Michael D. Gordin is Rosengarten Professor of Modern and Contemporary History at Princeton University, where he specialises in the history of modern science. He was awarded a National Endowment for the Humanities Fellowship and is a Guggenheim Fellow. His love of language cannot be matched: a veritable polyglot, he knows Russian, German, French, English, Ido, Esperanto and Latin among others. He has published on the history of science, Russian history and the history of nuclear weapons.

“Erudite and engaging” Nature

“It is hard to imagine that anyone with even the slightest interest in science or language will fail to find this book a treat” Prospect

“The many aspects of Western history touched on by Scientific Babel make it much more than a journey towards an international language for science ... a worthwhile and enjoyable read” Morning Star

“Fascinating ... a thoroughly enjoyable read; written with authority and enthusiasm, it vividly communicates the author’s great love of languages” Times Higher Education
SCIENTIFIC BABEL
THE LANGUAGE OF SCIENCE FROM THE FALL OF LATIN TO THE RISE OF ENGLISH

MICHAEL D. GORDIN

PROFILE BOOKS
CONTENTS

Introduction: Talking Science  1
Chapter 1: The Perfect Past That Almost Was  23
Chapter 2: The Table and the Word  51
Chapter 3: Hydrogen Oxygenovich  79
Chapter 4: Speaking Utopian  105
Chapter 5: The Wizards of Ido  131
Chapter 6: The Linguistic Shadow of the Great War  159
Chapter 7: Unspeakable  187
Chapter 8: The Dostoevsky Machine  213
Chapter 9: All the Russian That's Fit to Print  241
Chapter 10: The Fe Curtain  267
Chapter 11: Anglophonia  293
Conclusion: Babel Beyond  317

Acknowledgments  327  List of Archives  331
Notes  333  Index  403
Talking Science

Les savants des autres nations à qui nous avons donné l’exemple, ont cru avec raison qu’il écriraient encore mieux dans leur langue que dans la nôtre. L’Angleterre nous a donc imité; l’Allemagne, où le latin semblait s’être réfugié, commence insensiblement à en perdre l’usage: je ne doute pas qu’elle ne soit bientôt suivie par les Suédois, les Danois et les Russes. Ainsi, avant la fin du XVIIIe siècle, un philosophe qui voudra s’instruire à fond des découvertes de ses prédécesseurs, sera contraint de charger sa mémoire de sept à huit langues différentes; et après avoir consommé à les apprendre le temps le plus précieux de sa vie, il mourra avant de commencer à s’instruire.*

JEAN LE ROND D’ALEMBERT

You are able to read this sentence. That is obvious, but it is also quite an achievement. You read English; you may or may not speak it. Somewhere along the way, you learned the language, either relatively painlessly as a child or with significant exertion later (how significant depends a lot on who you are, how you were taught, and what other languages you already happened to know). This book is for both kinds of English-users, but it is not fundamentally a book about English. It is a history of scientific languages, the set of languages by means of which scientific knowledge has been produced and communicated. Whether

*“The scholars of other nations, to whom we have provided an example, believed with reason that they would write even better in their language than in ours. England has thus imitated us; Germany, where Latin seems to have taken refuge, begins insensibly to lose the use of it: I do not doubt that it will soon be followed by the Swedes, the Danes, and the Russians. Thus, before the end of the 18th century, a philosopher who would like to instruct himself about his predecessors’ discoveries will be required to load his memory with seven to eight different languages; and after having consumed the most precious time of his life in acquiring them, he will die before having begun to instruct himself.”
you are a scientist or have studiously avoided the sciences throughout your life (so far), the history of scientific languages is a constitutive part of your world. The story ends with the most resolutely monoglot international community the world has ever seen—we call them *scientists*—and the exclusive language they use to communicate today to their international peers is English. The collapse into monolingualism is, historically speaking, a very strange outcome, since most of humanity for most of its existence has been to a greater or lesser degree multilingual. The goals of this book are not only to show how we came to this point, but also to illustrate how deeply anomalous our current state of affairs would have seemed in the past.

For both ends, I have introduced what may seem the book’s oddest feature: the footnotes. Every quotation in the text, except the epigraphs, appears in English. (The epigraphs, as you can see right here, always appear in their original language, and are translated in the first footnote.) For any quotation that was originally composed in a language other than English I have, where possible, tracked down the original and reproduced it on the bottom of the page, in its original orthography, with my own translation in the text. (When I have been unable to do so, I explicitly credit the translator.) I do this not because I am a perfecttranslator, but rather because I am a flawed one. You may indeed find mistakes in some of the renderings, and that is precisely the point. Every history has those flaws, but I want to expose the reader to the friction caused by languages one knows imperfectly, the alienating quality of other people’s words, to make the active translation visible. The past did not happen exclusively in English, though many histories make it seem as though it did. The footnotes also make the historical trajectory evident: as the book progresses, fewer and fewer footnotes appear; that’s because the conversation in science has transitioned to English. (The footnotes can also be fun. Try reading Esperanto—you might like it!) Likewise, many of my sources wrote in foreign languages poorly. I have left their bad spelling and grammatical infelicities unadorned by the scholarly “*sic*,” except in cases of typographical error. You are also, of course, free to ignore the footnotes and read the text through entirely in English. That is, in truth, how most of science is done today.

But it wasn’t always that way: the languages of science used to be multiple. This is a book about *scientific languages*, and I use both terms with their most straightforward meanings. I certainly do not mean that some languages are intrinsically “more scientific” than others (although many have made such claims in the past and still do today, as we will
I define *science* rather narrowly, consistent with modern Anglo-American usage, to refer to what are often further specified as the natural sciences. To be even more precise, I focus on the comparatively small community of elite, professional scientists, a community that has engaged in international communication for centuries and maintains to the present the highest prestige among investigators of nature. (I exclude here medicine and certain applied sciences, such as agronomy, in part because those practitioners’ need to communicate with a nonscientist client base introduces significant complicating issues of popularization that are ancillary to the main issues in this book.) The narrowness of *science* in English is distinctive. Other languages, such as French (*science*), German (*Wissenschaft*), or Russian (наука, nauka), use the term to encompass scholarship in a broad sense, including the social sciences and often also the humanities. I follow English usage simply out of conceptual economy, although the ways languages have shifted in those disciplines are interesting and they exhibit a similar linguistic narrowing as the “natural sciences.” The natural sciences (physics, biology, chemistry) display the phenomenon I am tracking more vividly. I emphasize these sciences because they are at present almost exclusively in English, and they have been so for decades. If you are interested in what it would be like to live in a world with one language of communication, a world with no Babel, you should look to the natural scientists. They come from there.

At one level, the history of scientific languages is recorded in academic publications, as different scholars investigate nature and then try to persuade their colleagues of the detailed organization of the universe. But it is also a story of informal correspondence, friendly banter at conferences, government reports about the transformation of the scientific infrastructure, press releases, anti-Semitic diatribes, and muttering to oneself during a lonely night in the laboratory. This book ranges from the poetry of ancient Rome to attempts to communicate with alien civilizations, from the nationalist conflicts of the nineteenth century to the dawn of computerized machine translation, with a cast of characters including the greatest scientists of their day as well as (almost) anonymous librarians, politicians alongside linguists, frenzied debaters over the merits of artificial languages spoken by only a few dozen contrasted with attempts to standardize a language across the largest land empire the world has ever seen. It is an intimate and a public history, as befits language—something we all feel intensely about, while at the same time sharing it with communities of strangers.
Here is a truism: scientific activity is communicated in a language. I do not simply mean “in words”; I mean in a particular, specific language, shared by a community of speakers. People can have scientific thoughts, do scientific experiments, have scientific conversations, in whichever language they wish to use—in theory. But in practice, science has not been so conducted. Scientific findings are not usually communicated in Ibo, Bengali, or Polish, at least not at the dawn of the twenty-first century, and not at the dawn of the nineteenth, either. Science, as a lived human activity, has always traveled within a highly constrained set of languages. If we adopt the narrow stratum of elite science and look at the dominant languages in which it has been communicated to the international community of researchers from the beginning of recorded history to now, we end up with a rather limited list. Taking languages that register a statistically significant proportion of the world production of something we might now call science, we find (in alphabetical order): Arabic, Chinese (classical), Danish, Dutch, English, French, German, Greek (ancient), Italian, Japanese, Latin, Persian, Russian, Sanskrit, Swedish, Syriac, and Turkish (Ottoman). (I apologize for those I have excluded at the edges; even if you include them, the list does not grow significantly.) There is no other sphere of human cultural activity—trade, poetry, politics, what have you—that takes place in such a limited set of tongues.6 Behind the truism, therefore, is a fact of tremendous importance. This book is about life in Scientific Babel: how scientists managed to work among this (limited) profusion of tongues, how they hoped to conquer it, and how it came about that the Babel was no more.

Every time you utter something, you need to balance between two competing demands. On the one hand, you would like to express your internal notions, to say exactly what you are thinking or feeling. Of course, this is an ideal; we have all experienced the disconnect between what’s in our minds and the clumsiness by which we can formulate it.7 Yet, for most of us, we get closest to this ideal in our native language or in the language we use most fluently; it is, fundamentally, a speaker-centric choice. I call this identity, and it is surely possible for a particular speaker to have multiple distinct identities, speaking to children in her role as a parent most easily in one language, to a spouse in her role as a wife in another, at work as a lawyer in a third. Nonetheless, in this kind of speech, the speaker focuses on the capacity to express herself or himself in that particular role. But what about the audience? With
most utterances, you have some particular recipients in mind, real or imagined, present or absent. You want your interlocutor to understand what you say, and this is easiest to achieve by using the language your listener (or reader) understands best, or at least the strongest language you have in common—that is, using what is called by linguists a vehicular language. This choice is audience-centric, and I describe it as communication. Irreducibly, all utterances occupy a spot on the continuum, trying to express oneself as accurately as possible while at the same time making efforts to be understood correctly. The tension exists within a single language—I am not certain that even now I am presenting my thoughts accurately in what is both my native language and a vehicular language we have already established you understand—but the challenge is magnified significantly when you add language barriers to the mix.

Scientific utterances are no different from ordinary utterances in this regard. Today’s overwhelming dominance of one vehicular language may give the impression that science naturally trends toward communication and away from identity, since one’s scientific peers need to vouchsafe the validity of one’s claims—and, indeed, today science works this way, which helps explain the pressure toward fewer languages. But not necessarily to a single one, for there was a moment when European naturalists had a single language—it was called Latin—and they deliberately, consciously chose to give it up. Latin remained a language of communication, but it was joined by Dutch, English, Swedish, Italian, and some others. Identity was allowed in, to a certain extent, for a particular range of tongues. (One might also understand this as communication with a different, more local audience, as we will see.) Where communities fall on the spectrum between identity and communication is historically contingent; different tensions are tolerated differently at different times, but they have not gone away, even if scientific communication happens in a single language. It is, in fact, an omnipresent feature of all interchange, strongly dramatized in the case of science by its prominent intellectual creativity (identity) and its social organization (communication), and that allows us to see how creativity and social organization interact within the spheres of language and language choice. Yet the dilemma is not symmetric. If you are a native speaker of English, your language of identity equals your language of communication; your burden is reduced to the irreducible problem of saying what you mean, shared by all speakers everywhere, without the additional load of strug-
gling with a foreign tongue. That is an enormous privilege, but it is a privilege that Anglophones are largely blind to. One goal of this book is to make visible this asymmetry and its consequences.

English is dominant in science today, and we can even say roughly how much. Sociolinguists have been collecting data for the past several decades on the proportions of the world scientific literature that are published in various tongues, which reveal a consistent pattern. Fig. 0.1 exhibits several striking features, and most of the chapters of this book—after an introductory chapter about Latin—move across the same years that are plotted here. In each chapter, I focus on a language or set of languages in order to highlight the lived experience of scientists, and those features are sometimes obscured as well as revealed by these curves. Starting from the most recent end of this figure and walking back, we can begin to uncover elements of this largely invisible story. The most obvious and startling aspect of this graph is the dramatic rise of English beginning from a low point at 1910. The situation is actually even more dramatic than it appears from this graph, for these are percentages of scientific publication—slices of a pie, if you will—and that
pie is not static. On the contrary, scientific publication exploded across this period, which means that even in the period from 1940 to 1970 when English seems mostly flat, it is actually a constant percentage of an exponentially growing baseline. By the 1990s, we witness a significant ramp-up on top of an increasingly massive foundation: waves on top of deluges on top of tsunamis of scientific English. This is, in my view, the broadest single transformation in the history of modern science, and we have no history of it. That is where the book will end, with a cluster of chapters focusing on the phenomenon of global scientific English, the way speakers of other once dominant languages (principally, French and German) adjusted to the change, preceded by how Anglophones in the Cold War confronted another prominent feature of the midpoint of the graph (1935–1965): the dramatic growth of scientific Russian.

But, on second glance, one of the most interesting aspects of this figure is how much of it is not about English, how the story of scientific language correlates with, but does not slavishly follow, the trajectory of globalization. Knowledge and power are bedfellows; they are not twins. Simply swinging our gaze leftward across the graph sets aside the juggernaut of English and allows other, overshadowed aspects of these curves (such as the rise of Russian) to come to the fore. Before Russian, in the period 1910 to 1945, the central feature of the graph is no longer English but the prominent rise and decline of German as a scientific language. German, according to this figure, was the only language ever to overtake English since 1880, and during that era a scientist would have had excellent grounds to conclude that German was well poised to dominate scientific communication. The story of the twentieth century, which from the perspective of the history of globalization is ever-rising English, from the perspective of scientific languages might be better reformulated as the decline of German. That decline started, one can see, before the advent of the Nazi regime in 1933, and one of the main arguments in this book is that the aftermath of World War I was central in cementing both the collapse of scientific German and the ballistic ascent of English. We can move further left still, and in the period from 1880 to 1910 we see an almost equal partition of publications, hovering around 30% apiece for English, French, and German, a set I will call the “triumvirate.” (The existence of the triumvirate is simply observed as a fact in this book; I do not propose to trace the history of its emergence.) French underwent a monotonic decline throughout the twentieth century; one gets the impression (although the data is lacking) that this decline began before our curve does, but to participants in the scientific
community at the beginning of our modern story, it appeared stable. My narrative for this earlier period comes in two forms: the emergence of Russian, with a minor peak in the late nineteenth century, as the first new language to threaten to seriously destabilize the triumvirate; and the countervailing alternative (never broadly popular but still quite revealing in microcosm) to replace the multilingual scientific communication system with one conducted in a constructed language such as Esperanto. Long before all of this data, all of these transformations, there was Latin, and that is where the book properly begins.

For all the visual power of the graph, most of this book pushes against its most straightforward reading: the seemingly inexorable rise of English. Behind the graph lie a million stories, and it is history’s task to uncover them. There are other reasons for caution. For starters, we must be careful not to take its quantitative proclamations as gospel truth. The data comes from abstract journals: periodicals that supply an index of abstracts of scientific publications every year, an index to assist in taming the avalanche of information. (The history of these objects is an important subplot in this book.) A Japanese bibliographer named Minoru Tsunoda gathered a list of percentages of publications from numerous abstract journals (which he chose to publish in French, but in a Japanese journal), and then German sociolinguist Ulrich Ammon—the leading researcher today on the question of scientific languages—plotted the information in graph form, updating it as new information came in.11 Abstract journals are, however, already a simplification of global production, and what we see here is therefore a selection of which periodicals abstract journals have chosen to include, and this culling obviously biases the results toward the dominant languages. For example, 5,986 scientific and technical journals were published in Brazil in 2007, but only 17 were registered in the Science Citation Index, and therefore the majority do not show up in this kind of data.12 That obviously hurts the statistics for Portuguese (although quite a few of those journals might publish in several languages, or exclusively in English). Abstract journals, although they do reflect how elite scientists encounter the cutting-edge literature in their fields, do nonetheless generate some distortion, and we should view this curve more to gain a qualitative impression rather than a rigorous result. And that impression is extremely difficult to ignore.

As is evident from the above, I use the word language in a specific, but rather everyday, manner. I have not written a technical linguistic study, but neither do I use language in a literary fashion. There is a sense
in which we can talk about “scientific languages” metaphorically: that scientists use a jargon that is not the same as ordinary language; or that biologists and geologists “speak different languages”; or that each individual laboratory has its own particular idiolect that outsiders have a hard time penetrating. Much of the scholarship on the history of science and language concerns this metaphorical sense, and a good deal of it is of the highest intellectual rigor and utterly fascinating. However, precisely this sense, which I will refer to as discourse, is not my quarry here. I mean language in the brute forms of English, Swahili, Korean, or Russian. That is, I am interested in which languages people choose to use—and not use—in various contexts, at different times, in assorted places. I explore the history of these scientific languages mostly from 1850 to the present (although with a necessary excursus into Latin at the beginning), and with a principal focus on Europe and North America, with occasional visits to other parts of the globe. The comprehensive story is obviously bigger than that and could include all of the world over all of recorded history. I restrict myself to this narrower swath for two reasons: one intellectual and one practical. The first is that the phenomenon of global English started there, as did the basic institutions of modern science that were exported (sometimes forcibly, sometimes not) to other parts of the world. That is one significant justification for limiting this first pass, leaving you with a book of manageable size you can hold in your hands.

The second reason is no less important: the languages I happen to know are a subset of these languages of European origin, and I cannot write a history from sources I cannot read and understand. That is a frank admission of ignorance, and you don’t come across such things very often in books like this one, but without it you will lack a crucial piece for understanding not just this specific book, but any book on the question of scientific languages. To write this book I have used sources in English, French, German, Russian, Latin, Esperanto, and the latter’s offshoot, Ido. I hesitate to say that I “know” these languages, because competence in a tongue is always a relative matter, and I am more fluent and subtle in some of these languages (my native English and also Russian) than in others (French and German), and some, such as the Latin I learned in order to write this book, are still very much works in progress. I pen this confessional paragraph to illustrate several points that condition the following historical narrative.

The first is that knowing a language is measured by a standard that changes over historical time. Many of the scientists I discuss read and
published science in three or four languages as a matter of course. Was this a vanished race of polyglot naturalists? Of course not. Some of them were more linguistically gifted than others, to be sure, but most of them managed with a dictionary and consultation with those more adept (as I often did). Today, scientists expect their peers to be relatively fluent not just in reading and writing English, but also speaking it. The standard of fluency has gone up; the standard of quantity has gone down. The second point is that I happen to read these languages and not others. I chose to learn Latin to write this book. I wish I had the time and energy to learn Japanese, which has an important role to play in the history of scientific languages in the twentieth century, or Dutch, which was central in the seventeenth and eighteenth, or Italian, which continued its salience into the early nineteenth century. If I had, the story you read would be different. (I particularly regret the comparative neglect of East Asia in this account.) The few extant studies of scientific languages are written by those who do not know Russian, and those renditions look rather different than mine, which emphasizes that language quite a bit. I hope that those with different linguistic capacities—or even the same ones, calibrated to different degrees—will take the question of languages in communicating knowledge and run with it. We need more, and more diverse, accounts.

It is necessary to state all this up front because of the seemingly universal phenomenon of linguistic citation bias. Scholars disproportionately cite literature in the languages they feel most comfortable with, which are often their native languages. According to results cited in one 1981 study, American and Indian journals offer citations that are 90% to English-language literature, which was greater than the proportion (roughly 75%) of English material in the scientific literature in that day. Quality and even relative quantity, therefore, is not a full explanation. Likewise, the French cited 29% French, Germans 22% German, Japanese 25% Japanese, Soviet researchers 67% Russian—all in greater proportion than the baseline literature would suggest. (Articles in Chinese were cited only in China, for example.) I doubt I am an exception to this rule—many of my citations are to Anglophone literature, and I include almost no citations outside of my dominant linguistic core set. The scholarship you read is always biased by the linguistic capacities of the scholar. It’s only honest to admit it.

This is all the more important because of a very widespread notion that translation is trivial with respect to science, such that some studies neglect to mention a language barrier at all, or recognize that “[a]l-
though language of publication is an inescapable feature of scientific communication, it is most often treated as background noise.” ¹⁷ Or, in what amounts to the same thing, that science has uniform content and is therefore beyond translation: “Scientific prose has in fact a valuable and a not uninteresting characteristic—almost alone among all the different categories of prose it can be translated into languages other than the language in which it was first written, not merely satisfactorily but perfectly.”¹⁸ Such statements are based on a philosophical assumption that scientific claims represent the world unfiltered, and therefore scientific utterances are a kind of “metalanguage” that are only partially expressed in any individual tongue but are equally true in all of them. This belief is a central one to many of the scientists we will encounter in this book, but it is a view that is complicated by the experiences of those individuals who daily have to translate between and among various scientific languages. For them, translation has been a source of frustration, and often conceptual confusion.

The power of this notion of a metalanguage stems from the unquestionable success of mathematization of the sciences.¹⁹ When I have discussed this project with both scientists and humanists, I have often been told that there is no need to pay attention to the languages in which science is written because scientists can simply read the equations and figure out what is going on. This might be true in certain cases, but it is hardly true generally. Even for an ostensibly “hard” science like chemistry, papers contain more than isolated chemical formulae and mathematical equations. You read descriptions of the reaction, analyses of colors and odors, detailed explanations of method. This verbosity is one of the reasons why the emphasis in this book will be upon chemistry, which shares both in mathematical formulations and in more descriptive scientific traditions, and therefore exposes the capacities and limits of each. Even in cases of strongly mathematized sciences, like classical mechanics, a bare equation never tells you all you need to know. Consider this simple one:

\[ M \propto WgT^2/l \]

What does it say? Without further context, you can tell me that \( M \) is directly proportional to the square of \( T \), and inversely proportion to \( l \), but what does it mean? Mathematical equations are incredibly powerful tools, economically expressing detailed relationships and enabling stupendous manipulations that seem impossible without the formal-