

but by the cumulative effect of countless millennia of transformation proceeding through structured chaos. Although Japanese developed its hierarchical vocabulary alternatives because of the culture, in the meantime myriad sound changes, extensions, grammaticalizations, reanalyses, and random driftings in word meanings were taking place. The results of these processes constitute about 98.5 percent of the task in learning the language and had no more to do with Japanese culture than the rise of *Ich habe ein Bier getrunken* did to the German "soul." This is, and has always been, the case in all of the world's languages.

We cannot know what the words or structure of the first human language were, but we do know that one must have existed. Its outlines come gradually to us out of the mists of time. Our first indications are some possible handfuls of words deduced backward by comparing Proto-Indo-European and some other reconstructed family protolanguages. Our first *concrete* records of language come with the first written materials: Sumerian cuneiform inventories, Egyptian hieroglyphic narratives, Mayan inscriptions. By that time, the Ur-tongue, ever transforming over generations in each of the offshoots of the founding band of *Homo sapiens sapiens*, had already propagated into thousands of variations, which would all have been utterly incomprehensible to the East Africans who had developed their original progenitor: Egyptian obelisks, the *Rig-Veda*, the April 22, 1877, issue of the *New York Herald-Tribune*—all of this writing is mere snapshots of yesterday's cloud formations. Human speech is structured variation, like Haydn's string quartets or the images in a kaleidoscope after each shake: within the bounds of anatomy, human cognition, and the exigencies of social harmony, the first language took on a dazzling and infinite variety of permutations.

Now that we've come this far, would it beg the reader's forbearance if I revealed that, in the true sense, there is not even really such a thing as "a language" at all? It's the nature of language change that makes the concept of "a language" logically impossible, and your having read this chapter allows me to share the reasons for this impossibility in the next one and, in the process, fill in our story of what happened to the first language.

## 2

### The Six Thousand Languages Develop into Clusters of Sublanguages

The first time I went to Germany, I was pleased to find that my German was good enough to order meals, get a room at the hostel, and even understand enough to not "stick out" too much in German-speaking social situations. However, my joy was short-lived. I went to Konstanz, a town in the south of Germany, to stay with some German college students I had met in the United States. I soon found that, as soon as the beer started flowing and the conviviality level went up, I quickly lost any ability to understand a word anybody said. I will never forget the evening I spent at a local pub with them, as joke after joke in what might as well have been Navajo to me elicited rafter-raising howls of laughter and calls for more beer. Yet to my knowledge not a word of Navajo was uttered that night in Konstanz: all of these people were speaking German—or at least what they called German.

What I had run up against is a fact about languages that modern American life tends to relegate to the margins of our consciousness. Atoms are not the irreducible entities that scientists once supposed; instead, atoms are complexes of subatomic particles. In the same way, viewed up close, most "languages" are actually bundles of variations on a general theme, dialects.

By this, I do not mean that there is "a language" that is surrounded by variations called "dialects." As will end up being a kind of mantra for this chapter, "dialects is all there is." One of a language's dialects is considered "the standard," but this anointment is a mere geopolitical or cultural accident. Standard German for

"They're scuttling our ship! We're going under!" is *Sie machen unser Schiff kaputt! Wir gehen unter!* In the dialect that frustrated me, Schwäbisch, it would be *Dia machat onser Schiffe hel Mir gangat unter!* To the German, real-world sociological associations make the second sentence leap out as "other," "quaint," "rustic," and perhaps even "not 'real' German." To a foreigner familiar with standard German, it looks just plain weird—a kind of twisted rendition of what we were taught as "German" in textbooks. But a Martian, presented with both sentences, would find no way of designating one as "the real one" and the other one as "a variation"; they would just look like two similar systems, just as a Burmese and a Siamese cat are to us different but equal versions of the same basic entity.

And in fact there *is* no "default" cat; there are only types of cat. Language change parallels biological evolution not only in creating different "languages" equivalent to species, but in that most languages consist of an array of dialects equivalent to subspecies. As such, not only did the first language evolve into six thousand different ones, but most of these in turn evolved into what, taken together, is an untold number of subvariations on those languages.

### How Do Dialects Arise?

Dialects follow naturally from the inherently nondiscrete nature of language change. Latin developed into several distinct languages when populations of its speakers were dispersed throughout Europe. As we saw, however, the new languages appeared not abruptly, but by a gradual process in which there was no inherent dividing line between "Latin" and "new language." For centuries, the Latin of the region now known as France was a variety halfway between Latin and French, which I have facetiously called "Fratin." Similarly, in Spain there was once a "Spatin," in Italy a "Latalian," etc. In other words, had we toured Europe in roughly A.D. 1, we would have found varieties of Latin not so distinct as to strike us as "separate languages," but distinct from one another nevertheless. At the time, they were still, in our terminology, new *dialects* of Latin.

We have concentrated on language change as it operated in the past to lead to today's languages. Yet because human speech is inherently mutable, it follows that today's languages are slowly

undergoing the same transformative process that Latin did. In this light, a question arises about, for example, English.

### The English They Don't Want You to Know About

English is one of several languages that evolved from an unwritten ancestor linguists call Proto-Germanic; other Germanic languages include Dutch, Swedish, Danish, Yiddish, and (three guesses!) German. The different languages resulted from Proto-Germanic speakers settling in different locations, England being where English developed.

Okay, good—but obviously England is not a giant open field where all of its speakers interact with one another on a daily basis in a grand, mad, Woodstockian splash of teeming humanity. On the contrary, the Nordic invaders who conquered and took over the island in A.D. 449 quickly spread their language in all directions, and in subsequent centuries English came to be used in hundreds of separate regions, by people most of whom rarely ventured far from their villages.

We would expect, then, if Latin became several languages once it was spread among several separate populations, that English itself would have begun developing into different languages, just as Darwin's Galápagos finches began developing into distinct varieties once let loose on those islands—and this is exactly what happened. In the England of our moment, the process has not gone far enough to lead to separate languages. Instead, just as there was a time when today's Romance languages were still dialects of Latin, today's British English varieties are dialects of English, recognizable as "the same language" but quite distinct nevertheless.

It is important not to think of the regional varieties as having evolved *from* the Standard English we know today; the invaders of the continent did not arrive speaking like George Washington, or even Shakespeare or Chaucer. They arrived speaking Old English, and today's standard evolved *alongside* the ones that were eventually relegated to "regional" status. Thus the common source of all of today's English dialects is that queer-looking tongue in which *Beowulf* is written, a tongue that is no longer alive. Old English developed in different directions in each region, and thus each dialect developed its own sound changes, extensions, grammaticalizations, rebracketings, and semantic changes.

One gets hints of this on the British sitcom *Are You Being Served?* Ladies' Intimate Apparel sales assistant Miss Brahms is a Cockney and at first threw me a couple of times with sentences like *An' it's expensive an' all: an' all* has evolved semantically in Cockney to mean *too*. Ending a spectacular run in 1984, most of the cast members reunited to tape two more seasons in 1992, this time assigned to run an upcountry farm. The locals pronounced Mr. Humphries' name "Mister 'Omphries"—just as *Frasier's* Daphne would say it—because evolution of the *uh* sound here differed from that in other parts of England, and they said *sommet* for *something*, because the Old English source of *something* evolved through different sound changes in this area.<sup>1</sup>

Yet this kind of thing is just the tip of the iceberg. Here, for example, are some sentences in the English spoken until about a century ago in rural areas of the southwestern region of Cornwall:

Aw baint gwine for tell ee.

Th' Queeryans do s'poase the boanses ded b'long to a helk.

Ded um diggy ar no?

Billee, 'ome, d' b'long gwine long weth 'e's sister.

At first glance, the first sentence does not appear too opaque: "I'm not going to tell him," right? But no—it means "He isn't going to tell *you*": *aw* evolved from Old English's *hē* (pronounced "HAY") through different sound changes from those of Standard's *he*, and *ee* is an evolution of the initial *y* in *you*. And notice *for* instead of *to*. *Baint* is a rebracketing of *be* and *not* into a single form. Its ultimate source is extension: we tend not to notice that English uses no fewer than three different roots to inflect *to be*: *am/are/is*, *be/been*, and *was/were* derive from what used to be separate verbs, all of which inflected on their own in all six person/number combinations in all tenses. Standard English puts up with the odd division of labor among the three roots—no dialect, or language, irons out any but a fraction of the things that could be. But Cornwall English, like other

1. Amazingly, the revival, called *Grace and Favour* in England but *Are You Being Served Again* in the United States, was every bit as good as the original series, and it is to be hoped that the cast blesses us with at least one more go-round.

Southern dialects, extended *be* into negative constructions instead of only allowing *am/are/is* into that area; thus *I be not*, later *I baint*.

The second sentence is "The Antiquarians suppose the bones belonged to an elk." The *do* and the *did* are not meant for emphasis. The *do* was used quite neutrally, having been *grammaticalized* otherwise than in Standard English. Standard English uses *do* this way only in negative sentences: *I don't know*; in Cornwall you could also use it in affirmative sentences—what began as a full verb *do* evolved into a faceless little piece of grammar. The *ded* for *did* is a grammaticalized marker of past, just as our *-ed* marker is (it has been guessed that even *-ed* evolved from what began as *did*). *Helk* represents the extension, virus style, of initial *h* onto words that at first did not have it. In this as in many regional English dialects, sound change tended to erode *h* from the beginning of words (*Yella, mate!*). Yet because changes go through a phase where the old and the new form exist side by side, speakers who tended to let the *h* go were aware that there was both an *h*-less and an "*h*-full" version of the word: *hella/yella, horse/orse*. Notice that dropping the *h* leaves the word with a vowel up front: for many speakers, a feeling set in that *any* vowel-initial word might have an *h*-full alternate, and the result was words like *hellk*.

*Ded um diggy ar no?* is "Did he used to dig or not?" The *diggy* is not just a "cute" way of saying *dig* in the vein of "diggy-poo"; the sentence does not, for example, refer to a child. In Cornwall English, the suffix *-y* is a piece of grammar, used specifically to convey that an action is repeated. The fourth sentence is particularly confusing for us. The *b'long* does not literally refer to possession but is another marker of habituality, where the full meaning of *belong* has grammaticalized, and thus the sentence means "Billy, at home, usually goes with his sister."

This, then, is what happened to English as it mutated down in the southwest, as at the same time farther northeastward it was mutating into the standard variety we are familiar with. Meanwhile, up in the Midlands in Nottinghamshire, in local speech until not long ago one might hear *Tha mun come one naight ter th' cottage, afore tha goes: sholl ter?* Some readers might recognize this as one of the sentiments expressed by the gardener in *Lady Chatterley's Lover*; whereas Lady Chatterley at one point snaps, "Why can't you speak normal English?" the gardener's speech is simply the direction into which English had evolved in this region in contrast with her

standard-speaking one. In Farnworth, north of Manchester, two forms of both *yes* and *no* have evolved. *Yes* is *aye* under usual circumstances, but if one is contradicting a negative statement, then the form is *yigh* (pronounced like *aye* but with an initial y): A: *I can't find the scissors.* B: *Yigh, they're here.* *No* means *no* usually, but you contradict a statement someone just made with *nay*: *Nay, by gum! I'm not having that!* (They really do say "by gum.") In French, *si* is used similarly: A: *Tu ne l'aimes pas.* B: *Si, je l'aime!* (You don't love him. Yes, I do!), as is German *doch*. There's no reason why English shouldn't have this; Standard English just happens not to have taken this route. Farnworth English, however, did. All non-Standard English dialects use "double negation" (*I ain't got none*), but Farnworth English also allows particularly spectacular negation Dagwood sandwiches: *I am not never going to do nowt no more for thee.* (*Nowt is nothing*.)

One can think of dialects as different recordings of a pop song. "I Say a Little Prayer" was first recorded in a perfect three minutes by Dionne Warwick. Later, Aretha Franklin did a version, an equally perfect, but different, three minutes. One day, Mariah Carey, in her quest for true diva-hood, will most likely record the song, and that will be another fine three minutes. Even Luciano Pavarotti might give it a go some day. There is no "blueprint" "I Say a Little Prayer" in the way that there is a "blueprint" "Sempre libera" from Verdi's *La Traviata*, solemnly imprinted with melody and accompaniment precisely specified. There is a sheet-music version of "I Say a Little Prayer" for piano, but it's just an anemic little toss-off by some anonymous house arranger at the publishing company, hardly as colorful as Burt Bacharach's orchestration for the Warwick recording. But then Bacharach did not mean this first recording as "The One," and put together different orchestrations of the song for later recordings. And then the sheet music is just designed to accompany a singer anyway—neither Bacharach nor anyone else would consider Aretha Franklin's rather free, often improvised approach to the melody a "violation of the composer's original intentions."

Properly, there is no "Ur-text" "I Say a Little Prayer"; it exists *only* as various interpretations of the basic outline. Standard German and the Schwäbisch dialect I ran up against or Standard English, Cornwall English, and Cockney English exist in the same relationship as the various renditions of that song. Moreover, there is no "broken down" dialect that stands in relation to the others as

the party revelers' drunken, out-of-tune rendition of "I Say a Little Prayer" in the movie *My Best Friend's Wedding* does to the Warwick and Franklin recordings. Each dialect is just a different roll of the language-mutation dice.

Along these lines, as often as not, a language comes into existence split into different dialects from the very beginning, there never having existed any single original variety. Romans did not settle Gaul in a single clump, but spread out through the area, passing their language on to separate populations. There was enough intermigration and travel that the new Latins that developed in the northern area maintained a fundamental kinship with one another (whereas in the south things went so much their own way that a distinct language developed, now called Occitan); nevertheless, what became "French" differed significantly from region to region.

We saw how, in what I have called "French," the *k* sound often became the *sh* sound, such that Latin *canem* became French *chien*. This did not happen in all French dialects, though; it did in what is now the standard one, but in the northwestern regions of Normandy and Picardy, *k* stayed *k*. A piece of coal in Latin was *carbo*; in Standard French it is now *charbon* [shar-BAW<sup>ue</sup>], but in Normandy and Picardy it stayed *carbon*. At first, French dialects had *ei* where the standard now has *oi* (pronounced "WAH<sup>ue</sup>"); the midpoint between Latin *habere* and Standard French *avoir* was an earlier *aver*. Because language change is a chance affair, in the Norman dialect the change to *oi* never happened to transpire. That is why, in my fourth-favorite city in the world, there is a Mount Royal (*royal* being the Standard French form), but the city it is in, settled by people most of whom came from northwestern France rather than Paris, is called *Montréal*, *réal* being *royal* in Norman French. Over in the east in Lorraine, a special past tense has grammaticalized through the evolution of *oi*, which means roughly "now," into an ending: *j'étozor* means "I was" in situations where something else is about to break in on the proceedings. One of the oddest things about French for English speakers is learning that there is a special past tense used only in writing: one first learns that *he spoke* is *il a parlé*, but then finds *il parla* in novels and stuffy writings. In western and eastern dialects, however, these forms were commonly used in speech for centuries after they fell by the wayside in the standard.

Similar facts could be trotted out about most languages on earth spoken by groups larger than a village or two. Even Fijian, spoken

on a complex of islands by just seven hundred thousand people, has more than one dialect. If geography or culture ensures that subsets of a group speaking a language interact and identify more with one another than with the larger set of all people who speak the language, the inevitable result is language change, as in cooking, art, music, and dance, developing in divergent directions. Outcome: dialects—Szechuan, Hunan; *Did he used to dig? Ded um diggy?*

### All Systems Normal—At Least for Now

One thing that follows simply and ineluctably from this is that, despite the almost irresistible pull of the sociologically based evaluations that attach to dialects, there is no such thing as human beings speaking “bad grammar.” There are no dialects in any way analyzable as “decayed” versions of the standard or of anything else. Why would speech “decay” down in Cornwall but keep a stiff upper lip in the central Midlands? Why would Latin “crumble” in Picardy but for some reason just “evolve” around the Île-de-France?

It is almost sobering to realize that the social evaluations we place on how people talk are purely artificial constructs placed on speech varieties that neither a Martian nor often even a foreigner unfamiliar with our social terrain would arrive at on the basis of recordings of the speech alone. This observation goes down easy when we think about peasants in Picardy, of course—because we have not been steeped in the social evaluations particular to France. It is harder to truly wrap our heads around this here at home, though.

The Appalachian English that sounds so twangy, rustic, and full of “mistakes” (that is, ways in which it has mutated in directions other than Standard English has) would be just one more variety of English to our Martian—and we’re talking Snuffy Smith and Li’l Abner here, not just a “country” accent. Black English, America’s most controversial dialect, which even the most well intentioned of people often see as “bad grammar run wild,” developed through the same processes of change as those of any other dialect and thus stands equal to any other in the qualitative sense.

I have never heard the common conception that nonstandard dialects are “bad grammar” put as eloquently as when an elderly black woman in the Mississippi Delta once said to me that, from what she saw, “Seems like most people speak pretty good English, but some people, it seems like they just talk!” To her, it naturally

seemed as if the Southern Black English dialect spoken by many around her, especially in its “deeper” varieties, was “bad” English rather than alternate English. In this vein, Black English speakers are often accused of having “bad diction,” but this is mainly a trait local to many black male teenagers’ in-group identity and is common in dialects spoken by male teens in many societies (listen to them on the streets of Paris or Berlin); in any case, some whole language’s sound systems are just “crisper” than others’. Brazilian Portuguese, for example, is the antithesis of German in this regard, often sounding as deliciously gushy as a ripe slice of mango (think Astrud Gilberto, and this is even more pronounced in running speech). Yet we would be hesitant to accuse the entire country of Brazil of having “slurred speech” in comparison with that of Spanish-speaking countries.

And notice that dialectal differences run wide and deep. “Dialect” is not meant here just as a stand-in for accent and scattered regional terms: things like *Ded um diggy?* go far beyond this into the heart of sentence structures and shapes of even basic vocabulary. In America, the difference between dialects is rather slight in relation to how widely dialects often diverge worldwide. American dialect differences are largely limited to peripheral vocabulary (*pop* instead of *soda*), minor sound differences (*gray* versus *grey*), the fact that for young Californians, *John, Sean, Ron*, and *Dawn* all rhyme, the second and fourth pronounced “Shahn” and “Don”), and a kind of generalized set of nonstandard speech forms such as *ain’t*, double negations, use of *don’t* with the third person singular (*Don’t make no difference to me*), etc. Even Black English, though a tad more divergent from Standard English than this in sound and even in sentence structure, does not strain our sense of what “English” is.

But America is rather exceptionally bland in regard to dialect divergence, just as its flora and fauna pale in comparison with the riot of creatures in a tropical rain forest. In Konstanz, it wasn’t only the slang that was throwing me; I could barely understand when someone asked me whether I wanted another beer (which I of course took, which dampened my comprehension even more). Nor was it just a matter of twisting my ear to an “accent”: at one point in the evening, I seated myself behind one guy who was speaking in a clear, resonant voice at moderate speed and found that I could still barely make out a word he was saying. In Schwäbisch, on top of the slang, different sound changes and differing fates and

redistributions of endings have rendered even basic words into shapes related to, but significantly distinct from, the ones that developed in the standard and have created different endings: standard's *wir gehen* [veer GAY-un] is *mir gangal* [meer GONG-uhl], and so on.

We have never had the equivalent of this in America. We may have to adjust our ears to certain local dialects to an extent, but there is barely anywhere in the country that Standard English speakers could go where, after their hosts had had a few at the local pub, they would find themselves utterly at sea linguistically.<sup>2</sup> By the time America was founded, various aspects of modern societies that tend to retard dialects' mutation were long established. Widespread printing forces a decision about what "standard" speech will be and naturally has a way of enshrining that variety for future generations as "the language" because it lives immortally on paper. The spread of education, conducted in that standard variety, furthers this impression. Both factors pit the standard variety in competition with the local varieties and lead many speakers to lean their speech toward the former in deference to its association with prestige, which diminishes the extent to which the local variety changes in its own directions. Meanwhile, in the twentieth century, radio, television, and films imprinted the influence of Standard English even more deeply.

In Great Britain, English had more than a millennium to develop free from these artificial impediments to language change, the result being dialects such as Cornwall English, which would have baffled the foreigner trained in Standard English just as Schwäbisch did me and would have taken quite a bit of adjustment even for us. Meanwhile, though, the entire timeline of American English has taken place within the constraints of societal trends that have the by-product of retarding language change and thus the mutation rate of dialects.

Of course, printing, education, and the boob tube are as prevalent in Europe today as they are here. That is why I refer to Cornwall English in the past tense; throughout England, Standard English is strangling the old local varieties like *kudzu*, and thus, though I said that English dialects "have not" developed into sepa-

2. The exception would be places where either Gullah Creole or Hawaiian Creole English, generally called "pidgin," were spoken; on these and similar phenomena, wait until we get to Chapter 4.

rate languages with the implication that they eventually will, more properly, modern conditions are such that they never will. Increasingly, local differences are more a matter of local terms and accent than much else. It is now ever harder to find speakers under the age of 106 of regional French dialects as well: for example, no longer do you often hear *il parla* in small towns in the west or east of France.

The extent of this homogenization differs from country to country. Regional Italian dialects are still distinct enough that the student armed with Standard Italian can still have the same experience I had in Konstanz at a bar in Milan. The experience I had certainly did not suggest that Schwäbisch is exactly on the ropes, and recently in Leipzig, when electricians came to my office asking something in the local German dialect Sächsisch, the accent alone was so thick that I had to ask a German person down the hall to figure out what they were asking (I never did figure out what exactly they wanted to do). But even in Germany the dialects are diluting among younger generations. *Astrix* adventures have recently been published there in translations into more than a dozen "Mundarts" ("mouth-ways," or dialects) partly in an effort to help celebrate and preserve dialects widely felt to be in danger of imminent demise.

But the linguistically homogenizing tendencies of printing, education, and the communications revolution have set in only in the past few centuries, whereas human language has existed for about 150,000 years, as mentioned earlier. As such, for "languages" to consist of clusters of often highly divergent dialects has been the norm for human language for all but a final hiccup of its existence thus far. This is crucial in understanding that, because of the transformative nature of human speech, the concept of "language" is a mere terminological convenience. There is no intrinsically coherent entity that corresponds to our sense of what a "language" is. There is no heady, abstract, vaguely politicized philosophical argument behind this; it's really quite meat and potatoes.

## Why Dialects Are All There Is

Martians Couldn't Tell:  
"Standard" Dialects Are Just Lucky

Because the standard variety is the vehicle of almost all writing and official discourse, it is natural for us to conceive of it as "the

real deal" and nonstandard varieties as "other" and generally lesser, even if pleasantly quaint or familiar. This state of affairs also tends to foster the misconception that the standard dialect is developmentally primary as well: one can barely help operating on a background assumption that, at some time in the past, there was only the standard dialect but that, since then, nonstandard dialects have developed through the relaxation of the strictures of the standard.

But in fact standard dialects were generally only chosen for this role because they happened to be spoken by those who came into power as the nation coalesced into an administratively centralized political entity. What this means is that there is no logical conception of "language" as "proper" speech as distinguished from "quaint," "broken" varieties best kept down on the farm or over on the other side of the tracks.

**The Right Place at the Right Time** For example, today's Standard French began as just the dialect spoken in the area where Paris is today. It shared France with several other varieties, including those of Normandy, Picardy, and Lorraine, as well as the varieties of the south not even mutually intelligible with the northern ones and thus considered a different language, including Provençal, which was the vehicle of the love songs of the troubadours. Until the late 1700s, this linguistic heterogeneity was not considered a problem in France: in feudal times, the peasant's loyalty to the local lord, who probably spoke the same local dialect as the peasant did, was considered paramount and hardly depended on speaking the king's dialect.

But as the concept of nationalism began to arise, government officials began to be concerned that citizens of *La France* were not united by a common language. As one local official complained:

The multiplicity of idioms could be used in the ninth century and during the overlong reign of feudalism. The former vassals gave up the satisfaction of changing their master for fear of having to change their speech. But today, when we all have the same law for master, today when we are no longer Rougeras, Burgundians etc., when we are all French, we must have only one common language, just as we all share a common heart.

The observation about peasants' "fear of having to change their speech" indicates how distinct many of these dialects were from one another: we're talking Schwäbisch versus Standard German, not "greazy" pork chops. And today's Standard French dialect was indeed a minority dialect in numerical terms: France was a dialectal smorgasbord. Abbé Grégoire, a Catholic priest and revolutionary, was alarmed that:

France is home to perhaps 8 million subjects of which some can barely mumble a few malformed words or one or two disjointed sentences of our language: the rest know none at all. We know that in Lower Brittany, and beyond the Loire, in many places, the clergy is still obliged to preach in the local patois, for fear, if they spoke French, of not being understood.

And if there was a common language to be imposed, it was naturally to be the one used by those in power. Chauvinism was not the only root of this inclination: because the Paris dialect was the one that had been most written for centuries and used by those with the most power, it followed naturally to conceive of it as the "real" French.

From here on, this one of many dialects—now called "French"—was spread throughout France by education as well as an unfortunate dedication to eradicating the local dialects, dismissed as "patois," in the interest of national unity. Thus the dominance of today's Standard French in France resulted from an artificial perversion of an originally much more diverse scenario, rather as the rich fauna and flora of Madagascar have been significantly decimated by the roping of its residents into a global economy that drives them to clear-cut its forests.

In some cases, the standard dialect is even deliberately fabricated by picking and choosing from several local dialects. Today's Standard Finnish was deliberately codified first from southwestern dialects, with elements of eastern dialects interwoven in the 1800s with the popularity of the publication of the oral epic *Kalevala* in an eastern Finnish variety. The utter artificiality of this "standard" is shown by the fact that no Finn, of any educational or socioeconomic level, actually speaks this dialect casually. Newscasters, teachers, and sometimes politicians speak it as a deliberately neutral, "official" code, but there are no Finns who "offstage" speak a

"vanilla" Finnish as, say, the characters in *Thirysomething* or *That 70's Show*, who speak a faceless "generic" English. On the contrary, any Finn's regional origin is clear from his speech. Because no one speaks it except in formal contexts, the "standard" is considered not "the best Finnish" but a utilitarian strategy; Finnish dialects diverge significantly, necessitating an agreed-upon, even if arbitrary, common coin for official purposes.

Standard English developed through a kind of combination of the French and Finnish situations. Standard English incorporates elements from the Essex and Middlesex dialects that happened to be spoken in the London area. By the 1400s, London was the hub of manuscript copying and then printing. Because scribes and printers tended to come from the surrounding regions, features of dialects spoken in these areas made it onto the page more often than others, which in turn meant that people throughout England were more likely to see London dialect written than any other. Combine this with the cultural and commercial influence of London, and the result is a standard dialect perceived as the heart of "English," a star relegating the other varieties to character parts. The shift in attitude toward local speech is visible in the record. About 1490, England's first printer, William Caxton, depicted the differences between London and Kentish English as a matter of apples and oranges:

In my days happened that certain merchants were in a ship in Thames . . . and went to land for to refresh them[selves]. And one of them named Sheffield, a mercer, came into an house and axed for meat. And specially he axed after eggs. And the good wife answered that she could speak no French. And the merchant was angry, for he also could speak no French, but would have had eggs and she understood him not. And then at last another said that he would have eyren, then the good wife said that she understood him well. Lo, what should a man in these days now write: eggs or eyren? Certainly it is hard to please every man by cause of diversity and change of language.<sup>3</sup>

3. I have modernized the spelling and pronunciation to make it easier on the eyes. The original spelling is the likes of: *And theme at lasse a nother sayd that he woude haue eyren . . .*

(Note those *axed's*, which show how arbitrary our sense of "improper" English is; *axed* was accepted and ordinary even in written English at the time.)

But Caxton's equanimity was already on its way out; about 1400, a character in a play had casually dismissed northern dialects as "scharp, slitting, and frotyng and unschape," whose meaning comes through even without the modern equivalent "shrill, cutting, and grating and ill formed," and this kind of judgment was soon commonplace in England. Thus only chance determines that we have eggs Benedict rather than eyren Benedict and process *eyren* as vaguely "unschape."

Today's "Dialect" Is Tomorrow's "Language" Not only has one of many hitherto unranked dialects often been anointed the standard, but we even see dialects actively dismissed as " quaint vernaculars" at point A only to be enshrined as inherently noble vehicles of humans' loftiest thoughts at point B, with nothing but a decisive geopolitical shift at the root of the mysterious change in perception.

Dante wrote at a transitional period between point A and point B in Italy. Latin was still considered the appropriate vehicle of writing; officialdom, and educated discourse, but Italian had developed so far from Latin that its traditional classification as "village Latin" had come to strain natural perception. Dante, afflicted with that queer medieval southern European malady called courtly love, in 1293 dedicated a volume of poems to his adored Beatrice, who combined two traits unusual in a dedicatee of love poetry—namely, having never been touched by the author at any point in her life and being dead. Dante wrote this *La Vita Nuova* in Italian, but only as an extravagant gesture to a woman who understood Latin only with difficulty (not being alive probably made it even more of a chore). Dante usually wrote in Latin, and writing in Italian at all in this period was a nervy gesture. Yet his reason for writing *The Divine Comedy* in Italian in 1308 revealed a similar guiding sense that, at the end of the day, there was something fundamentally "kitchen sink" about Italian:

From this it is evident why the present work is called a comedy. For if we consider the theme, in its beginning it is horrible and foul, because it is Hell; in its ending fortunate, desirable, and joyful,



because it is Paradise; and if we consider the style of language, the style is lowly and humble, because it is the vulgar tongue, in which even housewives can converse.

Indeed, Dante's main grounds for championing Italian were practical; in his *De vulgari eloquentia* (note the Latin title), he urges that Italian be used in literature for the mundane reason that more people understand it than do Latin, which he meanwhile exults as nevertheless the "better" language. Yet though in the 1300s even one of the most masterful bards who ever blessed the Italian language essentially considered it a matter of "Comedy Tonight!"; by the late 1700s, Lorenzo da Ponte, writing the lyrics and libretti of operas such as *The Marriage of Figaro* and *Don Giovanni*, would have been surprised to be told that he was writing in "housewives' Latin." By then, Italian was considered by its speakers and beyond as one of the world's loveliest, most singable and romantic languages. Only the gradual unification of Italy and its ascendance as a world power made the difference.

The Romanian-speaking area extends eastward into a little hump of land called Moldova, much of which for decades was incorporated within the Soviet Union. Moldovan is not just "close" to the Romanian dialects in Romania proper: it is very much one of them, not differing from the standard dialect any more than any Romanian nonstandard one does. The only remotely salient difference between Moldovan and Standard Romanian is that a polite form of the pronoun *he* in the standard is used more informally in Moldovan. Otherwise, most of the differences are minor differences in vocabulary no more dramatic than the ones between American and British English. Within the Romanian-speaking orbit, then, Moldovan had been a "quaint," "rustic" dialect. The Soviets, however, in a quest to discourage Moldovans from identifying with their Romance-speaking neighbors to the west, directly required Russian linguists to foster a conception of Moldovan as a "different language" from Romanian, exaggerating the import of the minor differences inevitable between dialects of any language. Many grammar books of "Moldovan" were little more than translations of Romanian-language Romanian grammars into Russian. Now independent, the Moldovans continue to encourage a perception of "Moldovan" as a distinct "language" from Romanian, in part because Romanians tend to dismiss their dialect as sounding

uneducated. Hence the Moldovan "language," fully intelligible with Romanian right next door.

Don't tell the Scandinavians I said this, but "Swedish," "Norwegian," and "Danish" are all really one "language," "Scandinavian"—people speaking these "languages" can converse. Here is "He said he couldn't come" in all three:

Swedish: Han sade att han inte kunde komma.

Danish: Han sagde at han ikke kunne komme.

Norwegian: Han sa at han ikke kunne komme.

These are even closer than Standard German and Schwäbisch or Standard Italian and Milanese. The Danes used to run what is now Sweden and Norway, and there was no such thing as "Swedish" until Sweden became independent in 1526. What is today "Norwegian" was just "the way they speak Danish in Norway" until Norway broke with Denmark in 1814 and gradually began explicitly working out a standard form of what was an array of nonstandard local dialects.

I once asked two Bulgarians what Macedonian sounded like to them, and they said in unison, "It's a dialect of Bulgarian!" "Macedonian" is indeed so close to Bulgarian that Bulgarians crossing the border need make even less adjustment than Swedes make in going to Denmark. Many Macedonians would find my Bulgarian friends' comment a little irritating, which stems from the fact that "Macedonian" is considered a separate "language" owing to its speakers' distinct political and cultural identity from Bulgarians, reinforced by their incorporation until recently into the Yugoslavian federation.

Although one learns Hindi in different courses from those in which one learns Urdu, the two are dialects of the same language. Hindi, the indigenous lingua franca of India, has taken on a lot of vocabulary from Sanskrit, its ancestor now enshrined in liturgical writings, and Urdu, spoken in Muslim-dominated Pakistan, has done the same from Islam's vehicle, Arabic. Yet this is little more of a barrier to basic communication than, again, that between American English and British English. The sense of separateness conditioned by the profound animosity between Hindu India and Muslim Pakistan extends to linguistic identity and encourages

a sense of separateness between these two mutually intelligible varieties.

Today's "Language" Is Tomorrow's "Dialect" Conversely, there are also cases when a speech variety treated as a "language" at point A is suddenly a "dialect" in the history books at point B, again with mundane events rather than anything inherent to the speech itself having made the difference. Dante wasn't the only medieval European in the throes of courtly love: itinerant musicians in southern France made careers of composing songs to unattainable women of high rank, as the famous troubadors. They did so not in Parisian French but in the particular transformation of Latin that happened to have taken place in their region. Whereas dialects in areas surrounding Paris were similar enough to Parisian to be classifiable as "kinds of French," the dialects of the south were distinct enough to be processed as a different "language" altogether. The dialects of the north were called *langue d'oïl* "oui language," in reference to the word for *yes* in those dialects (not yet evolved to the modern *oui*), whereas the southern ones were correspondingly called *langue d'oc* (hence the region called Languedoc today). Note that in the Gallic consciousness of the period there was no inherent rank implied: the two "languages," or languages, were separate but equal. The dialect of the *langue d'oc* used by the troubadors was called Provençal, and from the 1100s to the 1300s was considered very much a "language," written as well as sung.

Troubadors are *troubadors* [troo-VAIR] in Standard French: the word is the standard dialect's descendant of a Latin root *trōbare* "to compose," whereas *troubador* is how the same root came out in Provençal. This difference is an indication that Provençal was no "kind of French" by any standard; it was very much a horse of a different color. Similarly, in the modern *langue d'oc* descendant (Occitan), *ueh* is eight where French has *huit*, and so on.

But as the center of power concentrated increasingly on Paris, southern France was deliberately yoked politically and administratively into the "French" orbit, complete with transplanted French-speaking officials. This nationalist tide turned against the nations-within-nations that could foster alternate standard dialects, and Provençal and its *langue d'oc* kindred dialects were effectively banished from writing and official contexts. The scene was set for these dialects to be classified as "lesser" rather than "different," con-

cretized by the post-Revolutionary language homogenization policy. By the 1700s, the once-prestigious "language" Provençal was a complex of rural dialects considered mere "patois." The general sense was that these dialects were a "kind of French" when, as we see, though there are no dividing lines to be drawn between "dialect of A" and "language B," Provençal was obviously different enough from Parisian French and the other northern dialects to fall on the "language B" side of the line. Thus the suppression of Provençal was less the silencing of one variety of "French"—sad enough in itself—but of a separate Romance language entirely. We'd rather that a particular subspecies of brown sparrow not become extinct, but the loss is perhaps even greater if we lose *all* subspecies of an entirely separate species of bird, such as pigeons.

Then, in the late 1800s, the poet Frédéric Mistral began a movement to revive these dialects, under the heading of a different name, Occitan. As a result, Occitan is now officially treated as "a language" again, complete with self-teaching materials, novels, and poetry. This cycle eloquently demonstrates that, in the end, dialects are all there is: the "language" part is just politics!

The Ukrainian "language" is a similar story. In Russia, one does not find Russian dialects as distinct from the standard as in Germany and Italy. The main reason for this is that history happens to have fenced off the regions where such "dialects" are spoken as separate cultural and political units, the "dialects" thus officialized as these units' separate "languages." Before the Ukraine was transformed into one of the old S.S.R.s, for instance, it was simply a region of Russia. As a matter of fact, when its city Kiev was considered one of the leading urban centers of Russia early in the last millennium, the dialects spoken there were the closest thing to any conception of "the best" spoken *Russian*, not "Ukrainian," because they were the ones most frequently written.

Starting in the 1300s, when Moscow became Russia's center of government, what is today the Ukrainian "language" became considered a peasant variety of Russian. The difference between Russian and Ukrainian is about the same as that between Standard German and Schwäbisch, often less. Russian for *get married* (when it's a woman doing so) is *vyiti zamuž*; the Ukrainian is *viti zamaž*; the woman herself is *žena* in Russian, while *žona* is one word for woman in Ukrainian.

In college I had a Ukrainian friend who would salutorily take her noble leave of me for the evening with Ukrainian for "good

night" [na doh-BRAH-nich!]. That *Dobranic!* is recognizable from Russian. Russian actually happens to use a different expression, *pokojnoi noci* "peaceful night," but if Russians did use their words for good and night as Ukrainians do, it would be *dobrya noc'*—not "doh-BRAH-nich" but "DOH-braya NOACHY." For a Russian, then, mastering Ukrainian is more a matter of adjustment than precisely "learning."

Edward Rutherford aptly dramatizes the revival of a sense of Ukrainian as "a language" suitable for writing in his page-turner saga of the history of Russia, *Ruska*. In 1827, the poet Karpenko has just recited tales of his Cossack ancestry, and after his friend Ilya suggests that he write them down, he reveals a heretical idea:

"Actually," he confessed, "what I really want is to write them in the Ukrainian language. They sound even better that way."

It was a perfectly innocent remark: though undoubtedly surprising. "Ukrainian?" Ilya queried. "Are you sure?" Olga, too, found herself puzzled. For the Ukrainian dialect, though close to Russian, had no literature of its own except one comic verse. Even Sergei, always willing to support his friend, couldn't think of anything to say in favor of this odd idea.

And it was now that Alexis spoke. . . . "Forgive me," he said calmly, "but the Ukraine is part of Russia. You should write in Russian, therefore." His tone was not unkind, but it was firm. "Besides," he added with a dismissive shrug, "Ukrainian is only spoken by peasants."

Belorussian is even closer to Russian (in fact its name means "white Russian") than Ukrainian is, and it, too, had status as "a language" as far back as the Middle Ages, when its speakers were administrated by Lithuanians who condoned official business in "White Russian," having no particular stake in elevating Moscow's Russian dialect. Thus geopolitics has elevated Ukrainian and Belorussian as official "languages": but if these regions had continued to be subsumed by Russia and if the Soviet Union had had less interest in suppressing unrest by fostering rather than repressing local speech varieties, then Ukrainian and Belorussian would be the Russian equivalents of German "Mundarts," celebrated mostly by a few local advocates and the occasional books of poetry, folktales, or jokes. Foreigners would tell stories about how "the Russian really

gets funky down there in the Ukraine—it was almost a different language!"

Several Languages as Different "Dialects" Finally, just as culture and politics can designate dialects of one "language" as separate "languages," they can also designate languages as distinct as French and Spanish as "dialects" of one. We often hear that a Chinese person speaks the Mandarin "dialect" or the Cantonese "dialect," but in fact the eight main "dialects" of Chinese are so vastly different that they are, under any analysis, separate languages. The Standard German speaker can gradually "wrap his ear around" and "get the hang of" Schwäbisch, but the Cantonese speaker must learn Mandarin as a foreign language. Here is a pair of sentences meaning *I've had my car stolen*:

Mandarin: Wō bei rén tōu le chēzi.

Cantonese: Ngóh béi yáhn tǎu-jó ga ché.

I by person stolen car

Taiwanese often speak yet other Chinese "dialects" in addition to Mandarin, which means that most of the Taiwanese immigrants we meet speak two Chinese languages, not just dialects of one.

The reason such different-varieties can even begin to be considered "the same language" is because the Chinese writing system uses not letters to represent sounds but symbols to represent whole words. Because the Chinese varieties did all evolve from the same original source, their grammars remain similar enough that they often line up word for word as we have just seen, and this allows the writing system to be suitable for all of the dialects (although because the writing system was developed for Mandarin, there are lacks of fit with the other "dialects").

And then, of course, each of the Chinese "languages" has several dialects, many mutually intelligible only with difficulty. Out in the countryside beyond Beijing, for instance, there are dialects of Mandarin that are as different from the standard as Ukrainian is from Russian and, under other circumstances, could easily be considered "languages" of their own. In Mandarin and other Chinese varieties, single syllables can have different meanings, depending on the tone they are uttered with. In Standard Mandarin, *shu* (pronounced more like "shrew") can mean *uncle* or *book* (among other things), depending on its tone. In the rural dialect of

Wuhan, however, the word for *uncle* is pronounced roughly as “sew” and, as for the word for *book*, shape your mouth to say *Sue* and then say *see*. The tone is no longer the most important distinction between the words, and their shapes are quite different from Standard Mandarin’s.

Most Chinese immigrants to the United States in the 1800s and early 1900s spoke a nonstandard dialect of Cantonese, such as the one spoken in the rural region of Sei-yap. Educated Standard Cantonese-speaking visitors or immigrants today often have some trouble understanding speakers in America descended from earlier waves of immigration. Thus, rather than having “eight dialects,” China actually has several dozen dialects of eight different languages!

### Subject Uncooperative: “Language” and “Dialect” and the Nondiscreteness of Language Change

Finally, we might well propose now that, even if cartographic and cultural labels display only fiftful correspondence with a conception of “language” and “dialect” based on mutual intelligibility, we might still save these useful taxonomic concepts by supposing that human speech varieties *are* distributed in tidy bundles of mutually intelligible dialects, regardless of how geopolitical and cultural boundaries obscure this. In other words, we might suppose that “in real life,” we can just include Macedonian in the “Bulgarian” bundle, ease Moldovan over into the Romanian one, think of Swedish, Norwegian, and Danish as one language, etc., and then everything would be nice and tidy.

But even this doesn’t work: in the proper sense, it’s not only the geopolitically contingent map that’s the problem. Even if the world had not been partitioned into countries, there would still be no intrinsically watertight concept of “language” that would stand up to how human speech varieties are actually distributed on the globe. This follows from the inherently nondiscrete nature of language change, which we already saw producing a “Fratin” phase between Latin and French.

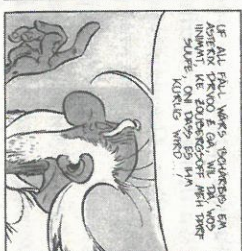
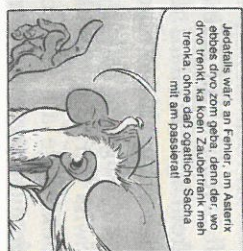
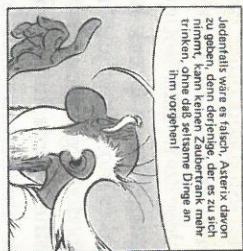
“Where Do You Draw the Line?” Redux: Halfway Between Language and Dialect Because the transformation of a language into a new one is an incremental process, there is a point in this transformation where the new speech variety is clearly akin to its

ancestor and other dialects still close to that ancestor, but only fitfully intelligible with them. A speaker of the ancestor or a dialect close to it does not quite process this new one as “a separate language,” as a Greek has to learn Hungarian, but then acquiring it takes more than just “making some adjustments,” as we would have to do to get along in rural Cornwall of the eighteenth century. In other words, it is common for a speech variety to stand in a relation to another one that is caught between what we intuitively think of as “dialect” and what we intuitively think of as “different language.”

The Schwäbisch that threw me in Konstanz is a useful example. It is just one of several “dialects” of German that are so different from the standard that, even for Standard German speakers, becoming able to function in them is almost a matter of learning a new language rather than adjusting to a variation on their own. Swiss German is another one of these “dialects,” and shown on the next page are identical panels from an adventure of the French comic character Asterix, in its translations into Standard German, Schwäbisch, and Swiss German.<sup>4</sup> Asterix has been clobbered by a scheming village traitor while standing on guard against invading Romans; the “magic brew” is the village druid’s trademark concoction that gives warriors superhuman strength.

Swiss German is particularly instructive. Miraculix’s two words *lage* for “let lie” would be *lassen* and *liegen* in the standard, but *la* is a long way from *lassen*, just as his subsequent *gü* (pronounced “GEH”) is from *geben* and his *ke* is from *kein*. In place of the Standard German version’s *falsch* for Miraculix’s “wrong,” the Swiss version has the local *schähbis* “screwed up,” alien to standard. The standard word for “drink,” *trinken*, is *saufe*, from a root that in Standard German is used only for “guzzle” or in reference to drinking alcohol in hearty fashion. Germans have to, more or less, “learn” Schwäbisch or Swiss German; regarding another non-Standard German variety, I have heard of Standard German speakers having to take classes in the Kölsch

4. I used to be baffled as to why issues of this formulaic series have sold briskly at newsstands across Europe in dozens of languages for decades (yes, the word-play but, really, there isn’t that much of it in any given episode), but by golly there is something about them that grows on you: I swear that I wish I could spend a month living in that village eating wild boars. Asterix does not really work in American English, however—the English translations, done by Englishmen, only begin to work if you imagine them speaking in British accents.



Oh, Miraculix! Do something! Give him some of the elixir that you used against the Romans!

Listen, my good Obelix! I think I left it at the lookout point.

Besides it would be wrong to give Asterix any of it, since everyone who drinks it can't drink any more magic brew without strange things happening to him!

dialect local to Cologne. Although I read Standard German on a regular basis as part of my academic work, when a friend of mine sent me e-mail messages in Swiss German, I found them so opaque that I literally could not grasp even the basic meaning and had to request “translation” into, at least, Standard German if not English.

Still, however, Schwäbisch, Swiss German, and Kölsch are more like other German dialects than like any other languages related to German such as Swedish or English and, after some exposure, one gains a sense of oneself in the “German” orbit in regard to word shapes and grammar. The question, then, is: As a Martian, would you treat Swiss German as a dialect of German or as a separate language if you knew nothing of what speakers call the varieties or where they are spoken? After all, remember that even a Standard German speaker can barely make out anything Schwäbisch, Swiss German, or Kölsch speakers are saying at first.

In making your decision, consider at the same time Spanish and Portuguese. We are accustomed to thinking of them as big, fat, distinct “languages” because they are spoken by formerly geopolitically dominant powers with distinct and rich literary heritages. Yet they are close enough that, if political boundaries had been drawn differently, they would be considered dialects of the “Iberian” language. Spanish and Portuguese speakers can get the gist of each other’s spoken languages (although Portuguese have a much easier time with Spanish than the other way around), and I long ago gave up trying to speak, as opposed to read, Portuguese because I find it impossible to keep it separate from Spanish in my mouth and always ended up committing the gaffe of seeming not to realize that Portuguese is not Spanish (which can be a particularly touchy subject in Brazil, where people are weary of Americans assuming that they speak Spanish as other South Americans do). The Spaniard would say *Ese hombre no tiene mis gatos* for *That man doesn’t have my cats*, whereas the Portuguese would say *Esse homem não tem os meus gatos*—the difference here is obviously quite akin to that between Standard and Swiss German. There are plenty of “languages” in, for example, Africa and Asia, as well as ones in Europe such as Italian, whose “dialects” are even more different.

The intersection between language identification and culture gets even messier in cases where group A considers itself to be speaking the same language as group B, whereas group B considers group A to be speaking a different one. On the island of New Britain just off of Papua New Guinea to the east, there are two varieties called Tourai and Aria that to our eyes and ears would appear to be dialects of the same language. That is also the way speakers of surrounding languages feel—they use the same “language” with speakers of Tourai and Aria. But though the Tourai speakers consider what the Aria speak a different language, the Aria

consider themselves to be speaking Tourai. The situation gets even more complicated because of the mixing of languages, which has created more than one kind of "Aria." The Aria speakers of the Bolo region, close to where another language called Mouk is spoken, have taken on a great deal of Mouk vocabulary. As a result, the other Aria speakers think of the Bolo Aria as speaking Mouk. However, the Mouk speakers see the Bolo Aria speakers as speaking, well, Aria.

Sometimes cultural distinctions rather than geopolitical boundaries or vocabulary mixture end up fencing off closely related varieties as "separate languages." Senegalese people who speak Mandinka are aware that there are "languages" called Bambara and Dyula spoken in nearby regions but also quite readily mention that they can understand both fairly well. On paper, these three "languages" reveal themselves to be about as close as the various German dialects, and in some linguistic descriptions are treated together as variations on one common theme. The speaker of Anyi in Côte d'Ivoire, if asked about other languages he knows, will usually mention that Anyi and Baule are really "the same thing"—about like Standard English and Scots English, as one Ivorian told me—but the distinct cultural heritages of Anyis and Baules conditions a sense that the two speak "different languages."

Scots English is, in fact, one of the only ways English speakers experience a variety of their own language that is so different from the standard that it strains the boundaries of what they consider their own language to be.<sup>5</sup> *Auld lang syne*, for example, is Scots for *old long since*. The words are different enough from Standard English equivalents that we usually sing this phrase as an undigested chunk rather than processing the meaning of each word in sequence, and furthermore, even when we know what they mean, there is the further distancing factor that we do not have a set expression "old long since" for "days of yore." This song is the only way most Americans ever encounter Scots; for a healthier dose, here is a passage from the Prodigal Son parable:

There wis aince a man hed twa sons; and ae day the yung son said til him, "Faither, gie me the faa-share o your haudin at I hae a richt

5. Again, creoles are another example: the "English" of many West Indians sits on the dialect/language line similarly; we will look at these varieties in Chapter 4.

til." Sae the faither haued his haudin atweesh his twa sons. No lang efterhin the yung son niffert the haill o his portion for siller, and fuir awa furth til a faur-aff kintra, whaur he spefelt his siller livin the life o a weirdless waister.

Now, we can follow that pretty well, but between far-out versions of words familiar to us like *aince*, *tea*, *nicht*, and *kintra*, and outright novelties like *atweesh*, *efterhin*, *niffert*, and *spefelt*, this is obviously quite unlike any "English" most of us in America ever hear. Actually, though, political unity with England has gradually brought Scots closer to Standard British English over the centuries. In medieval times, when Scotland was still a separate kingdom, the English dialect of Scotland was well on its way to becoming a separate language, as we see in a snippet from the first fully Scots text, written in 1376:

Thai defendit, and stude tharat,	They defended, and stood
Magré thair fais, gubhill	there,
the nycht	In spite of their foes, until
Gert thame on bath halfis	the night
leif the ficht.	Caused them on both sides
	to stop fighting.

This still looks like a sort of "English," more or less, when you look at it long enough, but differences this vast rendered mutual comprehension a dicey affair at best.

Thus what *are* the Schwäbisch and Swiss German in those *Asterix* panels, all considerations of cartography, history, and cultural identity aside—German "dialects" or separate "languages"? If Portuguese speakers can often get the gist of a Spanish news broadcast, in "God's eyes," are Portuguese and Spanish dialects of the same language? Today, there is an influential movement in Scotland to treat Scots as a separate language from English—well, from a linguist's perspective, which side of the line does Scots fall on? Or if it has been inching toward standard English in the past several centuries, which side of the line did it fall on in the Middle Ages, when it was a little farther from the standard than Swedish is from Danish?

The answer, really, is that there is no way to make the call in cases like these. We saw how close dialects can be compared to

"covers" of an original song. A case like Swiss German brings to mind an episode of *The Simpsons* lampooning *Mary Poppins*, complete with song parodies of "A Spoonful of Sugar," "Feed the Birds," and others. The songs did not use the melodies from the Disney movie, Weird Al Yankovic-style, but were specially crafted with basic shape, rhythm, and harmonic flavor paralleling the originals just enough to instantly recall the songs parodied without sparking a lawsuit.<sup>6</sup> Swiss German stands in a relation to Standard German analogous to that between these song parodies and their models.

If Belorussia and the Ukraine were still regions of Russia instead of separate countries, then Belorussian and Ukrainian would present the same conundrum as Schwäbisch and Swiss German. Ukrainian is definitely not Russian—but then it's more like Russian than like any other language, and enough like it that I could alternately entertain and annoy my college friend by "making up" Ukrainian words based on Russian. Cases like this show that speech varieties differ from one another along a continuum, on which no definite signpost can be placed distinguishing where "dialect" stops and "language" begins.

One "Language" Bleeds into Another "Language" The final reason there is no such thing as a "language" in any intrinsically coherent sense is because, in many cases, one runs into another one just as green is neither yellow nor blue, but a mixture of the two. An example is a group of dialects called Gurage [goo-RAH-gay] spoken in Ethiopia, one of several languages related to Arabic and Hebrew spoken in that country. Does anyone still play that game "Telephone" where people sit in a circle and one person whispers something in the next person's ear, and then that person, having heard something slightly different from what the first person said, whispers what she heard in the next person's ear, and so on, until what comes out on the other end is hopelessly different from the original sentence? A similar phenomenon occurs with what is called a *dialect continuum*. The way to say *He thatched a roof* in Gurage dialects differs slightly from one region to another (the *a* is called *schwa*, the sound of *a* in *about*):

6. Hats off to *The Simpsons*' house composer Hans Zimmer, who also composed what I consider the best theme song in the history of television, for the late, great *The Critic*, luckily still shown on Comedy Central.

Soddo:	kaddanam	Chaha:	khaderam
Gogot:	kaddanam	Gyeto:	khatera
Muher:	khaddanam	Endegen:	hattera
Ezha:	khaddaram		

Soddo and Gogot have the same word; Muher only differs in having an initial *kh* instead of *k*. Ezha makes the small change from here of substituting *r* for *n*; although this is cake for Muher speakers, it's already pretty odd for Soddo speakers. Chaha and Ezha are quite close, but then Gyeto changes the *d* to a *t* and lets go of the final *m*. Neither of these sound-change processes is at all unusual, but to someone who grew up on *kaddanam*, *khatera* is almost incomprehensible at first. Endegen substitutes *h* for *kh*, a natural little jump, and doubles the *t*. Endegen's *hattera* is so different from Soddo's *kaddanam* that Endegen (and Gogot and Muher) speakers process Soddo as essentially a different language. Thus what is "Gurage"? Relationships among many of the varieties are what we think of as "dialectal," but just as many relationships are akin to that between Spanish and Italian. Gurage is neither a bundle of "dialects" nor a bundle of "languages"—it is a conglomeration of varieties related to one another to various degrees.

This is not rare; linguists encounter dialect continua all over the world, often linking "what are conventionally known as separate "languages." In the Central African Republic and the former Zaire, from region to region there are various "languages" that differ from one another only to the extent that we would imagine of "dialects." Here is how to say *Me, I'm going to the village to build a house* in seventeen of these languages:

Bobangi:	Ngai, nakoke o mboka notonga ndako.
Nunu:	Ngai, namoke o mboka notonga ndako.
Libinza:	Ngai, nakakende o mboka nakatonga ndako.
Lusakani:	Ngai, namoke o mboka notonga ndako.
Mpama:	Ngai, nakei mboka nakatonga ndako.
Liboko:	Ngai, nakei o mboka nakatonga ndako.
Loyi:	Ngai, nakei mboka natonga ndako.
Impfondo:	Ngai, nakei o mboka mpfoa ya itonga ndako.

Enyele:	Nga, nakei mboka botonga ndako.
Bomitaba:	Nga, nakei mboka eke otonga ndako.
Likuaba:	Ngai, nasoke mboka otonga ndako.
Likuala:	Nga, nake o mbowa notonga ndako.
Moyi:	Ngai, nakeke o mboka notonga ndako.
Mboshi:	Nga, izwa mboa otonga ndai (ndao)
Koyu:	Nga, lizwa mbooka etonga ndako.
Makua:	Nga, ikendi mboga etonga ndago.
Bongili:	Ngai, nake mboka na kotonga ndako.

One could make a similar list of identical sentences in the various German "dialects," and they would often differ more than these "languages" differ from one another.

Turkish is one of several Turkic languages, many of which are highly similar to one another, such that, in many parts of the Turkic-speaking region (including many of new "stans" freed from the former Soviet Union), one "language" bleeds into another one through intermediate dialects. Here, for example, is *eight* in seven of these languages going from west to east:

Turkish:	sekiz	Kazakh:	segiz
Azerbaijani:	sakkiz	Kirghiz:	segiz
Turkmen:	sekiz	Uighur:	säkikiz (here, ä = the <i>a</i> in <i>cañ</i> )
Uzbek:	sakkiz		

You couldn't pay the Romance languages to match up that nicely. Even though their relationship is clear on all levels, here's what happens to the word *eight* in seven of them:

French:	huit	Romanian:	opt
Spanish:	ocho	Occitan:	uech <sup>7</sup>
Italian:	otto	Catalan:	vuit
Portuguese:	oito		

7. I like that one, too.

There is not a continuum of mutually intelligible dialects across all seven of the Turkic languages; there is one between Turkish and Azerbaijani, as well as between some other pairs. In general, however, all of these "languages" are closer than are all of the "dialects" in many "languages." Turkey, Turkmenistan, Kirghyzstan, Uzbekistan, Kazakhstan, and Azerbaijan are often referred to, as I did earlier, as united by being "Turkic speaking," with an implication that they all speak in some sense one "Turkic" language, even if that language is not Turkish proper. The roots of this concept of a "Turkic" hovering somewhere between "language" and "dialect" lie in the fact that a general "Turkic" system varies incrementally from one region to the next, confounding any attempt to apply taxonomic labels in any consistent way.

Predictably, Serbs and Croats have been known to treat "Serbian" and "Croatian" as different languages and even often claim to have difficulty understanding one another. As with the Romanian-Moldovan case, writing lends an artificial sense of distinction: Serbian is written in Cyrillic, whereas Croatian is written in the Roman alphabet. There are also, as always, some differences in vocabulary. Yet traveling from humble hamlet to humble hamlet across the former Yugoslavia, apart from the artificial division created by writing and cultural conflict, the linguist encounters a continuum of dialects changing Gurage-style from village to village. Among immigrants from the former Yugoslavia today it is quite common to see couples, one member Serbian and the other Croatian, conversing easily in a single language, Serbo-Croatian. Culture and politics make the call between dialect and language here and have continued to do so—after the Dayton Accords, a dictionary of the "Bosnian" language was published.

In larger view, this particular continuum encompasses "languages" even beyond Serbo-Croatian. Not only can Bulgarians understand Macedonians next door, but Macedonians on the border with Yugoslavia can communicate with Serbo-Croatian speakers on the other side, such that Serbo-Croatian, Macedonian, and Bulgarian form a grand continuum. Standard Serbo-Croatian and Bulgarian are as different as Spanish and Italian but are linked by a procession of dialects—and even a whole "language"—falling on a continuum linking them in a kind of living exhibit of one morphing into another in space just as languages morph into one another in time.



There are cases in this region where not only the sound of a word but its very meaning changes incrementally as well.

Dalmatian coast:	vridan	"industrious"
Bosnia, Montenegro:	vrijedan	"industrious"
Serbia:	vredan	"industrious"
Macedonia,		
Western Bulgaria:	vredan	"industrious" or "harmful"
Southeastern Bulgaria:	vredan	"harmful"

The question arises, then: If all of these dialects were spoken in some uncharted region rather than artificially corralled into "countries," where would you draw the line between one "language" and another one?

**Intelligibility: Taxonomic Quicksand** One manifestation of this "neither fish nor fowl" aspect of many varieties is that dialects that are in essence extremely close can still be just barely mutually intelligible. This results from small but sharp differences in the sound systems and from differences in the semantic evolution of vocabulary. For example, though Danes, Swedes, and Norwegians can converse, intelligibility is not all peaches and cream. Just as among Romance languages French has transformed the original material more than most of the others (from Latin *cantare* "to sing," Italian still has *cantare*, Spanish has *cantar*, but French has *chanter* [shaw"=TAY]), Danish is the "advanced" one in Scandinavian. For "to play," Swedish has *leka*, and Norwegian has *leke*, both pronounced approximately "LEH-kuh," but Danish has *legge*, whose archaic spelling masks that it is pronounced "LJE-uh." This means that, for Norwegians and Swedes, getting just what words a Dane is saying is a bit of a strain.

On the other hand, between Norwegians and Swedes, the similar sound systems make understanding what words are being used unproblematic, but problems arise because of different meanings of the same root: Norwegian *rolig* is "calm," but in Swedish the same root has drifted to mean "funny," just as *silly* drifted in English from "blessed" to "idiotic." *Dyrke* is "to cultivate" in Norwegian and "to worship" in Swedish; Norwegian's *blot* is "soft," Swedish's *blot* is "wet"; Norwegian's *tilbud* is "offer," Swedish's *tilbud* is "accident." Because Norwegian was still "Danish" more recently than Swedish was, Danish tends to have the same meanings as Norwegian, and

thus Swedes and Danes have the same problem. Because Danish is the odd one out in regard to sound system, whereas Swedish has gone its own way a bit in regard to word meanings, it has been said that "Norwegian is Danish spoken in Swedish"—that is, Norwegian, which parallels Danish's word meanings, is how Danish would come out if its sound system weren't so independently minded and were therefore more like the Swedish one.

This kind of ambiguous degree of intelligibility exists worldwide between speech varieties that look highly "close" on the page. Orungen and Evenki are two closely related—well, linguists don't know what to even begin to call them—that straddle a border between Russia and northwestern China. Line the two up on the page and they look as close as the Turkic varieties listed earlier. But in real life on the ground, the intelligibility matter is tricky. In line with what we would expect from what they look like in print, Orungen often claim to be able to speak with Evenkis—but then have been shown to not be able to understand a tape played of Evenki being spoken. Much of the problem appears to be that, in Orungen, accent always falls on the last syllable of the word, whereas in Evenki it can fall in various places as in English. Evenki has *ollo* for fish, Orungen has *olo*, and so on. When this difference is applied to every word in the language, the cumulative effect is considerable—imagine if we had to communicate with someone who, in addition to having what we processed as a thick accent in general, said things like "Yenn this DIFFRENS iss IBLIED to avREE wirt in dah langWIDGE, dah camalaTIFF EFFEct iss cansaderabULL." This would still be "English"—it sure isn't German; but you'd almost wish it *were* German so that you could claim not to speak the person's language!

Thus even the intelligibility issue is messy: any metric of intelligibility one tried to fashion would trip up on the fact that intelligibility and taxonomic closeness do not walk in anything approaching a lockstep. Certainly dialects that are not close on the page will also *not* be mutually intelligible—but then, when they *are* close on the page, they may or may not be mutually intelligible.

Linguists are often asked, "What's the difference between language and dialect?" Often, the answer they give is that it is a "difficult question," which depends as much on culture, history, and politics as on linguistic reality. This response, however, refers only to *language* and *dialect* as labels. The linguistic reality does not lend itself gracefully to *any* underlying conception of a language/dialect distinction. The geopolitical and cultural factors only make clearer

a problem that would exist even if there were no such subdivisions and humans simply coated the earth in little hunter-gatherer bands as they once did in Paleolithic times. Properly, the language/dialect distinction is, in the pure logical sense, meaningless.

Certainly there are "languages" with only one dialect, such as many languages spoken by only a few hundred hunter-gatherers in places like Papua New Guinea. And certainly there are bundles of closely related dialects where none stray particularly far from the basic template. Korean has its dialects (predictably, for example, North Koreans speak markedly differently from South Koreans), but all are readily mutually intelligible. Because Modern Hebrew is spoken in a tiny country, and only has been so for less than a century, there are no spoken dialects of it that strain an Israeli's sense of what "Hebrew" is. But the crucial point is that this is by no means the "default" situation: on the contrary, if anything, these situations are the exceptions, more typical of smaller groups of speakers. Typically, what looks from the air like "a language" is actually a much hazier business on the ground. Korean is relatively uniform, but Japanese speakers in Tokyo can barely follow speakers from the Ryukyu Islands, and on the page the Ryukyu dialect is about as far from Standard Japanese as Schwäbisch is from Standard German. Hebrew is pretty tidy, but its neighbor and relative Arabic differs so much from country to country that the various "dialects" differ about as much as Spanish, Italian, and Portuguese do from one another. Thus though there are cases where speech varieties *happen* to fit into an idealized language/dialect template, they usually do not. As such, properly speaking, there are no "languages."

If it is possible to save any remnant of our terminology, the best we can say is that there are innumerable *dialects* in the world, related to each other to various degrees, sometimes clumping into complexes particularly close to one another, but generally not so close that all are mutually intelligible, with distances often so great between some of them that their speakers do not consider themselves to be speaking "the same thing" in any sense.

It's all about dialects, then: language change has split the first language into tens of thousands of dialects that we arbitrarily group into "languages" according to approximate notions of intelligibility and the dictates of the cultural and political developments of the moment. Dialects are everywhere—and always have been: even Old English had them. I oversimplified a bit in depicting the

Angles, Saxons, and Jutes as speaking one original "language" when they invaded England. Because these peoples had lived in separate places in each of which the unwritten West Germanic ancestor of English had developed in slightly different directions, they spoke at least three dialects of Old English, as we see from three renditions of the first line of the Lord's Prayer:

West Saxon:            Fæder ðre, thū eart on heofonnum

Northumbrian:        Fæder ðrer, thū art on heofonu

Mercian:                Fæder ðre, thū eart on heofennum

Thus the dazzling variety among British dialects stemmed from at least three slightly variant founding *dialects*, not a single variety.

We would further predict that there would even be dialects of dialects, and this is exactly what happens, which in turn highlights once more that the original language has developed not into just six thousand more "languages" but, more properly, into tens of thousands of variations on variations corresponding to the tens of thousands of speech communities on the earth, obviously vastly outnumbering the mere six thousand "languages" that we can approximately delineate. When I mentioned the Bavarian German translations of *Asterix* to a Bavarian in Germany, her first response was a good-natured complaint that they had not translated into *her* dialect of Bavarian, and she proceeded to give me some of the differences between her speech and that depicted in the books. A few weeks later in America I mentioned the Swiss German translations to a Swiss person, and he immediately said, "Well, the one problem is that of course they didn't translate it into *my* dialect!" The word for "messed up" in the Swiss German *Asterix* excerpt, *tschärbi*, is from only one dialect of Swiss German, not used by all German-speaking Swiss. Dialects is all there is.

## Two Tongues in One Mouth

There is a nuance to be added to our developing picture of human speech across the globe. Just as many people in the world are bilingual in two or more "languages," a great many people control more than one dialect of a language. In particular, it is common for

members of a community or society to speak both a standard dialect and a nonstandard one, especially today with the spread of education and the centralization of economics lending increasing numbers of people more contact with the standard dialect than was the case in earlier periods of history.

The classic example is Swiss German. For German-speaking Swiss, Swiss German is the language of the home, the language learned first, the language of the casual, the familiar, the intimate—and for *all* Swiss rich and poor; Swiss German is not a class issue the way, say, Appalachian or Black English partly are. Standard German, however, is the language of writing, official announcements, and all scholastic endeavor. All students are taught in Standard German, its grammar is the one taught in the schools, and it is used almost exclusively of Swiss German on the radio and on television. Thus all German-speaking Swiss by the time they are mature speak and understand both Standard German and Swiss German, thinking of them as variants on a common theme, each with their particular sphere of appropriateness. My expounding about Swiss German, dutifully giving samples from it on paper, contrasts with its actual “place” in its speakers’ consciousness, which is as an integral but strictly informal part of life, usually only seen in writing in personal messages in local newspapers or in personal letters. At the Frankfurt airport I saw a man, apparently Swiss, chuckling and hooting while reading an *Asterix* edition in Swiss German, it being funny to him to see a book written in the dialect.

There are similar situations throughout the world, and they are called *diglossia*, from the Greek for “two tongues.” The Arabic an Egyptian of any class speaks at home is actually a different language from the Modern Standard Arabic used in writing and in scholarly instruction. Just as the Swiss German speaker has *suufe* at home and *trinken* in print, the Egyptian refers to his nose as a *manaxir* but would write *ʿanfun* in print. This is also true of different non-Standard Arabics spoken in Morocco, Algeria, Nigeria, Sudan, Syria, Lebanon, Saudi Arabia, etc. I once heard an educated Moroccan journalist describe his childhood saying casually that he had spoken “Moroccan” at home and then learned “some Arabic” in school, neatly demonstrating that, in the mind of a Moroccan Arabic speaker, Modern Standard Arabic is not just a hoity-toity way of speaking what he learned at his mother’s knee, but essentially a different language that must be carefully taught. I also saw three Finns

unable to agree on just how to say “Hey, look—a shortcut!” in large part because of dialect differences; one of them came from a region where the local dialect is different enough to inspire affectionate, locally produced jokebooks just as Schwäbisch and other dialects do in Germany. Yet all of them were fully functional in Standard Finnish as well. The standard dialect of Indonesian used in writing and taught in books is a scholarly creation designed to parallel Western languages in its grammar as much as possible; in their everyday lives, Indonesian speakers use an array of nonstandard varieties that are often only fitfully intelligible with the standard one. The layers of language in Javanese that we saw on page 50 are another example of diglossia (although in that case there are actually “middle class” forms as well, such that we are really dealing with triglossia, something else not unheard of worldwide).

Diglossia is a manifestation of a hierarchy of social domain that all human speech varieties observe to some extent; diglossia goes to the extreme of dividing the labor between two distinct grammars, but in many languages the same kind of distinction is indexed through vocabulary alternates and various set expressions. There is a song that longtime residents of Oakland, California, sometimes sing dedicated to the city, and one of the lines goes, “Where did all the people go when Frisco burned?/They all came here to Oakland and they never returned!” That line has always struck me, overthinking such things as I tend to, as a bit “off” in tone. It’s that word *returned*. The song has a red-blooded, rah-rah feel that leads you to expect it to end with the spelling out of O-A-K-L-A-N-D and a rousing “Whoop!” or the like (although it doesn’t, actually) and, as such, *returned* is too formal. We say *went back* or *came back* in casual English; *returned* is for writing and formal situations. Imagine asking your significant other, as you peel the potatoes, “When are you returning tonight?”—you’d either be (1) new to English, (2) striking an irritated or ironic tone, or (3) very, very strange. Our diglossia splits between *come back* and *return*, or *check out* and *examine*, or *kids* and *children*.

We can see an illustration of how diglossia plays out when distributed across two dialects with one more look at our friend Asterix, this time from the Bavarian German edition (see page 90). Asterix and Obelix, mistaken as Goths the Romans are chasing (don’t ask), are disguised as Romans to throw them off the scent, and Asterix instructs Obelix on how to greet any Romans they



horns); a *Brontosaurus*,<sup>9</sup> and so on. But if you really get into dinosaurs you see that each one of these standard *Flintstones* dinosaur types was actually one of an almost numbing array of variations on a theme. There were lots of kinds of stegosaurus—this one had spikes down half its back instead of plates, that one had smaller plates overall, etc.; there were lots of horned dinosaurs, one kind with one horn on its nose, another kind with a bump on its nose, still others with horns all around the frill; *Tyrannosaurus* was one of a couple of dozen similar but slightly variant creatures discovered worldwide; there were runty-sized brontosaurus types, and some of them had armor in their skin; etc.

This is what “languages” are like. The dinosaur parallel goes even farther in that similar dinosaur types often lived together in the same areas: the carnosaur *Allosaurus* shared its environment with the smaller *Ceratosaurus*, which had a little horn on its nose; the duckbill *Corythosaurus* with the helmet decoration on its head and *Parasaurolophus* (the one with the marvelous curved tube flying backward off the top of its head) lived in the same places as well. Evolution produced not only “types” but subtypes of dinosaur (and even these subtypes had differing species within them); language change has developed “subtypes” of language, not just occasionally—as we are sometimes misled into thinking by statements such as “Italian has a lot of dialects”—but almost always. Dialects are the norm; dialects are what happened under typical conditions. It’s all about dialects. For our purposes, forget “languages”!

So far, we have seen how the first language developed into turtles and cats, and then we looked a little closer and saw how it developed into snapper turtles and sea turtles; Burmese cats and Siamese cats. Now we will take the next step and see how the first language also developed quite often into mules and, well, my cat.

### 3

## The Thousands of Dialects Mix with One Another

So far, in describing how the first language split into thousands of subvarieties, I have implied that speech varieties have developed like a bush, starting from a single sprout and branching in all directions, each branch then developing subbranches, and so on, culminating in a dense web of a plant whose outer layer is crowned with thousands of leaves, symbolizing languages (or dialects). Allowing that the bush analogy cannot capture the fluid nature of the degree of relatedness between dialects, we could think of leaves lining the same twig as closely related varieties such as the German dialects, the leaves over on the next branch as the English dialects, and branches way over on the other side of the bush as, perhaps, the languages of Polynesia.

Indeed, linguists who study language change and the family relationships between languages have traditionally taken the “family tree” model as central in how language has developed. Yet particularly in the past twenty years, language change and classification specialists have come to realize that this model actually only takes us so far in describing the reality of what has really happened to the first language as it has spread across the planet.

Just as it is inherent to languages to change gradually into new ones, it is equally inherent for them to mix with one another. Moreover, just as language change is an unbroken process along which no lines of demarcation can be drawn, language mixture is along a continuum of degree. Viewed close up, not only are “languages” clusters of dialects but, in the past 150,000 years, the dialects of

9. Dino fans: Yes, I know, but really, how many other people know what an *Apatosaurus* is?

Also by John McWhorter

*The Missing Spanish Creoles: Recovering the Birth of  
Plantation Contact Languages*

*Spreading the Word: Language and Dialect in America*

*Towards a New Model of Creole Genesis*

*Word on the Street: Debunking the Myth of a "Pure"  
Standard English*

*Losing the Race: Self-Sabotage in Black America*

# The Power of Babel

A NATURAL HISTORY  
of Language

John H. McWhorter



Perennial

An Imprint of HarperCollins Publishers