Claims about the relationship between language and thought have a long and venerable history. The Holy Roman Emperor Charlemagne proclaimed, "to have a second language is to have a second soul." One of his successors, Charles V, declared, "a man who knows four languages is worth four men." Frederick the Great of Prussia had a more specific set of hypotheses. He quipped, "I speak French to my ambassadors, English to my accountant, Italian to my mistress, Latin to my God, and German to my horse."

Others, however, have been much less enthusiastic about the idea of language shaping thought. Shakespeare's Juliet tells us that "a rose by any other name would smell as sweet," suggesting that what you call something does not change your experience of it.

These days, the idea that language shapes thinking is most commonly associated with the writings of American linguist Benjamin Lee Whorf (1897-1941). By trade, Whorf was a fire prevention engineer for the Hartford Fire Insurance Company. His early interest in languages led him to study with the linguist Edward Sapir at Yale. Whorf went on to publish many influential academic articles and to serve as a lecturer at Yale (though he never gave up his full-time job at the insurance company). Fascinated by how languages differ from one another, Whorf became an ardent advocate for the idea that speakers of different languages think differently, and wrote many influential articles exploring the topic.

Whorf's ideas enjoyed popularity for several decades after his death in 1941. But then, the winds started to shift. Noam Chomsky emerged as a powerful force in linguistics, and Chomsky's approach was in many ways antithetical to Whorf's.

Chomsky envisioned a formal, mathematical approach to language. Instead of simply going around the world documenting different languages (an activity he dismissed as similar to "butterfly collecting"), he believed linguists should be trying to build a generative model of language. A generative model is a formal description of the grammar—a set of rules that (theoretically at least) could be used to produce all of the grammatical sentences in a language and none of the ungrammatical ones.

Human languages are exquisitely complex, and specifying a complete generative structure for any language proved an elusive task. But there is an apparent paradox. If language structures are so complex, how are children able to learn them so quickly? After all, every normally developing child, regardless of what culture they're born into, will master the complexities of the languages around them within a few years (though the amount of time necessary does depend on the language). How are children's brains succeeding at this?

Chomsky had a beautiful idea. He proposed that underlying all of the world's languages is a universal grammar. This hypothesized universal grammar would be an abstract structure that could be used to generate the grammars of all of the worlds' languages. And the reason children are able to induce the complex structures of their particular languages so efficiently is because they are born with a genetically specified language acquisition

device that comes with this hypothesized universal grammar already built in. That is, children aren't really learning language *per se*. The generative grammar engine and the blue prints were already in their minds at birth. All children need to do is learn the vocabulary and adjust a few parameters to match the language they hear, and voila!

But what about all the easily observable differences between languages, you might ask? Don't these argue against the idea of a universal grammar? Chomsky maintains that this linguistic variation is only on the surface, that languages only *appear* different. To illustrate the idea, he imagines a Martian scientist who comes to Earth to observe the world's languages: "A rational Martian scientist would probably find the variation rather superficial, concluding that there is one human language with minor variants."

Even when not appealing to extra-terrestrial intelligence, Chomsky is clear on his claim that languages do not importantly differ from one another: "It must be, then, that in their essential properties, and even down to fine detail, languages are cast to the same mold." Setting the call to arms for his research program, he writes, "we have to ⋯ show that the enormous apparent variety, is a kind of superficial variation."

The idea of a universal grammar gained great popularity, and for decades many students learned that Noam Chomsky discovered the universal grammar of human languages, or that he proved that all languages are indeed the same. Claims of linguistic universality are also rife in popular culture, and are often cited as an a priori counterargument to the idea that languages shape thinking. If languages don't *really* differ, then why would you even bother asking whether speakers of different languages think differently?

Let's examine why claims about universal grammar are either demonstrably false or simply irrelevant to the question of whether language shapes thought.

One oft-repeated claim is that linguists have examined the world's languages and have determined that they do not differ. This is simply false. At the moment we have good linguistic descriptions of only about 10% of the world's existing languages (and we know even less about the half a million or so languages that historical linguists estimate have existed in the past). There are regularities and trends across many languages, to be sure, but regularities are not the same as universals. And even in our limited sample, there are wild differences.

In 2009, linguists Nicholas Evans and Stephen Levinson examined decades of actual linguistic description work done around the globe, looking closely for purported universals to see how many might hold up. In a landmark paper titled "The Myth of Linguistic Universals" they summarized the body of evidence as follows: "Languages differ so fundamentally from one another at every level of description (sound, grammar,

lexicon, meaning) that it is very hard to find any single structural property they share." That is, far from finding that languages share all or even many structural properties, they had trouble finding even one structural property that could be meaningfully said to exist in all languages. In short, if someone tells you that linguists have examined all the world's languages and found they do not differ, hold on to your wallet.

So universal grammar is not an empirical fact. But that doesn't necessarily mean we must give up on it as an idea. One way to treat it is as a mathematical claim: perhaps there exists a more abstract formal level of description (one we haven't yet developed) at which apparent differences between languages disappear. This is the abstract structure that would be so readily apparent to our friendly Martian scientist. Other than asking a Martian, how would we know if this claim is viable? Someone would have to actually come up with such a formal description. We would need, for example, a computational model that could, from a single formal description (allowing a few adjustable parameters), generate all the world's languages.

No such formalism exists. In fact, there is not even a working formal description that captures the structure of English. Claims about the formal equivalence of languages are interesting, but grossly premature at best.

Whether languages are formally equivalent is also irrelevant to the question of whether language shapes thinking. When it comes to human cognition, even things that *are* formally equivalent are rarely *psychologically* equivalent. This insight earned Daniel Kahneman the Nobel Prize in Economics, transforming the field of economics along the way. For understanding human cognition, the question of whether languages really differ is an empirical one. Do differences between languages have measurable cognitive consequences?

In other words, does speaking one language versus another change how people attend to, remember, perceive, represent, and reason about the world? If we want to understand the human mind, this is what matters. Even if the world's languages *appeared* to be the same to a Martian, those appearances wouldn't change the facts of human psychology.

The data we will review in this book demonstrate that in the workings of *human* minds, languages really differ from one another. As for Martian minds, it remains an open empirical question.

Chomsky was and remains an intellectual giant in linguistics. With the rise of Chomskyan linguistics, Whorfian claims fell sharply out of favor. By the 1980s the topic was seen as taboo, and the claims dismissed as ridiculous. Philosopher Jerry Fodor colorfully expressed this view, writing: "I hate relativism more than I hate anything else, excepting, maybe, fiberglass powerboats."

While some of this was a matter of intellectual fashion, there was also the big problem of evidence—in particular, the lack thereof. Whorf's claims about linguistic relativity included many fascinating examples about how languages differ, but lacked the necessary ingredient of empirical evidence to show that speakers of those languages actually think differently. Just observing that languages differ is not enough to show that their speakers differ in how they think.

As an example, consider the distinction French makes between the formal or polite form of *you* (*vous*), and the informal or familiar version of *you* (*tu*). A typical argument might go like this: The French tend to address older people using the formal *vous* and not the familiar *tu*, therefore French speakers must be more deferential to elders than English speakers.

But how do we know whether French speakers actually feel more deference for elders? It's tempting to say, "Just look at how they talk! They address elders using the formal *vous* and not the informal *tu*." This just takes us in a circle. We haven't done anything to add evidence for our claim. We have only restated that French speakers talk differently from English speakers.

Just because people talk differently does not necessarily mean they think differently. We can think about things without saying them out loud. Indeed, language is very sparse. Most utterances include only a small part of what we know about a given situation. Many things are left implied, or are commonly understood, with no need to say them out loud. For example, I can say, "It rained yesterday," without adding, "only outdoors, not inside my house." Although I know it only rained outside and not inside, I am not likely to mention it. I may also know that it rained only in the morning, only in the mountains, that the rain was heavy, the droplets felt warm on my skin, and that it means I'll be hunting for chanterelles this weekend, but none of these things needs to be specified in my sentence for it to be a normal, grammatical English sentence. That is, I can know, notice, and remember many more things than my language strictly requires.

It is possible, then, that the distinctions we express in language are only a small subset of what we encode and remember. In fact, it could be that all people notice all the same things, remember all the same things, organize the world in their minds in exactly the same ways, but choose different subsets of information to talk about, because that is the norm in their language. Everyone could think the same, but talk differently.

To figure out whether languages indeed shape the way their speakers think, we need to measure how people think. A starter experiment would go like this: expose speakers of two different languages to the same stimuli, and see whether they behave differently as we might predict from the patterns in their languages.

For example, if Tarahumara doesn't make a distinction between the colors that English calls blue and green, might Tarahumara speakers judge blue and green patches as more

similar to one another than would English speakers? Might English speakers be less likely to confuse a blue and green patch in memory? In other words, we need some measure of how people think about these colors, beyond knowing that the two languages differ in what words they use.

Showing that people from two language groups differ on some cognitive task is an important first step and an interesting finding in its own right. But it is not enough to establish that language is a causal force. Maybe there is another reason that English and Tarahumara speakers treat colors differently? To see whether language really has causal power over people's color judgments, we need to show that changing people's language experience in some way changes their judgments. For example, would teaching Tarahumara speakers English color terms change the way they judge and remember colors?

Modern research on language and thought takes this two-step approach: First test whether there is a difference in behavior between people from two language groups on some cognitive task. If so, then test whether changing some aspect of their language experience affects their behavior on the task. The claim that differences between languages can lead to differences in thinking requires this complement of evidence.

With the rise of cognitive psychology in the 1970s, researchers attempted a few experiments to measure whether language shapes thinking. A few early studies produced equivocal results. Considering the taboo nature of the topic, most quickly stopped trying. To many, the topic was decided and dead; language did not shape thought. When I took my first university course on language, Steve Pinker's "The Language Instinct" was the textbook, and according to Pinker the idea of language shaping thought was "a conventional absurdity," straight from the dark ages. Pinker writes "The idea that language shapes thinking seemed plausible when scientists were in the dark about how thinking works or even how to study it."

Psychologists, for the most part, were happy to assume that cognition is the same everywhere. Assuming that cognition is universal makes the life of a research psychologist a lot easier. The vast majority of psychology experiments are carried out at elite North American universities. The participants are college students, usually freshmen or sophomores taking an introductory psychology course and participating in studies to fulfill a course requirement. Top-flight researchers don't think of themselves as studying how sophomores think, but how *humans* think. If cognition is universal, then studying American sophomores should produce results that generalize to all humans.

College students, for their part, are great experimental subjects. Not only are they conveniently located at your university, they are also a generally well-behaved homogeneous group of people who usually follow directions and produce data that are normally distributed and not too noisy. Plus, if you only test middle-class English-

speaking North-American college students, you are in little danger of discovering cultural differences that might challenge your assumption that cognition is universal.

Not everyone gave up on the idea of cross-linguistic differences. Linguistic anthropologists and anthropological linguists working at field sites around the world continued to collect fascinating observations about how languages and their speakers differ. Occasionally, a paper would appear in an anthropology journal or technical report. But in mainstream psychology the topic remained taboo. Psychologists live and die by controlled experiments. When linguists and anthropologists returned from their field sites with descriptions of intriguing linguistic structures and unusual cultural practices, the psychologists just shrugged it off. It wasn't data from controlled experiments, and so it didn't count as evidence.

Slowly, things within psychology started to change. Researchers studying learning were discovering the transformative power of practice, experience, and expertise. Study after study showed how experts see the world differently from novices. Cognition didn't seem to be so fixed anymore. It could change with experience.

If other types of experience could shape cognition, why wouldn't the experience of speaking a language shape thought? Considering the thousands of hours of practice we get with language, it would be a real anomaly if habits in language did *not* have some effect on thinking.

In other corners of psychology, researchers studying attention and memory were discovering striking limitations in what we could pay attention to and remember. It didn't look at all like people were perceiving reality in full detail and remembering things as they happened. Instead, it looked like we were able to attend to and encode only small pieces of the world around us, and our brains did a lot of work to fill in the details.

If each of the world's languages requires its speakers to pay attention to a different set of things, is it really possible that we all pay attention to the same things and just talk differently? For this point of view to work, every one of us would have to (at the minimum) pay attention to and encode all the information required by each of the world's 7000 or so languages. That is a very large (and often contradictory) set of distinctions to attend to! Do our brains really encode so much information?

Most of us intuitively feel like we get a complete and clear picture of the visual world. It seems so easy: you open your eyes and you see. Nothing to it. Other people have eyes too, so they should see the same things that you do. Yet decades of research on attention and memory have shown that this feeling we have of seeing the world clearly, as it really is, is a grand illusion. We are able to process only a small fraction of the information that bombards our senses, and we remember very little of what we experience.

Here's a quick test. Remember the U.S. penny? You have seen it thousands of times. You may even have one with you right now (but don't look at it – that would be cheating). Can you remember what it really looks like? Take a moment and try to draw the penny from memory. I'll wait.

Most people struggle to draw the penny. And it's not just because most people aren't very good at drawing. Would you be able to recognize a real U.S. penny if I showed you one? The picture on the next page shows many possible pennies. Only one of them is the real deal. Can you recognize the true penny?

If you didn't pick the right penny, don't feel bad. Very few people remember the intricate details necessary to succeed on this test. This is a healthy and normal adaptation. The world around us is infinitely complex, and every moment we are bombarded with much more information than our brains can process. We must select what to attend to and process further.

This idea applies not only to remembering visual details. Take Chekhov's *Uncle Vanya*. What kind of uncle is he? Maternal or paternal? Related by blood or marriage? Unless you're a real Chekhov aficionado, you are not likely to answer this question from memory. Even after reading *Uncle Vanya* in Russian, watching the film in English, and seeing it live on stage in St. Petersburg, I couldn't bring to mind what kind of uncle he was when I first posed this question to myself (we'll leave the question of why I pose such questions to myself for another time). But if you have ever come across *Uncle Vanya* in Mandarin, you would easily know that he is a mother's brother, because Mandarin has different words for all these different types of uncles and then some. In Mandarin, the very title of the play reveals the consanguine maternal link – you can't miss it.

Our experience in the world includes an ever-present stream of language. By the time infants are born, they have already learned a great deal about the sound properties and regularities of their language from the noises that reached them in the womb. From the very beginning of life, we are awash in parallel streams of linguistic and other perceptual information, and children are avidly learning and processing both at the same time. Studies monitoring infants' brains show that language starts to organize how infants process visual information long before they even start to speak.

As children develop and learn, language serves as an attentional guide. Languages encode aspects of the world that generations of people before us have found useful to pay attention to. Each language functions as a culturally created guide to attention, a way of highlighting certain aspects of the world, at the expense of others.

It took time for researchers from psychology, anthropology, and linguistics to start to converge. More psychologists started taking notice of the findings coming from field sites (and not just from sophomores). And more anthropologists and linguists started adopting the experimental methods of the psychologists. As this new cross-disciplinary intellectual

community began to emerge, the pace of research accelerated. After centuries of a priori arguments and debate, there was finally evidence.

Before we get to the studies that re-opened this question, let us first lay to rest Benjamin Whorf. The new wave of research is often called neo-Whorfian, and I am called a Whorfian, a neo-Whorfian, or a Whorf apologist (and that is just to my face). However, in my daily life and research practice, I rarely think about "the Whorfian question" or "the Whorfian hypothesis," for a number of reasons.

First, many people around the world wrote and thought about this set of questions long before Whorf. Whorf is certainly an important figure, but naming a topic with such a long and diverse intellectual history after a single 20th century American strikes me as a bit of revisionist imperialism. (Though if we're going to play that game, I would prefer we call Charlemagne a proto-Boroditskyan).

Second, there is no such thing as "the" Whorfian question or "the" Whorfian hypothesis. Whorf asked many different questions, proposed many different hypotheses, and made many different and sometimes contradictory claims on the topic (as have many others). If anyone tells you that they know what "the real" Whorfian question is, you will definitely want to double-check for your wallet.

As with many other early claims about the mind, Whorf's hypotheses about the relationships between language and thought are typically not specific enough to be testable by modern empirical science. Although he is often pilloried for this lack of specificity, let's keep in mind that cognitive psychology did not yet exist when Whorf was writing. Very little of the analytical tools and methodological arsenal we take for granted today were available when Whorf was alive.

Further, in the decades since Whorf's death (actuarial maturity?), we have learned many things about how the mind works that greatly inform and constrain our hypotheses about how language can shape thought. And we have developed much more precise theoretical and analytical tools for reasoning about cognition. As a result, the set of questions we can formulate and test today are far more advanced and precise than what was possible in Whorf's time. So in this book, when we formulate scientific hypotheses to test, we will do so informed by modern science.

Figuring out what Whorf really meant or thought makes for an interesting historical question. But for a cognitive scientist, the interesting question is how the mind actually works. All scientists, me included, are wrong about some things and right about others. A conclusion that says "Whorf was right" or "Whorf was wrong" is much less interesting to me than finding out how language and thought actually interact and what this tells us about the human mind.

There are some deep and fascinating mysteries in the intellectual history of this topic. But for the purpose of understanding the mind, the impetuous imp inside me is jumping up and down shouting, "Who cares what some old dude thought or didn't think? Let's do some science and figure out how things really work!"