; indeed, the child's (like the adult's!) cognitive machinery has been shaped by hundreds of millions of years of natural selection over complex nervous systems. But we adopt as a starting point the hypothesis that the child begins without innate knowledge or cognitive predispositions *which are specific to language*. That is, our sense of empiricism is that what children come to know *about language* comes through the senses—and, most importantly, comes from exposure to language produced by other people.

We recognize that the original arguments against an extreme empiricist approach still apply: it is self-evidently necessary for the mind to have some principles that organize and make sense of the data that comes through the senses. In what way does our suggested revival of the empiricist approach address these pitfalls?

1.6 Clarifying our program

A first clarification concerns scope. Classical debates between empiricism and rationalism blurred the distinction between two very different questions: on the one hand, issues surrounding the methodology by which knowledge can reliably be attained (problems, in modern terminology, of epistemology or philosophy of science); and on the other hand, issues concerning the psychological question of how children acquire their native language in practice.

Today, it is possible to see more clearly than earlier empiricists and rationalists did that there is a healthy distinction to be drawn between these issues. Questions about how to do science are questions of method, debated most profitably by the scientists engaged in research (though often with the help of sympathetic or critical philosophers who observe from the edges). Questions about psychology focus on how the human mind functions and operates. This distinction will be central to our discussion in this book; we wish to show ways in which current work in the cognitive sciences can better inform both our ways of doing science, and our theories about the human mind.

One important methodological question is whether the study of language is more similar to empirical science or to mathematics [Katz, 1981]. In some respects, it is self-evident that the study of language is an empirical science. Every language studied by a linguist presents new challenges that come unexpectedly, as far as the linguist was concerned. Methodologically, the field of linguistics *learns* about what a language can be by the study of each new language.

Furthermore, we know that the language we speak natively is a historically contingent and conventional system, subject to continual change, and the range of the world's languages exhibits stunning diversity (e.g., Evans and Levinson [2009]). Indeed, it is this diversity that leads many to become linguists early in their careers. The variety of languages has, since von Humboldt's day, been compared to the diversity of the living world, and scarcely governed by a priori mathematical principles. To be sure, biological diversity is not without limit: from D'Arcy Thompson onwards, biologists have also been interested in qualitative and quantitative patterns across species. Such patterns might be expected, by analogy, across languages also. Yet, despite such patterns, the study of language appears, at least at first blush, to be an empirical science par excellence: ~~our~~ limits of our imagination are always outdone by the next careful study of a newly discovered language.

From a psychological point of view, the wild and capricious variety of human languages appears to stand in contrast to the much more invariant character of number, perception, or geometry. Thus, although nativism about the latter may be credible, it seems *prima facie* implausible when it comes to language: it appears, rather, that the primary challenge of the child is to learn the spectacularly subtle and highly idiosyncratic intricacies of the specific language or languages to which she is exposed. Of course, a nativist would reply that the apparently vast differences between distinct languages are only apparent—that, at a deep level, all languages share certain strong commonalities or universals.

Resolution of this issue requires, ironically, as much empirical research as it does formal analysis. The formal and technical nature of much of this book should not make the reader misunderstand our project: this is mathematics in the aid of empirical science, not as an end in itself.

1.6.1 Our approach

Our general approach is strongly empiricist methodologically and weakly empiricist psychologically. We suggest that linguistics, as a science, will best progress by using a methodology that favors constraining and testing formal theories against data. Much of our focus in this book is on the first half of that (developing, defining, and testing formal theories), rather than the second half (acquiring and using appropriate data). This is because that is where our expertise lies, and where we can make the strongest contribution. Both halves, however, are key; and it is worth saying a few words about the data before we go on.

There are three sorts of data that are being actively employed in linguistics currently: (i) introspective judgments, reported by linguists; (ii) analyses of naturalistic corpora (that is, language use that existed before the linguist approached the subject); and (iii) controlled, experimental work in laboratories studying language processing in production and perception.

That data should not solely (or even mainly) consist of introspective judgments about linguistic intuitions, as is standard practice in much of generative linguistics; although these intuitions can be a useful tool in guiding the formation of theories, using them as the primary or only source of empirical support for a theory is problematic.[1]

Not only is there considerable variation among speakers, to the point where many native language users may find acceptable what others find thoroughly unacceptable, but intuitions may be murky even for a single speaker. Relying on linguistic intuitions—or even treating them as if they constitute the same degree of support as data arrived at in a more scientifically rigorous manner, such as survey data—has the effect, therefore, of reifying variable or marginal intuitions into something far more certain or well-defined than they actually are. It is a problem when the resulting theories, constructed to account for data that may not in fact even be accurate, become incorporated into the set of accepted principles of linguistics.

There are a variety of methodologies that *are* well-suited to the investigation of linguistic phenomena, many of which are already employed throughout

---

[1] See Wasow and Arnold [2005] and Gibson and Fedorenko [2012] for similar arguments.

cognitive science. These include reaction-time experiments, eye-tracking paradigms, corpus analyses, and survey data.[2]

All of these result in a more statistically valid and nuanced picture of grammatical acceptability than that provided by intuitions.[3] Nevertheless, these other methods have the drawback that (like linguistic intuitions) they often yield data only on the particular constructions or phenomena in question. Though this may be interesting in its own right because syntacticians are often focused on the question of which grammatical formalism or theory best describes an entire language, it is, of necessity, limited in scope: every theory includes some phenomena that it can explain easily and some that can only be accounted for by more *ad hoc* measures.

Syntacticians tend to focus on a narrow range of linguistic issues that are thought to be interesting or important: island constraints, parasitic gaps, quirky cases, and the like. Though we agree that these phenomena are indeed interesting, we think that an exclusive focus on these extreme cases is methodologically suspect, particularly if the underlying judgments have not been validated thoroughly. What is often desirable is some approach that can objectively decide between theories on the basis of how well they account for observed natural language usage, in its full variety: globally, rather than on the basis of a few cherry-picked special cases. This book discusses several variations on such an approach, which relies heavily on computational and mathematical machinery, sometimes in combination with empirical observations and linguistic corpora.

On the psychological side, we call ourselves *weakly* empiricist to differentiate from two approaches that ours should not be confused with. The first is that of the behaviorist, who has traditionally claimed a much weaker role for internal states—and a much weaker innate apparatus—than we are comfortable with. The behaviorist does not play a major role in cognitive science today. The second approach that we do not follow is that of the connectionist, to the extent that the connectionist claims a more impoverished representational ability than we do.

The term *connectionism* has been used to cover a wide range of approaches to problems of cognition, learning, and the modeling of neural processes, and more than one of the authors of this book have embraced, or at least seriously

[2] On reaction-times, see Spivey and Tanenhaus [1998]; on eye-tracking, see Just and Carpenter [1980]; Tanenhaus and Trueswell [1995], and Altmann and Kamide [1999]; on corpus analyses, Nunberg et al. [1994]; Lohse et al. [2004], and Levy [2008]; and on survey data, Terence Langendoen et al. [1973] and Wasow and Arnold [2005].

[3] See Sprouse and Almeida [2012] for a different view.

explored, properties of connectionist systems [Goldsmith, 1993; Christiansen and Chater, 2001]. Some connectionists are more aligned with psychologists (e.g., Rumelhart and McClelland [1986b]), while others are more aligned with computer scientists (e.g., Feldman and Ballard [1982]). All connectionists see their intellectual roots as going back to the pioneering work of McCulloch and Pitts [1943], and Hebb, in the 1940s, and to Rosenblatt's perceptron learning algorithm [Rosenblatt, 1958]. Many were influenced by the Parallel Distributed Processing Group several decades later [Rumelhart and McClelland, 1986b].

Broadly speaking, the connectionist perspective seeks to explain language (and cognition more generally) as emerging out of neural processes consisting of interconnected networks of simple units upon which statistical computations are performed. Most research within this perspective utilizes neural networks in which information is represented by the strength of connection weights between units, and learning consists of modifying those weights. Formally, connectionist networks are equivalent to nonlinear function approximators, with the weights corresponding to the parameters; learning is equivalent to searching through the space of weights for a function that minimizes error on a training dataset.

There are two claims associated with the connectionist perspective that are especially relevant to our purposes here. First, although connectionism is sometimes discussed as if it assumes no prior biases or constraints at all, this is not true: as we have already mentioned, there is no such thing as an unbiased learner. For connectionists, prior assumptions are built implicitly into the initial architecture of the networks, the initial setting of the weights, and the learning rule (which generally favors uniform weights or smaller ones corresponding to smoother and simpler functions). Second, the underlying representational assumption is that there *is* no explicit representational structure; representation is implicit and emergent. In particular, the connectionist perspective does not take the existence of formal linguistic entities like grammars seriously as a construct to be modeled. It is here that we depart most radically from that tradition. Interest in connectionism grew in part in response to the nativist viewpoint of generative linguistics but threw out the representational baby along with the nativist bathwater. We believe that it is important to investigate the possibility that knowledge is structured (perhaps in the form of grammars, perhaps in some other form), while still being learnable from data in the environment, given only domain-general constraints on that learning.

We have discussed what we are *not*: behaviorists or connectionists. In particular, we do not believe there is such a thing as an unbiased learner. The criticisms of classical empiricism, dating back to Descartes and Kant, are not without merit, needless to say; we believe it is indisputably true that all learning takes place within the context of principles that organize the sense data we receive. Indeed, learning language—as with any problem of induction—is logically impossible to solve without the existence of some sort of overarching constraints [Goodman, 1955; Quine, 1960; Wolpert and Macready, 1997]. For us, the real question is what the *nature* of these constraints or biases are. Where we depart from the more nativist tradition in generative linguistics is that we see no reason to presume that all or most of the interesting constraints on language learning are language specific.

Because we are biological organisms, derived via a process of evolution from ancestors who had rich cognitive abilities but no language, we believe that the more parsimonious explanation is that our language abilities—even (or especially) the abilities underlying any linguistic universals that might exist—are built on an already-existing cognitive and perceptual infrastructure. This is not an ideologically firm position; if it were to be established that some phenomenon or ability could only be explainable by the existence of a language-specific mechanism, we would accept it; but we do not believe that such a standard of proof has been reached. As we will see in the next chapters of this book, at least one argument that is typically taken to prove the necessity of innate language-specific knowledge (the famous "poverty of the stimulus" argument [Clark and Lappin, 2011]) in fact only proves the necessity of innate constraints of *some* sort. We believe that it is most sensible and parsimonious to proceed under the assumption that our linguistic abilities are *not* the result of a language-specific mechanism and then see how far that takes us.

In this sense, we share "the desire to reduce any language-specific innate endowment, ideally to a logical minimum" expressed by Berwick et al. [2011]. But although in this respect we are in harmony with the expressed principles of modern Minimalist and Biolinguistic thinking [Boeckx and Grohmann, 2007], we differ radically in the methodologies we use and the conclusions we draw. What this means in practice is that we begin with the assumption that human learners are equipped with relatively powerful learning mechanisms, involving the ability to search (possibly through the use of heuristic methods) through a large space of possible explanations, theories, or grammars, to find the one that best explains the linguistic data they see; that these learning mechanisms rely at least in part on statistics, enabling graded generalizations; and that the mechanisms are constrained by initial assumptions or biases that are domain general, deriving (at least initially) from other aspects of our cognitive or perceptual system. We conceive the objective and nature of language acquisition in a probabilistic way: we suggest both that the nature of the *learning system* is inherently probabilistic (i.e., that it consists of performing statistical inference about the observed data) and that the nature of *linguistic knowledge* is also probabilistic (that "knowing" a grammar does not mean being 100 percent certain that it is the correct explanation for the data but simply that it is highly likely that that is the case). We also conceive of the *grammar itself* as containing probabilistic information—information not just about what can be said but also about how likely particular words and sentences are to occur. That said, for technical reasons it is sometimes convenient to switch to a nonprobabilistic grammar, as this can simplify the mathematical analysis, as we do in Chapter 4.

We adopt a methodological approach that derives from Bayesian and Minimum Description Length approaches to learning and relies strongly on an abstract notion of simplicity. Abstract in this case does not mean vague or imprecise—on the contrary, we are strongly committed to using mathematically and computationally precise models. In the absence of this technical detail, discussions at such a high level of abstraction run the risk of becoming mere speculation. This precision pays off in two respects: one computational and one mathematical. From a computational perspective, we can implement, at least in part, the proposed learning mechanisms and see the extent to which these are successful on natural language corpora. Mathematically, we can give proofs that show that, under certain assumptions, such mechanisms are guaranteed to learn languages. These approaches provide objective and rigorous ways to assess what is learnable given the information in a child's linguistic input and the hypothesized biases and learning mechanisms.

1.7 Linguistics

What brings the four of us together, and what unites the work that we describe in this book, is the belief that learning plays a central role in the way language is acquired and that the study of learning should play a central role in the way linguists do their work. This is not a statement of credo but rather a conclusion based on our experience. When we speak of "the study of learning," we refer to what has been established about learning in a number of fields and approaches that are different from linguistics and also to what has been discovered about learning that is specific to language. By its very nature, learning involves the interaction of an organism—let us simply say a *person*—with what is going on around her, and learning takes place when the person can internalize some structure or organisation that she is able to discern in that experience.

A good deal of emphasis over the last several decades has been laid upon the ways in which linguistics can shed light on what aspects of mind might be innate. The general principles that might be innate differ a good deal in different linguists' estimation, but clues to innateness lie both in the implausibility of ever finding a learning theory that could account for the principles and in the appearance and reappearance of these principles in many languages. The logic of that research is undoubtedly attractive, but it seems to us that what the science of linguistics needs is a forum in which claims about innateness and claims about what is learned can be judged in the light of day, without one side or the other claiming the high epistemological (or philosophical or mathematical) ground.

There are any number of voices in linguistics expressing similar sentiments, and those perspectives have had an impact on work done under the rubric of laboratory phonology, for example, or experimental syntax. But there is more that we could hope for. Advances in computational linguistics have rarely been taken—as we think they should be—as challenges to linguistics to see if tools developed in empiricist contexts might inform and restructure the way mainstream linguists think about language [Abney, 2011]. In a few cases, this has indeed happened: there are linguists who develop models of inflectional morphology, for example, with full awareness of the computational structures that have been developed for practical ends, to mention just one example. But syntacticians rarely if ever think about what syntactic theory might look like if the language learning faculty led to a grammar of English or Swahili in which there were far more categories than are countenanced in contemporary syntactic theory.

But we should not be taken to be championing a view of language with many more categories and fewer explanatory principles. That might be the way reality works; it might not be. An empiricist perspective, as we show in detail in this book, is deeply committed to exploiting the power of simplicity. That perspective puts so much emphasis on it because it operates not only on the scientific level in which one theory competes with another, it operates as well in the reasoning used by the learner who is looking for the best account of the data she is presented with.

Our goal, then, is to bring learning back into the set of tasks that the linguist's Universal Grammar must be deeply involved in. We are the species that learns better and faster than any others; our history in the last ten thousand years has shown that clearly, as each generation has surpassed the one that preceded it. Perhaps the complexity of language that linguists seek to analyze has nothing to do with our abilities to learn. But we would not bet on it.

1.8 The field of linguistics

A word on what we take the term *linguistics* to cover. We intend it to be interpreted in a broad way, to include the systematic and scientific study of language and the ways in which language is used. In practice, the ways of studying language have focused on psycholinguistics, the study of individuals using language in real time; on sociolinguistics, the study of how language is used by individuals as members of social groups, often as members of several groups simultaneously; and on language as a structured system, abstracting away from the context in which utterances are used by individuals and groups. This third category, the proper domain of general linguistics, includes three principal subparts. First, there is the study of sounds, manual signs, or written language as the external manifestation of language, which is to say, phonetics and phonology. Second, there is the study of how small, meaningful, or, more generally, structured pieces of expression are put together (by concatenation or by methods more complex) to form words, phrases, and sentences. This is the domain of morphology and syntax. And third, there is the study of how the meanings of words, of subword pieces, and of larger phrases composed of words can be systematically analysed, and this is the domain of semantics. General linguistics, understood as these last three parts, can be, and is, studied in a multitude of ways, varying a good deal in the degree to which proposed accounts are couched in formally explicit ways. Just how formally explicit an account is may sometimes be hidden or left as an open question to be answered in the future. This is often the case that we find when a researcher cannot determine what aspect of his analysis is intended to hold for all languages and what aspect is intended to be specific to the language he is analysing; which is to say, all kinds of linguistic analysis, but most especially the work done in general linguistics, must be mindful of the distinction between, on the one hand, characteristics that we believe to hold of all languages, by virtue of either logic or empirical fact, and on the other, characteristics which we believe hold of one or more individual languages but which we understand are not universal across all languages and which must therefore be explained as learned by speakers in the course of their acquisition of their native language.

The reader may be puzzled by the lack of detailed analyses of particular languages in this book, and so a word or two of explanation is in order to describe the relationship, as we see it, between the traditional fields of linguistics and the research program(s) presented here. This book is about approaches to language learnability and acquisition; Chomsky was the first to put language acquisition at the center of linguistic theorizing and for good reason. The range of possible analyses for a given linguistic phenomenon is really endless; and since the beginning of linguistics, this has posed a serious methodological challenge. As, Bloomfield [1933] believed that when universal linguistics finally comes, it "will be not speculative but inductive," our intent has been to provide a way to balance between the two. The work presented here focuses on the procedures of analysis, as we think that it is only by integrating the study of learnability and language acquisition into linguistics that real progress can be made.

1.9 Going forward

There is a certain amount of technical apparatus needed in order to develop in detail the proposals that we will make over the course of this book, and Chapter 2 Offers a brief overview of these conceptual tools.

Chapter 3 discusses how notions of probability and simplicity have been used to model both the linguist's and the child's problem of building a grammar of language and builds the linguistic case for a new empiricist approach to language. Following that is Chapter 4, which addresses learning and computational complexity from an abstract perspective, presents mathematical results relevant to the learnability of specific classes of languages, and formalizes the notions of generalization and analogy; in this chapter, we draw links between the ideas of distributional learning and a specific notion of simplicity of a grammar.

This is followed by Chapter 5, which presents two famous problems in language acquisition—the argument from the poverty of the stimulus and the problem of no negative evidence. We will present theoretical results showing that an "ideal" simplicity-based learning can in principle learn from positive data only, and we illustrate briefly how this approach can be scaled down to examine the learnability of specific grammatical structures. This leads naturally to Chapter 6, in which we present a specific implementation of a model that addresses both of these famous problems and illustrates one implementation of our general modeling approach. We show what can be learned from the corpora of typical child-directed speech, given certain built-in representational assumptions, and discuss how those assumptions constrain learning and to what extent they drive our results. Finally, in Chapter 7, we conclude with a general summary and integration of the perspectives presented throughout the book, and end by drawing some conclusions for the direction of future research.