LANGUAGE VS. REALITY

WHY LANGUAGE IS **GOOD** FOR LAWYERS AND **BAD** FOR SCIENTISTS

N. J. Enfield

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Enfield, N. J., 1966– author.
Title: Language vs. reality : why language is good for lawyers and bad for scientists / N. J. Enfield.
Description: Cambridge, Massachusetts : The MIT Press, 2022. | Includes bibliographical references and index.
Identifiers: LCCN 2021010579 | ISBN 9780262046619 (hardcover)
Subjects: LCSH: Communication—Social aspects. | Cognition. | Psycholinguistics.
Classification: LCC P95.54 .E54 2022 | DDC 401/.9—dc23
LC record available at https://lccn.loc.gov/2021010579

Contents

Preface and Acknowledgments

The world as we know it is saturated with language: our minds, our systems of knowledge and belief, our patterns of reasoning, our values, our narratives, our relationships. For us, these things could not develop as they do, in sustained and intensive social interaction, if language did not provide moorings for us to coordinate around. In aiming to understand how this works, my research has taken me to worlds far from the one I grew up in. There has been no better lesson than to be shaken from the idea that my background assumptions—including the kind of language I grew up speaking—are adequate as reference points for all of humanity. Eighty years ago, the linguist and anthropologist B. L. Whorf advocated the study of diverse languages—especially, Indigenous languages—precisely for the purpose of gaining insight into our own cognitive and cultural prejudices and limitations. He asked us to imagine ourselves as a people with "the physiological defect of being able to see only the color blue." Such a people "would hardly be able to formulate the rule that they saw only blue," he wrote. "In order to formulate the rule or norm of seeing only blue, they would need exceptional moments in which they saw other colors." The quest to see other colors has led me to spend years in the wilds of mainland Southeast Asia working with languages and cultures quite unlike my own. This book draws in part on lessons I have learned in the upland rainforests of central Laos and similar lessons that others have learned looking for new colors with speakers of sometimes radically different languages of the world. But I also draw on many other areas and methods of language science—from controlled psychology experiments to media and narrative analysis—without which there would be no hope of reaching a

comprehensive understanding of language, our species' most unique combination of art and instinct.

*

This work has had many inspirations. I mention just three here.

First, in 1997, my PhD adviser, Nick Evans, handed me a book edited by Esther Goody, *Social Intelligence and Interaction*. That book set the coordinates for most of the questions I have asked about language in the quarter century since. Among many gems in the book is a chapter by Ed Hutchins and Brian Hazlehurst, titled "How to Invent a Shared Lexicon: The Emergence of Shared Form-Meaning Mappings in Interaction." In trying to understand language, Hutchins and Hazlehurst wanted to "push the boundaries of a genuinely cognitive unit of analysis out beyond the skin of the individual." They grapple with the question: How can a cognitively shared language arise among members of a community when there is no telepathy? They planted this question in my mind, and the question has guided, in one way or another, most of the research I've done since. This book is another of my efforts to provide an answer.

Second, on an evening in 2003, Steve Levinson hosted a research retreat for the Language and Cognition Department of the Max Planck Institute for Psycholinguistics in the Netherlands. We were deep in discussion about the properties of language, both its strengths and weaknesses: How is it that the distinctions languages make for describing the world are so minimal compared to the many fine distinctions we can perceive with our senses? The average human language has only five or six basic color words, yet the human senses are able to discriminate some 2 million or more distinct colors.[1] Steve remarked: "Language is hopelessly blunt for capturing the details of reality".[2] Tanya Stivers—now Professor of Sociology at UCLA—retorted: "Yes, but it's exactly as sharp as it needs to be for managing social relations." In the years since, I have worked on trying to understand both of those contentions (which are, essentially, the two contentions in the subtitle of this book: language is too blunt for scientists but just right for lawyers), working together with many of the people who were there in the room that evening. In time, I have become convinced that the concept of language as a social coordination device helps to explain why—and how—both Steve and Tanya are right. In the subsequent years that I spent collaborating with them and other incomparable colleagues and visitors—especially Melissa Bowerman,

Mark Dingemanse, Paul Kockelman, Asifa Majid, and Jack Sidnell—I have explored language in ways that shed light on both of these ideas.

Third, in a more distributed way, I have been inspired by more than two decades of interactions, real and virtual, direct and indirect, with an incredible array of researchers and thinkers beyond those just named. These include (but are not limited to!) Paul Bloom, Penny Brown, Eve Clark, Herb Clark, Emma Cohen, Sotaro Kita, Betty Couper-Kuhlen, Daniel Dor, Robin Dunbar, Nick Evans, Dan Everett, Susan Goldin-Meadow, Bill Hanks, John Heritage, Chris Knight, Hugo Mercier, Seán Roberts, JP de Ruiter, Manny Schegloff, Michael Silverstein, Marja-Leena Sorjonen, Dan Sperber, Mike Tomasello, Anna Wierzbicka, and Chip Zuckerman.

For generous comments and suggestions on the manuscript of this book, which have greatly helped improve it, I am grateful to Emma Cohen, Kensy Cooperrider, Daniel Dor, Robin Dunbar, Margie Enfield, Phil Laughlin, Weijian Meng, Hugo Mercier, Susan Perry, Danielle Pillet-Shore, Jack Sidnell, Oliver Traldi, Tom van Laer, Samantha Williams, and three anonymous readers for MIT Press.

For their expert and professional assistance, I thank Gus Wheeler (who prepared the figures), Weijian Meng (who assisted with preparation of the manuscript), Naomie Nguyen (who assisted with proofs), and Angela Terrill at Punctilious Marvelid for compiling the index.

I am sincerely grateful to my agents, Katinka Matson and Max Brockman, for their guidance and support and to my editor at MIT Press, Phil Laughlin.

And finally, with love, I thank the other Ns at Goldie—Na, Nyssa, and Nonnika—for brightening the locked-down world in which this book was written.

*

I dedicate this book to Nyssa and Nonnika. May they be good lawyers and better scientists.

or dispassionate as we would like to think it is, that our inner scientist is in fact an inner lawyer.[5] The scientist seeks to know the truth, while the lawyer seeks to persuade. And in persuading, the lawyer seeks not to get at the truth but to get her way (or to get the way of those who pay her fee). She seeks not to explain but to defend. And notice that while the scientist may sometimes work alone, the lawyer's job is a necessarily social one, and language is her primary tool.

The idea that language is an infrastructure for social coordination and not for the transfer of information per se will help us understand some of its shortcomings, which we shall encounter in the first two parts of this book: why language seems to fail us in the ways it does, why it is so ambiguous and approximate, why it distracts and detracts, why it falls short when we try to describe an experience or capture an innermost feeling. At the same time, the idea that language is a coordination device will help us understand why it can be so good at the things it is good at: directing people's attention, framing situations in arbitrary ways, playing to people's biases, tuning our interactions, managing reputations, and regulating social life.

Arguments for the social function of human intellect go back to pioneering ideas in evolutionary psychology,[6] and language is central to those ideas. In this book, I focus not on the natural history of language's role in socially oriented reasoning but rather on the properties of language that shape its functions today. Language excels at reason giving, storytelling, and sense making: the quintessentially social qualities that characterize our species.

One of the most dangerous properties of language is that it allows us to say things that aren't true. The danger is not just that people may be misled, but that falsehood may be more effective than truth. Truth becomes a collateral victim of human sociality. The strength of human commitment to beliefs in supernatural entities and conspiracy theories—a kind of commitment found in human groups worldwide—draws precisely on the disconnect between a statement and the reality it claims to describe. If a group of people collectively state a belief in something that is likely to be false, then the statement, far from seeding doubt, will work as an honest signal of each individual's commitment to the group.[7] Author Curtis Yarvin explains the attraction of improbable ideas in building social movements. For the purpose of social allegiance, it's actually better if the belief that people coordinate around is patently false: "Nonsense is a more effective organizing tool than the truth. . . . To believe in nonsense is an unforgeable demonstration

of loyalty. It serves as a political uniform. And if you have a uniform, you have an army."[8] This is all very well if your only goal is to secure loyalty in defending a position, but reality will come for you at some point. While real soldiers may pledge allegiance to magical ideas, they are ultimately in the business of physical force, not magic but brute reality par excellence. Once a bullet is flying, neither words nor the beliefs they express can stop it.

This is why you cannot say that there is no reality beyond our ways of talking, or that reality is whatever we say it is. That caricature of postmodernist thought—as if anybody really lived by it—makes no sense in a world in which our species evolved by natural selection, in which we depend on food, air, water, light, and avoidance of injury to live through each new day. Those who claim to doubt objective reality will still defer to that reality in lawful and predictable ways. As the philosopher David Hume quipped, if you are skeptical that a real world exists, then you are welcome to leave via the second-floor window.[9]

But heeding physical reality as a matter of survival is quite a different matter from coordinating around reality for social purposes. When we talk, our words create the *versions of* reality—whether social or physical—that we agree to coordinate around, for example, when we want to affiliate with someone, influence someone, recruit somebody's help, or collectively evaluate a situation and work out what has happened, why, and what action to take. It's only through our publicly *shared versions* of reality that social coordination is possible. And it is always just *one* version of reality that we coordinate around at a time. That version is the one we create with words.

*

Do you control language or does language control you? Does your act of describing an experience alter that experience forever? Are you putty in the hands of a compelling narrative or can you think for yourself? In this book, we find answers to these questions in far-apart fields of the cognitive and social sciences. We delve into research in linguistics, anthropology, cognitive psychology, sociology, and communications to find out what language is good for, what it does for us, and what it does *to* us. We discover some of the many ways in which language giveth and taketh away.

In part I (chapters 1 to 3), we begin with the question of how words relate to physical reality. There is, of course, a reality beyond human perceptions and beliefs. But as evolved creatures, we have access to only a thin slice. Our

perceptual systems deliver a massively stripped-back version of the real world. This is good because it reduces the computational complexity involved in interacting with our surroundings. Simplifying matters—without oversimplifying them—is necessary. It gives us the information we need for survival without being overloaded or paralyzed. Right now you are surrounded by electromagnetic radiation on a spectrum ranging from the shortest gamma rays to the longest radio waves, but you can only perceive a tiny band of that radiation, visible to you as the color spectrum. In turn, human language imposes on this tiny band a second radical simplification. We *name* only a fraction of the perceptible distinctions available in that already-narrow range.

We shall see that the two processes of abstraction—from reality to perception and from perception to language—serve different functions. While perception reduces reality to make it easier to navigate the physical world, word meanings further reduce complexity for navigating the *social* world. If perceptual categories provide individuals with anchors to reality, *linguistic* categories provide shared moorings for social coordination in groups. Perception is private; language is public.[10] The two things provide solutions to very different kinds of problem.

Using language to simplify reality is not only the business of ordinary social sense making; it is also important in scientific progress. When a scientist makes a breakthrough, they yield a sharable narrative about some piece of reality. This narrative reduces the work involved in making predictions about that piece of reality. In turn, it enables people to coordinate around questions of how to intervene in our environment in useful ways. Language is key. We use it to reduce computational complexity in features of our world—real and imagined—when we need to coordinate around those features.[11]

All human languages categorize natural phenomena ranging from plants and trees to colors to smells to the structure and movements of the human body. As we shall see, the languages of the world show a form of *constrained diversity*. All languages capture only a tiny part of reality, and they do so in similar ways—up to a point. Beyond that point, languages can vary widely in what they capture. Constrained diversity means that languages can vary but not without limit. Words for physical objects and events—from plant species names to color terms to kinship terms to words for parts of the body and human emotions—will always show some degree of respect for the structure of physical reality.

In part II (chapters 4 to 7), we explore how language can tamper with our perceptions, memories, and processes of reasoning, in sometimes surprising ways. More than a century of ingenious experimentation in the psychology of language reveals that the meanings of words do not have transparent or straightforward links to the reality they describe nor to the mental images and ideas they denote. Far from it. We shall see that not only do our words simplify and skew our perceptions, they can alter and overwrite them, changing our memories and beliefs about things we've experienced firsthand. Choice of words can prime us, directing our attention toward some things and away from others. And the effects of misdirection are profound. They underpin the incorrigible biases that characterize human decision making, and in turn they can be exploited in influencing people's attention and reasoning. One of our most important weapons against misinformation is to be aware of these biases, in others and in ourselves. Language is *choice architecture* for thinking and social action.[12] We may yield to it quite unknowingly or we may take control of it through mindful attention to its design. And because different languages feature different kinds of meanings and structures—different architectures—this introduces the possibility of *linguistic relativity*: the idea that if a language can nudge our thoughts and perceptions then different languages can nudge us in different ways.

These effects of nudging and skewing our thinking at a microlevel can be viewed as bugs from the point of view of the person who is affected. But they are also features because they reduce the costs involved in processing language, and because they can be exploited in influencing people. Nowhere is this more apparent than in the use of framing, a key tool of persuasion.

In part III (chapters 8 to 11), we scale up, examining the power of framing, stories, and narratives in persuasion, sense making, and social cohesion. We open with two powerful principles in the manipulation of information. One is the tendency to use language strategically, in an artful form of "conjugation" described by the philosopher Bertrand Russell. We might say, *I am firm*, but *He is pig-headed*, choosing markedly different ways to describe the same quality in two people. A corollary of Russell's principle is that when you choose a framing, that choice reveals something about *you*. By a principle of inversion noticed by the philosopher Ludwig Wittgenstein, we can use a person's words as a measure not of the thing this person describes but as a measure of their stance toward that thing, and thereby as a measure of the person. Through the chapters of part III, we explore the power of these

principles in public discourse, first through case studies of strategic wording in media language and then through an examination of the structure and power of storytelling and narrative.

We shall see that storytelling is central to human affairs, not only in the creative arts but in the everyday narratives that pepper our conversations. Narrative informs us, transports us, and engages us through emotional connections that attract and hold our attention and bind us socially with our co-audience members. Stories convey the practical and cultural wisdom of myths and teachings and help us make sense of our social worlds. And more locally, we use storytelling to manage reputations, an important part of our species' particularly complex and intense form of social group organization.

In the book's concluding chapter, we reflect on implications of what we've learned about the things that language does for us and to us. These are implications for human agency and freedom of thought and action in a physical world that demands our respect. In order to coordinate around reality for social and practical purposes, we rely on language for shared moorings. Most importantly, I think, the implications of language's shortcomings concern our obligations to ourselves and to our fellow users of humankind's most powerful and transformative invention.

1 Coordinating around Reality

> Solving a problem simply means representing it so as to make the solution transparent.
>
> —Herbert Simon (1981)

The scene is a tropical dry forest floor in the Lomas de Barbudal Biological Reserve, northwest Costa Rica.[1] Two wild white-faced capuchins—a small-ish species of monkey around a foot and a half tall and weighing about the same as a newborn baby—scuttle along looking for food. They stop and together turn their attention to something nearby. One hops on the other's back, and they form the overlord posture, their little heads stacked one on top of the other like a totem pole.[2] As one, they scream at their common focus of attention.[3] This coordinated action is how capuchins signal and build coalitions in their societies. Once coalition-mates have built a bond, they may later protect each other or call on each other for help when needed. But who—or what—are these two capuchins threatening?

It's a patch of dirt. As it happens, capuchin monkeys engage more often in pseudo-aggressive incidents than in real encounters. When they get together to display aggression toward something harmless like dirt or an eggshell, this shows us that the function of coordination is not just about defeating a threat but also about building a social structure.[4] When individuals build coalitions through coordination, they become merged as agents. They work as one. To achieve this, they need landmarks to coordinate around. Even a patch of dirt will do.

This is one sense of *coordination* that we will encounter in this book: adopting a common stance toward a landmark, a shared focus of attention. Many species engage in this, but we humans have remarkable capacities to

use such landmarks as cognitive tools for achieving coordination even in the absence of communication, in a kind of virtual mind reading.

Imagine you are a parachutist coming down in unknown territory. You are with another parachutist who is somewhere in the same area, but neither of you knows where the other has landed. You have no way of communicating. But you need to find each other, and fast, if you are going to be rescued. Your one chance is the map shown in figure 1.1.[5]

You are both in possession of the map, and you each know that the other has it. Can the two of you figure out how to coordinate your movements and meet in the same place? 'Does the map suggest some particular meeting place so unambiguously that each will be confident that the other reads the same suggestion with confidence?'[6] Where would *you* go to meet the other parachutist?

The scenario and map are from a 1960 study by economist Thomas Schelling. In Schelling's experiment, most people said they would meet at

Figure 1.1
Schelling's map: Two parachutists, at points X and Y, must coordinate and find each other by going to where they think the other will go, each knowing that the other has the same map. After Thomas C. Schelling, *The Strategy of Conflict* (Cambridge, MA: Harvard University Press, 1960), 55.

the bridge.[7] It doesn't seem like rocket science, but it's actually remarkable that we can pull this off so reliably. We appear to be the only species that can achieve entirely ad hoc coordination of this kind *without communication*.[8] This demands a certain kind of "mind reading" or, more accurately, not reading but imagining or speculating.[9] We take into account what the other person knows, what we believe that person knows about what we know, and how that mutual knowledge should factor into our current shared task of trying to coordinate our behavior. The things we say provide landmarks of this kind, and we actively imagine where our minds are supposed to meet. If you hear someone say, *It was too slow*, this vague proposition could mean a million things, but if you know the context, you will readily use the words as a map to arrive at a much more specific understanding. We have a subjective sense that everything comes through in the words, that the words themselves carry all the meaning. But in truth, our words are as sketchy as Schelling's hand-drawn map, leaving out almost every feature of the landscape, and requiring us to imagine where in the world we are supposed to come together.

This is a second sense of *coordination* that we will encounter in this book: when we use language in interaction, we draw on landmarks to infer common solutions to the problem of converging in thought and action. The reason we are forced to rely on inference is that we have no recourse to telepathy. When we coordinate around the maps that language draws for us, we naturally fill in the details by imagining what must be there. Language is a portable device for constructing such landmarks at will.

As a coordination device, language is a tool of human *agency*, both individual and shared. Agency can be defined as our capacity to do things and make things happen, where this capacity entails some degree of control and some degree of accountability. Words are our most powerful tools. Students of the Stanislavski acting method are taught to regard their every spoken line—along with their every move on stage—as pursuing a task, solving a problem, carrying out an action. At every point, an actor must ask: What am I trying to make the other person *do*? With words, we act both *on* other people and *with* other people, for agency is situated in the individual *and* is distributed across individuals.

We find it natural to think of ourselves as individuals when it comes to our personal agency. By coordinating socially, separate individuals may mutually influence one another, but in reality, our agency is distributed.

At every step, we enhance our individual agency using person-extending technologies from levers and wedges to bicycles and smartphones. These technologies are *interfaces* that translate or transform our actions beyond what our bodies alone could do.

An interface for coordination requires mutual tuning between those who would coordinate. Here is an example of such tuning in the natural world. In the jungles of Kibale National Park in southwestern Uganda, primates and birds play a special role in the ecology. These day-active creatures are the primary seed dispersers for rainforest plants, eating ripe fruits and spreading the seeds in droppings. Three thousand kilometers away, an entirely different ecology is found in the rainforests of Ranomafana National Park, across the Mozambique Channel in southeastern Madagascar. There, the main seed dispersers are lemurs. They are different from Kibale's primates and birds in a specific way: they are active at night and are red-green color-blind. A team of biologists measured the color contrast between ripe fruits and leaf backgrounds in the two parks and found a difference: "Fruits in Uganda have higher contrast against leaf background in the red–green and luminance channels whereas fruits in Madagascar contrast more in the yellow–blue channel."[10] What explains this? The answer is evolution by natural selection. If an individual plant's fruit happens to be more visible to the animals that would eat it and thereby disperse its seed, then that plant is more likely to have its seeds dispersed. As long as differences in fruit color are genetically inherited by each generation of plants, the fruits in time will become more clearly visible to the relevant seed-dispersing animals in each area.

The Madagascar fruits are *tuned to* the local lemurs' visual systems. The fruits show higher color contrast in the yellow-blue channel because that's what makes them more easily noticed by the red-green color-blind lemurs. The fruit's color is an evolved interface between the seed-bearing edible fruits and the visual systems of the lemurs. By producing fruits of a certain color, these forest plants use light waves to go beyond their skins, influencing another species at a distance, in this case for mutual benefit. With the fruit-color interface, these two life forms—the plants and the lemurs—can coordinate.

We humans show a similar form of interdependence for coordination. Among the resources we mobilize to extend our agency are other people. Language is our interface with them and our tool for mobilizing them. Language is the most flexible, powerful, and all-pervasive agency-extending

say what you like about these statements but your beliefs or opinions will have no effect on whether they are true. This is not to say that our grip on reality isn't subjective. Of course, how you access reality depends on what kind of body you have.[17] Thanks to the nature of the human auditory system, the average person can't hear sounds at a frequency of 30 kHz. But this doesn't make those sounds any less real. A dog can hear them just fine. We might say that they *aren't real for humans*. They don't affect us, and so they can't interest us or influence our decisions, at least not in everyday life. It's the things we can detect that have captured our attention and interest to our advantage through our species' evolutionary history: voices, hailstorms, rocks, bears, breakages, heat, earth, wind, fire, water, chairs, guns, parties, and especially, other people.

Let's see just how narrow a slice of reality we normally have access to. We are each of us surrounded by electromagnetic radiation, with wavelengths that vary by many orders along a continuous spectrum. Through our evolved visual systems, our bodies deliver only a small slice of that reality, the slice that we call visible light. The rest is literally invisible to us (figure 1.2).

Perception is a simplified user interface between us and the world.[18] Our perceptions "have not been shaped to make it easy to know the true structure of the world but instead to *hide its complexity*."[19] They "have been

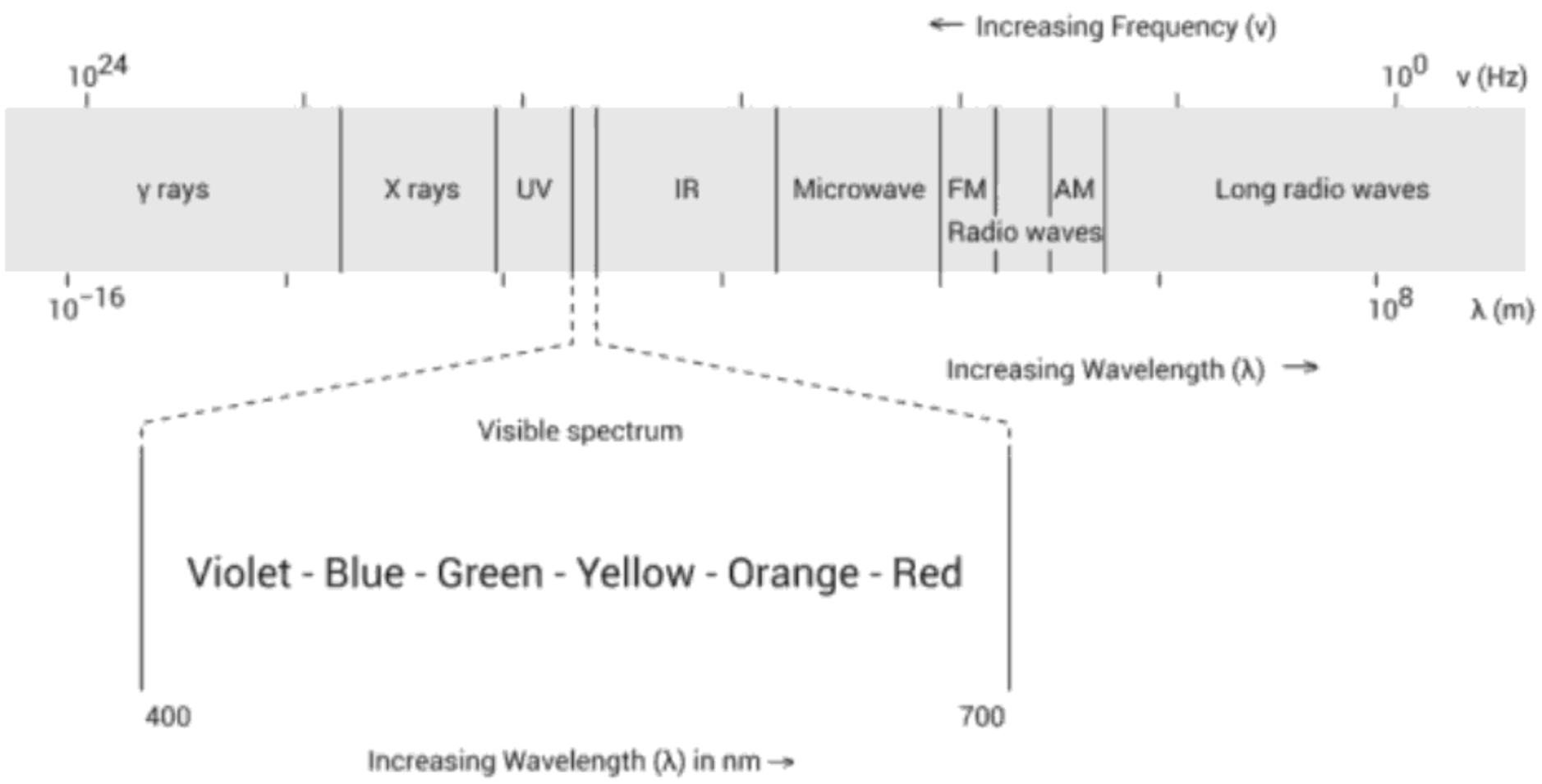

Figure 1.2
The spectrum of electromagnetic radiation, highlighting the part that is directly perceivable by humans. Image adapted from https://upload.wikimedia.org/wikipedia/commons/3/30/EM_spectrumrevised.png.

shaped by natural selection to make it easier for us to act effectively in the world."[20] This means that it is more important for our perceptions to be useful than to be true. Of course, it is more useful to know the truth than to believe a lie. But it's also often more useful not to know the *whole* truth. We are better off knowing only know what we *need* to know.

Think back to your car's user interface. You don't engage directly with the engine, the hydraulic system, the electrical system. You engage with much simpler mediating structures: the gas pedal, the brake pedal, the steering wheel, the indicator lever. The user interface is many orders simpler than the true workings of the car. You don't need to know how the movements of your feet and hands cause the car to slow down or speed up, reverse or turn. In fact, you are better off not knowing those things, at least for the purpose of driving around. The simpler the interface, the fewer demands it makes on us, freeing us up for other things. It reduces cognitive cost. Our senses of sight, smell, hearing, taste, and touch—designed by the merciless processes of evolution—follow this principle.

With vision, do we see what is actually there? "The frog does not detect flies, it detects small, moving, black spots of about the right size."[21] Vision scientist David Marr explains why evolution might result in perceptions that do not present reality as it is. The perception need only provide an organism with "sufficient information for it to survive":[22] "Natural selection is a search procedure that yields *satisficing* solutions, not optimal solutions."[23] Satisficing can be defined as motivated stopping:[24] When you've arrived at a solution that's *good enough*, you cut your losses and move on to the next thing.

If our systems of perception provide interfaces that simplify reality in order that our decision making may be more efficient, this would explain why our *perception* delivers such sparse mapping of our world, but it wouldn't explain the *linguistic* facts. While our perceptions capture only a small set of possible distinctions in brute reality, words make *many orders fewer distinctions* again. The visible spectrum of light is a tiny slice of the full spectrum of electromagnetic radiation, though even within that spectrum, more than 2 million distinctions in color can be discerned.[25] In turn, our vocabularies for color reduce that number of distinctions by many orders, usually coming down to only a handful of words for basic color categories (*red, green, blue*) or at best a few dozen or so technical words for color distinctions in a language (see figure 1.3).

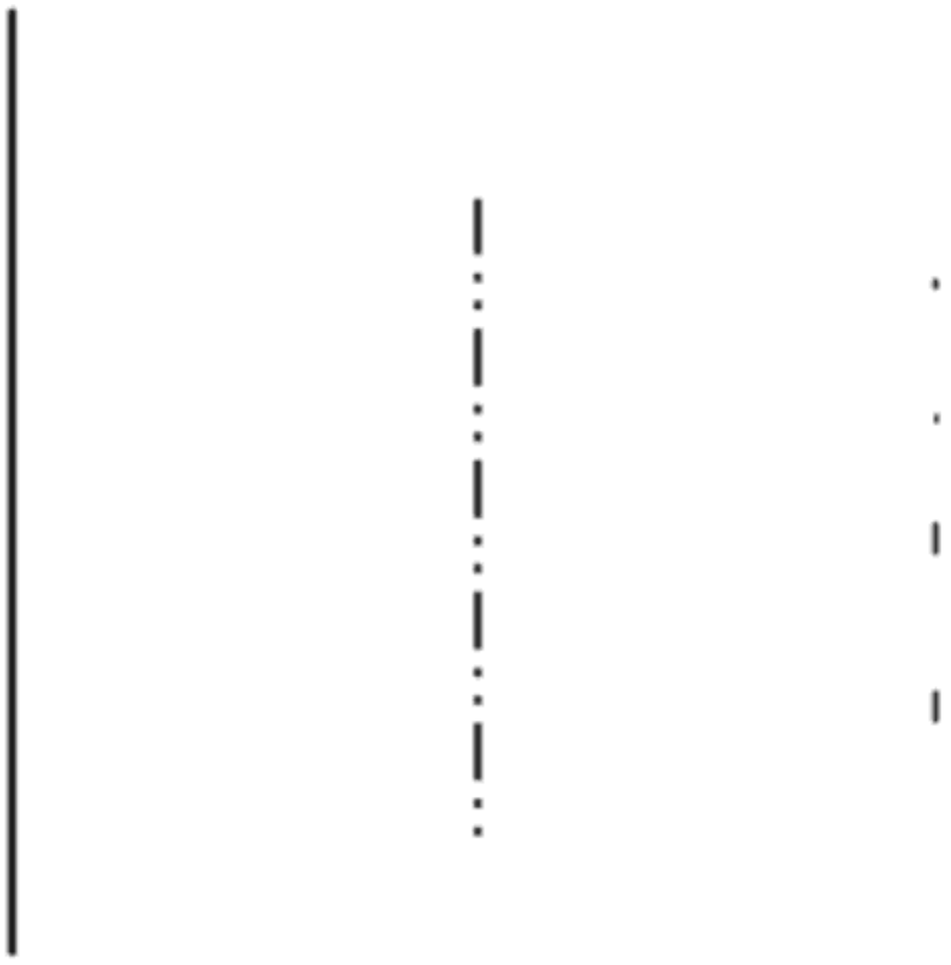

Figure 1.3
Two layers of simplification: reality is reduced by perception, perception is reduced by language.

Why have two interfaces that successively simplify and hide reality? Because the two interfaces serve two different purposes.

Perception is an interface between reality and an individual. It delivers a version of reality that is good enough for individual decisions and action—good enough for *individual agency*. Perception simplifies reality in a way that reduces costs for individuals, without reducing these costs so much as to cause an intolerable level of error in our interactions with the real world.

In turn, and by contrast, language is an *interface between individuals*. It delivers a rendition of reality that is good enough for coordinating *joint* decisions and actions. Good enough for *distributed* agency. Language simplifies our perceptions and concepts in a way that reduces costs, yet without reducing those costs so much as to prevent us from coordinating successfully in our interactions with other people. A system for social coordination can serve its function with information that is much more abstract and partial than what is needed for individual decision making and action. The two interfaces—perception as an interface with the world and language as an interface between people—successively simplify reality.

It's difficult to overstate how poor language is at capturing brute reality, or at transferring the richness of people's perception or experience. Suppose I say to you, *Jo drives a red car*. Now, what if I ask you to do a painting of the car as you see it in your mind? Which exact shade of red do you choose?

What sort of car is it? What if I asked you to pick out Jo's car in a parking lot? Could you do it? Not likely if there were ten red cars of slightly different shades. Or suppose I describe Jo's face. Would you be able to pick her out from a lineup of ten similar-looking people, based solely on my verbal description? Even assuming you are a skilled portrait artist, could you draw her face accurately from my words alone? Language provides nowhere near the required level of information for the task. And not only that, as we'll see, language distracts us in various ways. It can divert our attention and mess with our memories.

Now let's turn to the relation between language and the second type of reality: social reality. A true statement about brute reality—for example, *These potatoes will fall if I drop them*—will remain true no matter what people believe. By contrast, a statement about *social* reality—for example, *These potatoes belong to me*—is a statement about my rights and duties. The statement is entirely dependent on people's beliefs. It is true *because of* an agreement among people to treat it as true, whether that agreement is tacit or explicit. I can sell the potatoes or give them away. These are both ways of transferring the rights and duties I have over them, and thus of rendering untrue the statement that "they belong to me".

The philosopher Elizabeth Anscombe introduced this distinction between brute and social reality by comparing two ways of describing a single situation.[26] She imagines that someone delivers a shipment of potatoes to her house. Two descriptions are: (1) *He had potatoes carted to my house and left there* and (2) *The potatoes came into my possession*. The first sentence is a physical description of a person's movements in relations to some objects and places. It describes brute reality—things that can be physically observed. The second sentence describes a different kind of reality. You would not be able to tell by physically examining the potatoes who they belonged to. The potatoes are edible when cooked. They can block someone's path. These are matters of brute reality. But when we say that somebody has the right to cook and eat them or the duty to move them out of the way, we are talking about social reality. My right to eat them or my duty to remove them is not caused by brute facts but by social, interpersonal agreements about ownership—agreements that are created by, or at least backed up by, language.

The philosopher Wilfred Sellars presented a similar contrast using the example of chess.[27] One situation, two descriptions: (1) *There is an open diagonal*

space between this white piece of wood and that red piece of wood, (2) *My bishop is checking his king*. The first describes a physical situation and could be roughly rendered in any language and understood and evaluated by anybody in the world on the basis of the words in the one-sentence description (assuming we know what the words *diagonal, wood, white, red,* mean). But the second requires knowledge of the regulations of chess and of the social rights and duties that chess entails. Not only that, but the facts at hand—what a bishop can do, how check is defined—are true only because people agree to treat them as being true. People can stop Sellars's second statement from being true if they simply change the rules of the game—for example, by deciding to agree that bishops cannot move diagonally. But no form of agreement would stop the first one from being true (unless, of course, the meanings of the words in the statement are changed in meaning; we will encounter examples of this later in the book). Whatever the game's rules, the white piece of wood will stand in the same spatial relationship to the red one.

The relation between the two realities is not symmetrical. Legalities are a classic example of social reality. How are laws upheld? By the threat of physical force. Brute reality. This is why political power—mostly a matter of institutionalized rights and duties—is ultimately grounded in physical facts. A baton to the head will hurt, no matter who you vote for. The political sociologist Max Weber defined the state as an entity with a monopoly on the legitimate use of physical force.[28] Whether force is legitimate can be contested (as social reality), but the physical effects of the force cannot. The effects of force are in the realm of brute reality. They are not affected by our interpretations or our perspectives, and this is what makes physical facts the ultimate arbiter of power in human affairs. In this sense, brute reality is the bedrock of all of our social institutions.

It is sometimes argued that brute reality is trumped by social power, that "facts" are illusory, as if anyone with power can decide what is true and what is not. But social power can only change social facts. Ownership, for example, can readily be negotiated—rendered untrue, if you like—by the use of force and its nonnegotiable effects. If someone with a knife deprives me of my rights to the cash in my wallet, this works because of the physical fact that the knife would cause me harm. A mugger's power to overturn the social fact of ownership comes directly from physical facts. I can contest it by producing a bigger knife. Or if I threaten the mugger with prosecution,

and the text on your phone screen appears in identical form on mine). Language is unlike either of these things because it is neither direct physical manipulation[41] nor direct reading or replication of a mental state. Any plausible explanation of how language works has to appeal to nature and not magic. Hence, a *no-telepathy assumption*: "No mind can influence another except via mediating structure."[42] In the case of language, that mediating structure is provided by words, gestures, and the rules of grammar. It is a shared structure for calibration and coordination of thought and action. Individuals must of course still individually process and represent meanings in language.[43] Ultimately, when we have learned a language, our minds have become "organized by structure created by other individuals."[44] We are all mutually implicated in the collective significance of our languages.

2 Schelling's Game

> Every living creature is in fact a sort of lock, whose wards and springs presuppose special forms of key.
>
> —William James (1884)

Imagine you are participating in an experiment. The experimenter puts you in a room on your own and tells you that you have a partner in the experiment, someone you've never met, who is sitting in a room nearby, also on his own. You have no way to communicate with that person. Here is your instruction:

> You are to divide $100 into two piles, labeled A and B. Your partner is to divide another $100 into two piles labeled A and B. If you allot the same amounts to A and B, respectively, that your partner does, each of you gets $100; if your counts differ from his, neither of you gets anything.

Can you guess what the most popular solution is? That's right. Ninety percent of people in Schelling's experiment converged on the obvious: fifty-fifty.

Schelling's coordination problems are famous for showing that strangers are able to converge in brand-new situations with no discussion or conferral. But the situation that Schelling put his subjects in is just one special case of coordination problem.[1] The philosopher Margaret Gilbert explains: "Where it is not possible for the parties to make an agreement about who will do what, or otherwise to communicate their intentions, and there is no special background knowledge, the parties have a genuine problem: Should each make a random choice of action, the chances of achieving the desired coordination of actions are not good."[2] But of course, most of the time we do have the opportunity to make an agreement, and we do share relevant background knowledge. Most of the time we are able to follow convention. Take the case

of traffic rules. The most important traffic rule of all tells us which side of the road to drive on. As the philosopher David Lewis wrote, "It matters little to anyone whether he drives in the left or the right lane, provided the others do likewise."[3] We don't leave this to chance.[4] It is a convention, codified in law.

Whether we rely on conventions, laws, or clever guessing, coordination has two crucial elements. First, whether one person's action is the correct one will depend on what the other person does. In the case of the parachutists, your choice to go to the bridge will be the right decision only if the other parachutist also goes to the bridge. Second, the two people involved have aligned incentives: they will both prefer it if they succeed in meeting in the same place on the map, regardless of what place that is.[5]

Schelling's findings led him to a *theory of interdependent decision*:[6]

> People *can* often concert their intentions or expectations with others if each knows that the other is trying to do the same. Most situations—perhaps every situation for people who are practiced at this kind of game—provide some clue for coordinating behavior, some focal point for each person's expectation of what the other expects him to expect to be expected to do. Finding the key, or rather finding *a* key—any key that is mutually recognised as the key becomes *the* key—may depend on imagination more than logic, it may depend on precedent, accidental arrangement, symmetry, aesthetic or geometric arrangement, casuistic reasoning, and who the parties are and what they know about each other.

There are many more examples of these kinds of coordination problems.[7] Here are some:

- Call "heads" or "tails." If you and your partner call the same, you both win a prize.
- Circle one of the numbers listed in the line below. You win if you all succeed in circling the same number.

  7 100 13 261 99 555

- Name some amount of money. If you all name the same amount, you can have as much as you named.
- Suppose you and I are talking on the telephone and we are unexpectedly cut off after three minutes. We both want the connection restored immediately, which it will be if and only if one of us calls back while the other waits. It matters little to either of us whether he is the one to call back or the one to wait. We must each choose whether to call back, each according to his expectation of the other's choice, in order to call back if and only if the other waits.

We are so practiced at solving coordination problems like these that we hardly notice we are solving them at almost every turn.

Solutions to Schelling's games have something in common: the thing that people coordinate around will always have "some kind of prominence or conspicuousness."[8] That conspicuousness should be apparent to both parties. Recall that there is only one bridge on the map in figure 1.1. It is near the center of the image, and most roads lead to it. People readily see it as the best thing to coordinate around. This is not just because people find the bridge prominent. It is because they understand that anybody else would also recognize this prominence.

When openly or mutually prominent solutions to coordination problems are available, then people can not only increase the likelihood of successful coordination; they also reduce the costs. Schelling's coordination games are remarkable because people solve them without needing any kind of communication at all. These sorts of solution relieve us of the need for explicit bargaining or negotiation so that we may get on with whatever is next in the stack of things to deal with.

You might wonder how often in everyday life you are faced with a coordination game of this kind. The answer is: all the time. Every time you talk to somebody. Language is our most important tool for achieving social coordination, and using language is itself a coordination game.[9] Becoming highly practiced at using language means recognizing the prominent features that word meanings provide, just as the parachutists' map provides prominent landmarks for anyone who looks. And with language, as with the parachutists, the map is shared and known to be shared. When it comes to using language, we are all, as Schelling put it, highly practiced at this kind of game.

This point underpins one of the central ideas I want to convey in this book. The idea is that words and other bits of language are none other than *highly practiced solutions to coordination games*.

Take a simple word like *spoon*. Imagine the sentences I might use it in. *Can you pass me a spoon? This spoon's dirty. You need a fork and spoon to eat spaghetti*. If the function of this word were to capture a piece of reality, it would be a rather blunt tool for that. The examples of *spoon* that you just read gave you only the most schematic information about the objects I might have been referring to. Was I referring to metal or plastic spoons? Stainless steel or silver? What size? What color? Damaged or in mint condition? The word *spoon* alone gives you none of this, although I might have all those details

in mind. The word's function cannot be to convey the details of my mental image—or at least it would be inadequate for that function. But if the function of the word were to provide a useful landmark for us to coordinate around, it would be just good enough.[10] If you're fetching a spoon for me, you scan the possibilities for something that will fit the description and provide the solution to the coordination problem at hand. The word *spoon* abstracts from the differences between individual spoons and is a historically practiced solution to the frequently occurring need for coordinating our behavior around the things we agree to call spoons. This is what psychologist Roger Brown meant when he said that the meaning of the English word *spoon* is grounded in "the community-wide practice of treating spoons as equivalent but different from knives and forks." The word is a coordination device.[11]

What information does a word like *spoon* need to encode? It needs to hit a happy medium between being not so vague as to result in a failure to coordinate but no more specific than that. A basic level of meaning is "just right." Just right for what? For achieving coordination with others in social interaction. Most of our words are established coordination devices, and when we use them, we receive immediate feedback on how effective they have been in context. That feedback is not incorporated into purely referential views of meaning because they focus on the links joining word, mind, and thing, and they bracket out the actual functional context for words as public entities.

For frequently recurring coordination problems, words like *spoon* are so easy to access and use that we often feel as if the word simply encodes the information we have in mind. We don't notice how imprecise the word is because it is so well tuned to the situations we use it in. Those situations occur over and over. But there are still plenty of situations that aren't just repeats of what we've seen before. You know the feeling of wanting to say something and having trouble finding the right word. That's when the problem of coordinating through language comes to the surface.

Suppose you are on the phone with a friend and you want to mention someone you both know. You need to agree on who you're talking about. As it happens, we have well-practiced solutions to this coordination problem. Personal names serve the function.[12] But what happens when you can't remember a person's name? Here is an example from a recording of a phone conversation:[13]

In this seemingly simple situation, one person's choice of words results in another person's successful actions. One way to interpret what is going on is that the speaker is helping the listener. He is giving the listener what she needs in order to solve the problem of picking the right item. Another way to think about it is that the speaker is influencing or manipulating the listener, using words to direct or even control the listener's thinking and behavior. The difference between these two views depends on whose goals we are focusing on. If we are concerned with the matcher's goal of picking the right item, then it's natural to think of the director as helping that person. If we see the selection of the correct item as essentially the director's goal, then the matcher is the means toward that end.

But surely both of these things are true. It's not my goal versus your goal. The people in this interaction share a goal. As in any other coordination game, they have aligned incentives, so the director is helping the matcher precisely *by* influencing her. This is how language works.

A corollary of the minimalist language that emerges when coordination problems are solved is our tendency to use simplified descriptions when possible. Imagine you are in an experiment and you have been asked to read this paragraph:[17]

> Rachel parked her car outside the supermarket. She got out of her car, collected a grocery cart and wheeled it inside. She checked her list and went down the aisles. She put the items that were on her list into her grocery cart until she had them all. Then Rachel went to the checkout where she joined the fastest line. She waited in the line, and then unloaded her items onto the belt. The cashier rang up the items on the till and told Rachel the total. Rachel gave the cashier some money and the cashier gave Rachel her change. Rachel put the shopping into the bags and put the bags into the grocery cart. She wheeled the grocery cart out to her car and put the bags into the boot before driving away.

The experimenter asks you to write down what you can remember from what you've just read. Here's what you write:

> Rachel went shopping, parked her car at the supermarket, got out of the car, got a grocery cart, went into the supermarket and collected the food she wanted. She went to pay for the goods, gave the cashier the money, he gave her change and a receipt. Then she took the grocery cart back and then drove off in her car.

Now your paragraph is given to the next person in the experiment, who in turn is asked to write what she can remember from it. This is then passed

on to another person, and another person. Here is how the experiment plays out:

Round 2. Rachel drove to a supermarket, parked her car, got a grocery cart, and chose some food. Then she went to the cashier to pay for her food. The cashier gave her some change. Then she put back the grocery cart and drove away.

Round 3. Rachel drove to the supermarket, parked her car, got a grocery cart, and chose some food. She paid the cashier and drove home.

Round 4. Rachel went to the supermarket, got some food, and went home.

Each round has the same ever-simplifying language we saw in the ice skater experiment. But there is an important difference. In this supermarket story task, each contribution is an act of recall, not an instruction to a matcher who has to pick from alternatives. In the tangram study, the phrase that resulted from a series of rounds—*the ice skater*—emerged because it was the phrase that worked for solving the coordination problem at hand. All of the unnecessary stuff was whittled away over the course of each round. Now, note that the first version of the supermarket story features twenty-one distinguishable events. In the end, almost all of them are captured with a single phrase: *she went to the supermarket*. It's no accident that this phrase ended up as the description of the original complex description. Why? Because the phrase *go to the supermarket* is a landmark on the Schelling map of cultural knowledge that is already shared by everybody in a particular slice of human culture and society. Unlike the ice skater case, the solution wasn't worked out on the spot. This particular solution has been worked out over generations of community language use: if you say that someone went to the supermarket, that conveys all of the usual stuff that one does—you park your car, get out of your car, collect a grocery cart, wheel it inside, and so on—as laid out blow-by-blow in the first version in this experiment.

The truth is that people would hardly ever say anything like the first version of the supermarket story. The reason we have phrases like *go to the supermarket* is precisely to save us that trouble. We so often need to coordinate around the idea of going to the supermarket and all it entails—for example, when I say, *I'm going to the supermarket. Do you want me to pick anything up?*—that we have converged on a simple and brief convention: the phrase *going to the supermarket*. The details of that complex chain of events—for

example, whether I used a grocery cart or a hand basket, whether I pay by cash or card, whether I put the bags in the trunk or on the passenger seat—seldom matter to the basic function of the phrase that describes it, so the convention is not to mention them at all. And in fact, when we do mention them, this departure from the normal economy of expression signals to our addressee that we are adding this detail for a reason.

Consider this line in the first version of the story: *She joined the fastest line*. Of course she joined the fastest line. If I was talking about a trip to the supermarket, why would I even mention this? It would have to be specifically relevant to what I'm telling you. From what we know of how people actually talk, the most likely reason to mention this detail would be as a prelude to some kind of disruption to normal expectations. (We get into this in chapters 9 to 11 on storytelling and sense making.) Maybe the line looked as if it would be the fastest one, but then there was some delay and it turned out I had to wait longer than I would have at any other line. That could be the basis of the kind of everyday anecdote we tell each other all the time. But if I'm not launching a narrative about how something went differently from normal, I should stick to the simple, default way of referring to the situation and leave out the details. That's why in the tangram experiment, you can be sure that if they kept playing, players would stick to the phrase *ice skater*, never again bothering to mention that the skater happens to be sticking two arms out in front. And it's why (at least in English) when we refer to people in conversation, we tend to stick to personal names—like *Gina*—rather than more roundabout ways of referring.[18]

When we optimize coordination through language in these ways, we are following principles of *recipient design*[19] and *mutual responsibility*. As Herbert Clark and Deanna Wilkes-Gibbs, the authors of the tangram study, put it, "The participants in a conversation try to establish, roughly by the initiation of each new contribution, the mutual belief that the listeners have understood what the speaker meant in the last utterance to a criterion sufficient for current purposes."[20] What those "current purposes" are will vary depending on the community of people who are conversing. When certain kinds of coordination problem recur in a community, then rich, culture-specific concepts can emerge, like *going to the supermarket*. These concepts will "accurately portray crucial relations in the thinker's physical world."[21] This explains why specialist terminology exists in various subcultures. Consider

this example of 1980s' surfer terminology from psychologists Earl Hunt and Mahzarin Banaji:

> Surfers speak of waves as being "hollow" or "walled." A hollow wave is one that breaks sequentially along its crest, so that the wave break may roll roughly parallel to the beach for perhaps a mile. A good surfer will ride a hollow wave just in front of the break, moving almost perpendicular to the wave's path towards the beach. By contrast, a walled wave has a nearly vertical rise, and breaks simultaneously at all points. A wall can only be ridden directly towards the beach. These concepts have functional distinctions. Surfers can perform acrobatics on their boards while riding hollow waves, so beaches with hollow waves are considered more desirable for surfing. The ability to manipulate hollow waves, however, depends upon the design of one's surfboard. In the 1950's, before surfing technology developed, surfers did not speak of hollow and walled waves, for all waves were ridden directly toward the beach. The surfer example is an example of a situation in which a single referent can be used to describe a whole sequence of events. A surfer's statement "I rode hollow waves all day" implies a whole style of surfing in addition to specifying a wave form. The concept has obvious predictive utility; saying the waves are hollow informs the surfer of the sort of day, type, and probably intensity of surfing. Indeed, one of the benefits of having a single word for a schema is that two surfers can, briefly and succinctly, explain to each other why they are not going to work or class: "It's hollow."[22]

Hunt and Banaji argue that concepts "are culturally satisfactory if they succeed in explaining and predicting the problems that a culture faces."

The key insight is not about how individual people think or behave. It is about how groups of people coordinate in their thinking and behavior. This point pushes back against an individual-centered view of language that has dominated cognitive science for decades.[23]

In chapter 4, we discuss the idea of basic-level categories in language, a degree of specificity in referring to things that captures a sweet spot between too general and too specific. As with maps, there is a Goldilocks principle in getting the resolution just right: providing not too much information, and not too little, for the purposes at hand. People categorize things at the basic level (for example, "He was bitten by a dog," not "by a mammal" or "by a pug") in order to optimize the benefits of information, while minimizing the cost of gathering and holding on to that information. Now note that this argument is about categories we *think with* rather than categories we *communicate with*. Cognitive scientists have emphasized that this trade-off is to "the organism's advantage," focusing on the individual.[24] But words are not for individuals. They exist for achieving convergence among two

or more people. If we are talking about word meanings—and not just any concepts—then our unit of analysis needs to be the dyad, the two people whose aligned incentives are met by the unique solution that a basic-level category provides. Our examples of *the ice skater*, *going to the supermarket*, *hollow wave*, and *dog* aren't just ideas inside our minds. They are publicly shared. They exist only because they have successfully functioned in the history of certain communities as landmarks for social coordination.[25]

*

If our words capture concepts for explaining and predicting things, they do this in order to share explanations and to *coordinate* around shared predictions. As we'll see, words capture only a tiny slice of all the things that one might conceivably perceive or think. Language is not good for capturing and transmitting the details of reality. What it's good for is providing landmarks we can coordinate around. Unlike with any other communication system, with language we can tell people about things that they have not experienced themselves. We can use language to update people on unseen realities. But this does not mean that our experiences are literally transferred to others simply to be downloaded and viewed or experienced. We are now going to examine the limitations of language in its function of describing reality. Not only are there oak trees, fir trees, and spruce trees; we have words for them. But the link between words and the world is not direct. Our minds must provide that link.

gets me to the same place using the concept of minute and the common counting system. Same reference, different senses.

The contrast between *twelve* and *a dozen* is internal to English. There are also sense-to-reference contrasts *across* languages. Suppose I am telling you the time, and it is 1:20 p.m. In English, both *one twenty* and *twenty past one* would get you there. Dutch speakers have another option: they can use a mode of presentation that works just as well but makes no sense in English. The phrase *tien voor half twee* literally means "ten before half (of) two."[1]

The sense of linguistic expressions is stripped back and partial in comparison to the richness of either the references or the ideas that those senses correspond to.[2] As the psychologist Dan Slobin notes:

> Language evokes ideas: it does not represent them. Linguistic expression is thus not a straightforward map of consciousness or thought. It is a highly selective and conventionally schematic map.[3]

So, the mode of presentation—the word's sense, or content—gives instructions for reaching the idea or thing being described. Words with different senses but ostensibly the same reference—think *freedom fighter* versus *terrorist*, *half empty* versus *half full*—is the essence of framing, a powerful principle in language's capacity as a tool for influence.

Language links to the world in highly flexible ways. How this works has been the most studied question in the psychology of language for at least the past 175 years.[4] Researchers have asked: How do our words relate to things in the world? How do we identify objects from their names? What is the link between our knowledge of objects and their linguistic labels? They have found that words and other bits of language touch on virtually every element of our mental and social lives. As cognitive scientists George Miller and Philip Johnson-Laird put it in their landmark 1976 book, *Language and Perception*:

> The meaning of a word can tell you what is, and what is not, an entity that can be labeled with that word. . . . [It] can tell you the function or purpose of the entity that the word labels. [It] can lead you to all you know about an entity. It has access to encyclopaedic information in long-term memory. [It] can tell you about relations between what the word labels and what other words label. [It] can tell you about what other sorts of words can occur with it in sentences. It can place syntactic and semantic constraints on other words.[5]