

OTHER BOOKS BY STEVEN PINKER

Language Learnability and Language Development

Learnability and Cognition

The Language Instinct

How the Mind Works

WORDS AND RULES

*The Ingredients
of Language*

Steven Pinker

Weidenfeld & Nicolson
LONDON

DISSECTION BY LINGUISTICS

Regular and irregular words have long served as metaphors for the law-abiding and the quirky. Psychology textbooks point to children's errors like *breaked* and *goed* as evidence that we are a pattern-loving, exception-hating species, explaining everything from why children have trouble learning simple laws of physics to why adults make errors when using computers or diagnosing diseases. In 1984 George Orwell has the state banning irregular verbs as a sign of its determination to crush the human spirit; in 1989 the writer of a personal ad in the *New York Review of Books* asked, "Are you an irregular verb?" as a sign of her determination to exalt it.

Science is not always kind to folklore from the natural world. Elephants do forget, lemmings don't commit mass suicide, two snowflakes can be alike, we use more than 5 percent of our brains, and Eskimos don't have a hundred words for snow. We had better give irregular and regular verbs a closer look before using them as evidence for a language faculty that works by words and rules, or more generally, a mind that works by lookup and computation.

Regular and irregular forms do not work in isolation; they are part of the integrated living system we call a language. This chapter will tease out regular inflection from the linguistic organs and tissues in which it is embedded. The next chapter, on irregular verbs, will have a different feel. Living creatures can be dissected, but creatures dead so long that only a trace of the living organs

remain must be excavated. Our tour of the irregular verbs will uncover them from layers of historical sediment laid down over thousands of years.

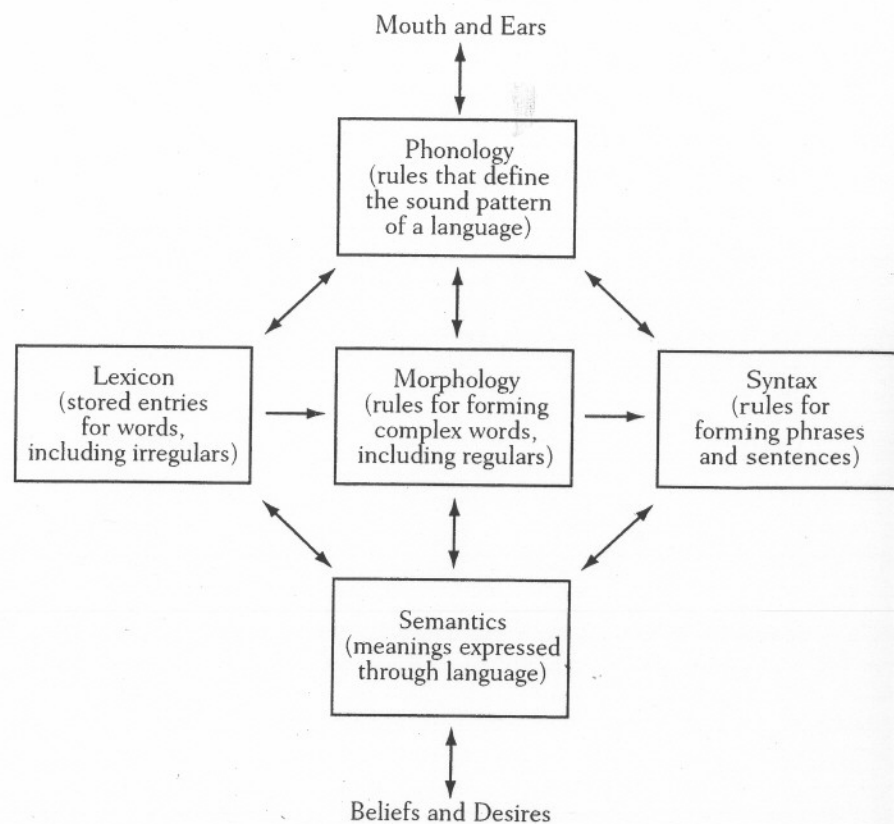
Does language even *have* an anatomy? Many people think about language in the following way: We need to communicate, and language is the fulfillment of that need. For every idea there is a word and vice-versa, and we utter the words in an order that reflects the connections among ideas. If this common-sense view is true, there would be little need to speak of language being a complex system. The complexity would reside in the meanings, and language would reflect that complexity directly.

The point of this chapter is to show that this view is mistaken. I will put regular verbs under a microscope to reveal the delicate anatomy that makes them work. Language does express meaning as sound, of course, but not in a single step. Sentences are put together on an assembly line composed of mental modules, shown on the following page. One is a storehouse of memorized words, the mental lexicon. Another is a team of rules that combine words and parts of words into bigger words, a component called *morphology*. A third is a team of rules that combine words into phrases and sentences, a component called *syntax*. The three components pass messages about meaning back and forth with the rest of the mind so that the words correspond to what the speaker wants to say. This interface between language and mind is called *semantics*. Finally, the assembled words, phrases, and sentences are massaged by a set of rules into a sound pattern that we can pronounce when speaking or extract from the stream of noise when listening. This interface between language and the mouth and ear is called *phonology*.

Many people are suspicious of box-and-arrow diagrams of the mind. The walls of the boxes and the paths of the arrows often seem arbitrary, and could just as easily have been drawn differently. In the case of language, however, these components pop out as we tease apart the phenomena, and at least some of the divisions are now becoming visible in the living brain, as we will see in chapter 9.¹ This chapter will explore the kinds of discoveries that have led linguists to divide language into parts, using only the facts of regular and irregular words. First, we will see why the lexicon is different from the two boxes of rules to the right, then why morphology is in a different box from syntax, and finally, why phonology and semantics each gets a box.



The easiest boxes to keep separate ought to be the boxes containing words and rules. From the discussion in the preceding chapter, it should be clear that a



simple word like *duck* belongs in the lexicon to the left in the diagram. Just as clearly, a sentence like *Daffy is a duck* is assembled by the rules of syntax in the box on the right. According to the words-and-rules theory, irregular forms such as *swam* are also words that come from the lexicon, because they are as arbitrary as *duck*. What do we do then with regular forms like *quacked*? They look like words and sound like words, but I have been insisting they don't have to be stored in the lexicon. They don't seem like words, but they don't seem like sentences either, which are the clearest products of rules.

The problem is that the terms *word* and *rule* come from everyday parlance and are as scientifically fuzzy as other vernacular terms, like *bug* and *rock*. On closer examination, the word *word* has two very different senses.² The first sense matches the everyday notion of a word: a stretch of sound that expresses a concept, that is printed as a string of letters between white spaces, and that may be combined with other words to form phrases and sentences. Some of these words are stored whole in the lexicon, like *duck* and *swam*; others are assembled out of

smaller bits by rules of morphology such as *quacked* and *duck-billed platypus*. A technical term for a word in this sense is a *morphological object*, to be distinguished from phrases and sentences, which are syntactic objects.

The second sense of *word* is a stretch of sound that has to be memorized because it cannot be generated by rules. Some memorized chunks are smaller than a word in the first sense, such as prefixes like *un-* and *re-* and suffixes like *-able* and *-ed*. Others are larger than a word in the first sense, such as idioms, clichés, and collocations. Idioms are phrases whose meanings cannot be computed out of their parts, such as *eat your heart out* and *beat around the bush*. Collocations and clichés are strings of words that are remembered as wholes and often used together, such as *gone with the wind* or *like two peas in a pod*. People know tens of thousands of these expressions; the linguist Ray Jackendoff refers to them as “the Wheel of Fortune lexicon,” after the game show in which contestants guess a familiar expression from a few fragments. A chunk of any size that has to be memorized—prefix, suffix, whole word, idiom, collocation—is the second sense of *word*. It is the sense of *word* that contrasts with *rule*, and the sense I had in mind when choosing the title of this book. A memorized chunk is sometimes called a *listeme*, that is, an item that has to be memorized as part of a list; one could argue that this book ought to have been called *Listemes and Rules*.

So *walked* is a word in the first sense (a morphological object) and not a word in the second sense (a listeme); its listemes are *walk* and *-ed*. These one-part listemes—prefixes, suffixes, and the stems they attach to, such as *walk*—are called morphemes, a term coined by the nineteenth-century linguist Baudouin de Courtenay to refer to “that part of a word which is endowed with psychological autonomy and is for the very same reasons not further divisible.”³



What about the rules? Why divide the rules of morphology, which build complex words (including regular plurals and past-tense forms), from the rules of syntax, which build phrases and sentences? Both are productive, recursive, combinatorial systems, and some linguists see them as two parts of a larger system.⁴ Yet all linguists recognize that they are not identical. This may seem of no interest to anyone but a student cramming for a Linguistics 101 final, but in fact it has been a source of countless barroom arguments, late-night dorm-room debates, and irreconcilable differences.

What is the correct word for people who pass by: *passerbys* or *passersby*? Do nervous fiancées dread the first meeting of the *mother-in-laws* or the *mothers-in-law*? Who did Richard Nixon force to resign: a series of *Attorney Generals*, or a series of *Attorneys General*? Here are a few real-life examples:

Dear Ms. Grammar,

A member of the Friday Night Couples League . . . had a *hole in one* on the third hole and another on the fifth. Did he have two *holes in one* or two *hole in ones*? One of us believes that the pattern should be the same as in *attorneys general* and *passersby*. The other disagrees, believing that *holes in one* would indicate that the golfer gained multiple holes in one shot. A Diet Coke has been wagered on this, and we have agreed that Ms. Grammar shall be the final authority.⁵

SPOONFULS

From a recipe: “Now throw in two tablespoons full of chopped parsley and cook ten minutes more. The quail ought to be tender by then.” Never mind the quail; how are we ever going to get those tablespoons tender? The word, of course, is *tablespoonfuls*, no matter how illogical it seems. One dictionary contains the entry *spoonsful*, but this is not generally accepted.⁶

Gin and tonic season (no hyphens, please) is just about finished, but Joe Galeota of West Roxbury would still like to know how to order when he’s having more than one. “Friends advised me that the answer is ‘gins and tonic’ because alcohol is the main ingredient,” he writes.⁷

Never has the U.S. faced a worse crisis than in 1887, after the invention of the Jack-in-the-Box. It had become a fad overnight, and everyone was having a whale of a time when someone asked, “What is its plural?” “Jack-in-the-Boxes!” claimed some. Others hotly insisted, “Jacks-in-the-Box!” Civil war seemed inevitable, when Zeke Kelp’s Crusade won a compromise on “Jacks-in-the-Boxes.” Unthanked for forty-three years, Kelp will be honored next week when N. Y. City unveils a hydrant in his name.⁸

All right, the last example isn’t from real life; it’s from the *Early Cartoons and Writing of Dr. Seuss*. The others are from well-known language columnists. *Hole-in-one* is from Ms. Grammar, the nom de plume of Barbara Walraff when presiding over “Word Court” in the *Atlantic Monthly*. *Spoonful* is from Theodore Bernstein, the late *New York Times* editor who wrote the syndicated

column “Bernstein on Words.” *Gin and tonic* is from Jan Freeman, who dispenses “The Word” in the *Boston Globe*.

People disagree on how to pluralize nouns, and they care about who is correct. Purists insist that the *-s* belongs on the noun in the middle of the expression (*notaries public*, *runners-up*), and those with the common touch are content to leave it at the end (*notary publics*, *runner-ups*). “Ms. Grammar” advised her beseechers that *holes in one* is technically correct, but added, “to say ‘two holes in one’ is to ask to be misunderstood.” Her Solomonic suggestion was to say *a hole in one twice*, and to buy two Diet Cokes.

For my purpose—figuring out how the human mind deals with language—there is no correct answer. Most disputes about “correct” usage are questions of custom and authority rather than grammatical logic (see “The Language Mavens” in my book *The Language Instinct*), and in these disputes in particular, both parties have grammatical logic on their side. Their agony highlights the distinctions among lexicon, morphology, and syntax, and illustrates the theme of this book: that the mind analyzes every stretch of language as some mixture of memorized chunks and rule-governed assemblies. How people pluralize an expression depends on how they tacitly analyze it: as a word or as a phrase.

With a simple word the plural suffix goes at the end: one *girl*, two *girls*. Now what happens in a compound word composed of two simple words, such as *cowgirl*? The plural still goes on the end: *two cowgirls*, not *two cowgirl* or *two cows-girls*. That is because the word *girl* inside *cowgirl* is special. It is called the *head* of the word, and it stands for the word as a whole in determining its meaning (a *cowgirl* is a kind of *girl*) and in determining its plural: The *-s* goes on *girl*. A *phrase* also has a head, and it too determines the meaning and gets the plural. But now we discover the major difference between a word, the product of morphology, and a phrase, the product of syntax: In the phrase, the head is on the *left*, not the right. If you meet more than one *girl* from Ipanema (head = *girl*), they are *girls from Ipanema*, not *girl from Ipanemas*. With a word the plural is on the end (*cowgirls*); with a phrase the plural can be in the middle (*girls from Ipanema*).⁹

The seeds of the *mother-in-law* dispute were sown by a special option of English: Occasionally a phrase gets repackaged into a long word. For example, a hangover victim may complain of a *bottom-of-the-birdcage taste* in her mouth; the phrase *bottom of the birdcage* has been packaged as a word that modifies *taste*. When a word-made-from-a-phrase is new and fresh, speakers still can perceive the anatomy of the phrase inside the word. For example, we parse the modifier *bottom-of-the-birdcage* to understand that it means something as foul as the bottom of a birdcage.

But when the phrase is used as a word repeatedly, the original meaning can recede from collective memory. The phrase boundaries melt into a glob, and speakers no longer sense its parts. No one thinks of *Thursday* as *Thor's Day* anymore, or of *breakfast* as *breaking a fast*. Modern English has thousands of former phrases and complex words that have congealed into what people now perceive as simple words, such as *business* (*busyness*), *Christmas* (*Christ's Mass*), and *spinster* (*one who spins*). The meltdown, of course, does not happen overnight or in all speakers at once; there must have been a time when some English speakers still heard *Christmas* as *Christ's Mass* and others heard it as the arbitrary name of the holiday, just as today's older speakers hear the *awe* in *awesome* where younger speakers hear the whole word as a synonym for *good*.

Most of our disputed plurals originated as phrases and then became words. Long ago people might have thought, “she is not my *mother in reality*; she is only my *mother in law*” (that is, according to canon or Church law). But the concept of a spouse's mother needs a word, and eventually the phrase got re-analyzed as that word: “She is my *mother-in-law*.” Similar meltdowns occurred in these phrases:

Jack is in the box → That is a *Jack-in-the-box*.

Phyllis completed that hole in one shot → She got a *hole-in-one*.

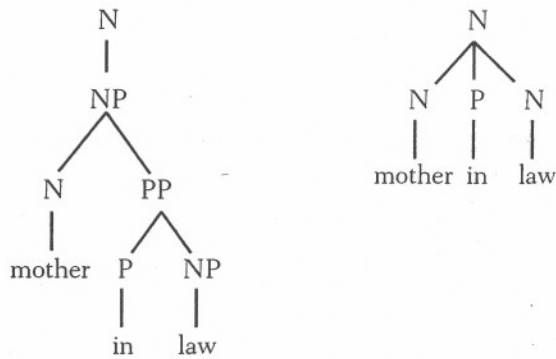
Barry passed by → He is a *passerby*.

I set aside a spoon full of parsley → I set aside a *spoonful*.

If some speakers still hear the phrase inside the word, they will be tempted to put the plural marker on the head of the phrase: *two mother + s in law*, *Jack + s in a box*, *hole + s in one*, *passer + s by*, *spoon + s full*. But if speakers glom the words together in their minds, they will be tempted to put the plural marker at the end: *motherinlaw + s*, *jackinthebox + es*, *passerby + s*, *holeinone + s*, *spoonful + s*.

It's not that phrase hearers interpret these expressions literally (for example, that a *mother-in-law* is a mother as recognized by the law), or that the phrase-deaf treat them as any old string of consonants and vowels; both surely recognize them as complex words built out of familiar words. It's just that they grow different kinds of connective tissue when piecing these expressions together. Those who would describe themselves as *sons-in-law* hear *mother* as the head of a phrase inside the word (shown in the left tree in the diagram); those who

would describe themselves as *son-in-laws* hear a string of little words inside the big word (right tree):



A proof that the *in-law* expressions have congealed into words may be found in the umbrella word *in-law*, which can stand alone and be pluralized in the usual way: *The in-laws are coming over*. It is a good bet that many of today's commonly used phrases will also become opaque some day and turn into words; the giveaway will be a plural at the end. Don't be surprised if one day you hear about *grant-in-aids*, *bill of ladings*, or *work of arts*.

This ambiguity—one stretch of sound, two ways of building a tree in the mind—also started the controversy raised by reports such as the following:

While Mo Vaughn should finish well over .300 with close to 40 home runs and more than 100 RBIs, Mike Piazza has not been producing anywhere close to what he did last season, when he hit .362 with 40 homers and 124 RBIs.¹⁰

Baseball purists who deplore artificial turf and the designated hitter get equally incensed by the plural form *RBIs*. *RBI* is an acronym for *run batted in*, a run scored by a teammate as a consequence of one's batting the ball. An *RBI* and then another *RBI* are two runs batted in, and the acronym for *runs batted in* is just *RBI*—so it should be *124 RBI*, not *124 RBIs*. (The purists are not mollified by the sportscasters' common alternative, *ribbies*.) But the purists fail to recognize that acronyms, like phrases, can turn into bona fide words as a language evolves, as in *TV*, *VCR*, *UFO*, *SOB*, and *PC*. Once an acronym has become a word there is no reason not to treat it as a word, including adding a plural suffix to it. Would anyone really talk about *three JP* (justices of the peace), *five POW* (prisoners of war), or nine *SOB* (sons of bitches)?

An additional puzzle surrounds *governors-general*, *solicitors-general*, and *attorneys-general*. The speakers who bequeathed the plurals to us must have

analyzed the words as phrases, which have their heads on the left. Indeed, a governor-general is a general governor, namely, one who has several governors under him. The puzzle is, why didn't they simply call him a *general governor*? After all, the adjective comes before the head noun in English, not after it. The answer is that these words, together with many other terms related to government, were borrowed from French when England was ruled by the Normans in the centuries after the invasion of William the Conqueror in 1066. In French, the adjective can come *after* the head noun, as in *États-Unis* (United States) and *chaise longue* (long chair, garbled into the English *chaise lounge*). The earliest citation in the *Oxford English Dictionary* is from 1292: "Tous attorneyz general purrount lever fins et cirrographer" (All general attorneys may levy fines and make legal documents). Anyone who insists that we eternally analyze (hence pluralize) these words as they were analyzed in the minds of the original speakers of Norman French also should insist that we refer to more than one major general as *majors general*, because a *major-general* was once a general major (from the French *major-général*). Long ago our linguistic foreparents forgot the French connection and reanalyzed *general* from a modifying adjective to a modified noun.

So if you are ever challenged for saying *attorney-generals*, *mother-in-laws*, *passerbys*, *RBIs*, or *hole-in-ones*, you can reply, "They are the very model of the modern *major general*." They come from reanalyzing a phrase into a word, a common development in the history of English, and a nice demonstration that we treat stretches of language not as sounds linked directly to meanings but as structured trees. People who put different trees on the same sound will use the sound in different ways, even if the meaning is the same.



Let's now peer into the morphology box. Morphology may be divided into *derivation*—rules that form a new word out of old words, like *duckfeathers* and *unkissable*—and *inflection*—rules that modify a word to fit its role in a sentence, what language teachers call conjugation and declension. The past tense and plural forms are examples of inflection.

English inflection is famous among linguists for being so boring. Other languages exploit the combinatorial power of grammar to generate impressive numbers of forms for each noun and verb. The verb in Spanish or Italian comes in about fifty forms: first, second, and third persons, each singular and plural, each in present, past, and future tenses, each in indicative, subjunctive and conditional moods, plus some imperative, participle, and infinitive forms.

Languages outside the Indo-European family, such as those spoken in Africa or the Americas, can be even more prolific. In the Bantu language Kivunjo, for example, a verb is encrusted with prefixes and suffixes that multiply out to half a million combinations per verb.¹¹ But English speakers subsist on only four:

open
opens
opened
opening

Strangely enough, English grammar does not have only four roles for verbs to play. It has at least thirteen different roles, but it shares the four forms among them, as if suffixes were expensive and the designers of the language wanted to economize.

The first suffix is a silent bit of nothing, $-\emptyset$, which when added to the stem *open* turns it into the inflected form *open*. You may wonder: Why say that speakers hallucinate an imaginary suffix at the end of a word? The reason is that it distinguishes the root or stem—the irreducible nugget found in the mental dictionary that captures the essence of a verb and upon which suffixes are hung—from a particular incarnation of that verb with a particular person, number, and tense. In English they can sound the same—to *open* and *I open*—which disguises the fact that they are different versions of the verb. In other languages the form of the verb that you look up in a dictionary cannot be pronounced. For example, in Spanish you can say *canto*, *cantéis*, *canten*, and so on, leaving *cant-* as the stem, but you can never say *cant-* by itself. Stems are therefore not the same things as pronounceable verb forms, and that distinction is useful to preserve in English—to *open* versus *open* \emptyset —even though the two forms sometimes sound the same.

The suffix, $-\emptyset$ is used in four variations of the verb in English:

Present tense, all but third-person singular: I, you, we, they *open* it.
Infinitive: They may *open* it, They tried to *open* it.
Imperative: *Open!*
Subjunctive: They insisted that it *open*.

The suffix *-s* is used for only one purpose:

Present tense, third-person singular: He, she, it *opens* the door.

The suffix *-ing* is used in at least four ways:

Progressive participle: He is *opening* it.
Present participle: He tried *opening* the door.
Verbal noun (gerund): His incessant *opening* of the boxes.
Verbal adjective: A quietly-*opening* door.

Finally we come to our friend *-ed*, which has four jobs:

Past tense: It *opened*.
Perfect Participle: It has *opened*.
Passive Participle: It was being *opened*.
Verbal adjective: A recently-*opened* box.¹²

Why make all these distinctions among verb forms that sound the same? One reason is that the list of phrases calling for a form such as *opened* have nothing in common: To capture the behavior of *-ed*, we have no choice but to list four phrase types separately. Another reason is that some distinctions that are inaudible for regular verbs are audible for irregular ones, and this shows that English speakers register these distinctions as they speak. About a third of the irregular verbs have different forms for the stem, the past tense, and the perfect participle: *I sing, I sang, I have sung; I eat, I ate, I have eaten*. A few make a further distinction and have a special form for the verbal adjective—a *newly wedded couple; a drunken sailor; a shrunken head; rotten eggs*—which is not used for the participle: people say *They have wed*, not *wedded*; *He has drunk*, not *drunken*; *It has shrunk*, not *shrunk*; *The eggs have rotted*, not *rotten*. And one verb comes in *eight* different forms:

Infinitive; subjunctive; imperative: To *be* or not to *be*; Let it *be*; *Be* prepared.
Present tense, first-person singular: I *am* the walrus.
Present tense, second-person singular, all persons plural: You/we/they *are* family.
Present tense, third-person singular: He/she/it *is* the rock.
Past tense, first- and third-person singular: I/he/she/it *was* born by the river.
Past tense, second-person singular, all persons plural; subjunctive: The way we/you/they *were*; If I *were* a rich man.

Progressive and present participle; gerund: You're *being* silly; It's not easy *being* green; *Being* and Nothingness.

Perfect participle: I've *been* a puppet, a pauper, a pirate, a poet, a pawn and a king.

With nouns, too, different grammatical forms have to dip into the same small pool of suffixes. The naked stem *dog* must be distinguished from the singular *dog* + \emptyset because a *dogcatcher* doesn't catch just one dog and a *dog lover* doesn't love just one. The *dog* inside these compounds refers to dogs in general and thus differs in meaning from the singular form in a *dog*. The plural *dogs* uses *-s*, which we have already met in the verb system in *She opens the door*. The possessive forms *dog's* (singular) and *dogs'* (plural) use it too; the three noun forms *dogs*, *dog's*, and *dogs'* differ only in punctuation.

All this redundancy suggests that regular inflection in English is remarkably simple. All the inflections are suffixes; none of the grammatical roles call for a prefix or some other way of decorating or tinkering with a word. And every word has at most one inflectional suffix. We never get *opened* or *opening*, nor do the plural *-s* and possessive *'s* stack up when several owners own something: *the dogs' blanket*, not *the dogs's (dogzez) blanket*. Finally, each niblet of sound making up a suffix has a life of its own and combines with several verb forms, noun forms, or both, rather than being a slave to only one role. This suggests that instead of crediting English speakers with seventeen verbose rules like "To form the past tense, add *-ed* to the end of the verb," we can credit them with just *one* rule:¹³ "A word may be composed of a stem followed by a suffix," like the simple rule shown on page 16. All the other details can be handled by assuming that suffixes are stored in the mental lexicon with entries like those for words, perhaps something like this:

- ed*
- sound: *d*
 - part of speech: suffix
 - use 1: past tense of a verb
 - use 2: perfect participle of a verb
 - use 3: passive participle of a verb
 - use 4: adjective formed from a verb

By factoring seventeen verbose rules into one austere rule and four lexical entries, one per suffix, we not only save ink but get some insight into the men-

tal organization of language. English *could* have used seventeen different forms for its seventeen slots in the noun declension and verb conjugation: prefixes such as *ib-*, *tra-*, and *ka-*, suffixes such as *-og*, *-ig*, and *-ab*, and so on. Instead the slots share a few sounds ($-\emptyset$, *-ed*, *-s*, *-ing*) and one position (immediately following the verb). This miserliness, called syncretism, is found in language after language. Syncretism suggests that the mind keeps separate accounts for the templates that build words (for example, "word = stem + suffix"), for the scraps of sound that may be added to words (*-s*, *-ed*, and *-ing*), and for the roles these additions can play (for example, plural, participle, imperative).¹⁴ A particular construction like the English past tense is a mix-and-match affair, assembled by hooking together parts also used in other constructions. No one knows why languages like to recycle their suffixes and other ways of modifying words. It's certainly not to save memory space, because the savings are trivial. Perhaps the reason is to help listeners recognize when a word is composed of a stem and a suffix rather than being a simple stem. Whatever its purpose, syncretism shows that in the language system, combination is in the blood; even the tiniest suffixes are combinations of smaller parts.

Syncretism—one form, several roles—is one kind of violation of the simplest conceivable system in which every sound has one meaning and vice-versa. The other kind of violation—one role, several forms—is rampant in languages as well; linguists call it allomorphy.¹⁵ Take the regular past-tense suffix—or is it suffixes? Though always spelled *-ed*, it is pronounced in three different ways. In *walked*, it is pronounced *t*. In *jogged*, it is pronounced *d*. And in *patted*, it is pronounced *ɪd*, where *ɪ* is a neutral vowel called "schwa." We also find allomorphy in the regular plural: The suffix *-s* has three different forms in *cats*, *dogs*, and *horses*.

Are there in fact three past-tense suffixes and three plural suffixes? In some languages, we are forced to this messy conclusion. Dutch speakers, for example, select either *-en* or *-s* as the regular plural, depending on the sound of the end of the noun. But in English the three-way variation has a simpler explanation, worked out by the linguists Arnold Zwicky and Alan Prince. *One* past tense suffix is stored in the lexicon, not three, and a separate module fiddles with its pronunciation: the rules of phonology, which define the sound pattern or accent of a language.¹⁶

Why do we pronounce the past tense suffix as *t* in *walked*, *d* in *jogged*, and *ɪd* in *patted*? The choice is completely predictable, and can be stated as a list of rules:

1. Use *ɪd* if the verb ends in *t* or *d* (for example, in *patted* and *padded*).
2. If it doesn't, use *t* if the verb ends in an unvoiced consonant—that is, a consonant in which the vocal cords don't buzz, namely *p*, *k*, *f*, *s*, *sh*, *ch*, and *th* (for example, *tapped*, *walked*, *passed*, *sniffed*, *passed*, *bashed*, *touched*, and *frothed*).
3. Use *d* for all other verbs: those ending in vowels, such as *played* and *glowed*, and those ending in the voiced consonants *l*, *r*, *m*, *n*, *b*, *g*, *v*, *z*, *j*, *zh*, and *th* (for example, *smelled*, *marred*, *slammed*, *planned*, *scrubbed*, *pegged*, *saved*, *buzzed*, *urged*, *camouflaged*, and *bathed*).

This sounds like something out of the tax code. Let's see if we can do better.

The first thing to notice is that nothing in these rules is specific to the past tense. Other constructions that use *-ed* work the same way:

	<i>t</i>	<i>d</i>	<i>ɪd</i>
Past tense:	kicked	flogged	patted
Perfect participle:	has kicked	has flogged	has patted
Passive participle:	was kicked	was flogged	was patted
Verbal adjective:	a kicked dog	a flogged horse	a patted cat

Outside the verb system entirely is yet another *-ed* construction that comes in the three variations; it turns a noun that means "X" into an adjective that means "having X":

	<i>t</i>	<i>d</i>	<i>ɪd</i>
Nominal adjective:	hooked	long-nosed	one-handed
	saber-toothed	horned	talented
	pimple-faced	winged	kindhearted
	foulmouthed	moneyed	warm-blooded
	thick-necked	bad-tempered	bareheaded

The regular plural *-s* also comes in three forms, which you can hear in *hawks*, *dogs*, and *horses*. The variation mirrors the past tense uncannily. Use *ɪz* when the noun ends in a sibilant sound: *s*, *z*, *sh*, *zh*, *j*, or *ch*. If it doesn't, use *s* if the noun ends in an unvoiced consonant. Use *z* for all other nouns. In fact, not only does this pattern appear with the plural, it appears with the other *-s* suffixes as well:

	<i>s</i>	<i>z</i>	<i>ɪz</i>
Plural:	hawks	dogs	horses
3rd person singular:	hits	sheds	chooses
Possessive:	Pat's	Fred's	George's

The variation even appears in versions of *-s* that aren't genuine suffixes. English speakers commonly contract the verbs *has*, *is*, and *does* to their final consonant and glue it onto the end of the subject, as in *Mom's left* or *Dad's home*. Sure enough, the contraction is pronounced in three ways, depending on how the noun ends:

	<i>s</i>	<i>z</i>	<i>ɪz</i>
<i>has</i> :	Pat's eaten.	Fred's eaten.	George's eaten.
<i>is</i> :	Pat's eating.	Fred's eating.	George's eating.
<i>does</i> :	What's he want?	Where's he live?	

That's not all. English has an *affective -s* that can be used to form nicknames in some dialects and argots, as in *Pops*, *Moms*, *Fats*, *Pats*, and *Wills* (the prince second in line to the British throne). That *-s* can also show up in emotionally colored slang such as *bonkers* and *nuts*, similar to the *-y* and *-o* that give us *batty* and *wacko*. (Sometimes the two suffixes are even used together, as in *Patsy*, *Bugsy*, *Mugsy*, *footsie*, *fatso*, and *Ratso*.) Still another version of *-s* appears in adverbial forms such as *unawares*, *nowadays*, *besides*, *backwards*, *thereabouts*, and *amidships*. A final use for *s* is as a meaningless link joining the words in compounds such as *hunterman*, *statesman*, *kinsman*, *bondsman*, *Scotsman*, and *grantsmanship*. And yes, all of these *-s*'s can be pronounced either as *s* or as *z*, depending on the preceding consonant (it's hard to come up with examples for the third column):

	<i>s</i>	<i>z</i>	<i>ɪz</i>
Affective:	Pops, Patsy	Wills, bonkers	

Adverbial:	thereabouts	towards, nowadays
Link in compound:	hunter	landsman

So we have *fifteen* suffixes that show the same three-way or two-way variation. Forty-one suffixes that happen to fall into fifteen parallel sets of alternatives is too much of a coincidence to stomach. More likely, *one* set of rules creates the three-way variation, and the set applies in at least fifteen situations.

There is a second, equally striking set of coincidences that runs across the suffixes. If the variation came from any old set of *if . . . then* rules, we would expect to find all kinds of pairings between stems and suffixes: for example, "Use *s* after the vowels *a* and *e* or after the consonants *th* and *g*," "Use *d* after a *k*," and so on. But the rules are far more lawful than that. The *t* sound comes after unvoiced consonants, and the *t* itself is unvoiced. The *d* sound comes after voiced sounds, and the *d* itself is voiced. The *-s* suffixes show the same chameleonlike behavior: We find unvoiced *s* after unvoiced consonants, and voiced *z* after voiced consonants. It looks as if something is trying to keep the consonants at the end of a word consistent: All of them are voiced, or all of them are unvoiced.

Indeed, something is—the sound pattern of the English language. English never forces speakers to turn their vocal cords on for one consonant then off for the next, or vice-versa. We see the restriction in force in one-piece words that end in a cluster of consonants. These words never received a suffix; they just happen to be built that way, so any sound pattern they display cannot have come from a suffix rule, but rather from the way English speakers like to pronounce words in general. In all but one of these words, the vocal cord switch can be left in the "off" position:

After <i>k</i> (unvoiced):	<i>s</i> can occur <i>ax, fix, box</i>	<i>z</i> cannot occur —
	<i>t</i> can occur <i>act, fact, product</i>	<i>d</i> cannot occur —
After <i>p</i> (unvoiced):	<i>s</i> can occur <i>traipse, lapse, corpse</i>	<i>z</i> cannot occur —
	<i>t</i> can occur <i>apt, opt, abrupt</i>	<i>d</i> cannot occur —
After <i>t</i> (unvoiced):	<i>s</i> can occur <i>blitz, kibitz, Potts</i>	<i>z</i> cannot occur —
After <i>s</i> (unvoiced):	<i>t</i> can occur <i>post, ghost, list</i>	<i>d</i> cannot occur —

In one English word, *adze*, the vocal cord switch is left in the "on" position:

After <i>d</i> (voiced):	<i>s</i> cannot occur —	<i>z</i> can occur <i>adze</i>
--------------------------	----------------------------	-----------------------------------

In *no* English word is the voicing switch toggled on and off, in an ending like *zt, gs, kz, or sd*.

These difficult-to-pronounce clusters *can*, however, be created by a dumb rule of morphology that pins a suffix onto the end of a word without regard for how the resulting train of consonants is to be pronounced. That is what happens when a rule adds a *d* sound to *walk* or an *s* sound to *dog*. English cleans up these awkward mismatches with a different kind of rule. The rule says, "When there is a cluster of consonants at the end of a syllable, adjust the voicing setting of the last consonant to make it consistent with its neighbor on the left." (In other words, change *kz* to *ks*, *pd* to *pt*, and so on.) The rule does not care whether the syllable was formed by a past-tense suffix, a plural suffix, a contracted *has*, a nickname with *-s*, or anything else. It kicks in *after* the syllable has been assembled, in the cleanup module we call phonology.

Can we now tell whether the suffix stored in the lexicon is *-d*, and is converted to a *t* when it finds itself at the end of *walk*, or whether it is *-t* and is converted to *d* when it finds itself at the end of *jog*? A little detective work can settle the question. Not every sound cares about the consonant that follows it. Those that do are consonants in which the airstream is obstructed, namely *p, b, t, d, k, g, s, sh, ch, z, zh,* and *th*. But the vowels, and the vowel-like consonants *r, l, n,* and *m*, are indifferent to what comes after them; they tolerate either *s* or *z*, either *t* or *d*, as we see in these one-piece words:

After <i>n</i> :	<i>s</i> can occur <i>fence</i>	<i>z</i> can also occur <i>lens</i>
	<i>t</i> can occur <i>lent</i>	<i>d</i> can also occur <i>lend</i>
After <i>r</i> :	<i>s</i> can occur <i>force</i>	<i>z</i> can also occur <i>furze</i>
	<i>t</i> can occur <i>fort</i>	<i>d</i> can also occur <i>ford</i>
After <i>l</i> :	<i>s</i> can occur <i>pulse</i>	<i>z</i> can also occur <i>Stolz</i>
	<i>t</i> can occur <i>guilt</i>	<i>d</i> can also occur <i>guild</i>

After a vowel:	s can occur <i>niece</i>	z can also occur <i>sneeze</i>
	t can occur <i>goat</i>	d can also occur <i>goad</i>

Here we have laissez-faire environments in which the suffixes can show their true colors, untouched by rules of phonology. What do we find? That the virgin suffixes are pronounced *-d* and *-z*, not *-t* and *-s*:

After <i>n</i> :	we don't say s —	we say z <i>grins</i> (grīnz), <i>pins</i> (pīnz)
	we don't say t —	we say d <i>grinned</i>
After <i>r</i> :	we don't say s —	we say z <i>wears</i> (wērz), <i>cores</i> (kōrz)
	we don't say t —	we say d <i>feared</i>
After <i>l</i> :	we don't say s —	we say z <i>calls</i> (kōlz), <i>balls</i> (bōlz)
	we don't say t —	we say d <i>smiled</i> , <i>well-heeled</i>
After a vowel:	we don't say s —	we say z <i>flees</i> (flēz), <i>fleas</i> (flēz)
	we don't say t —	we say d <i>flowed</i>

The *-t* and *-s* we hear in words with choosy sounds such as *walked* and *cats* must be the aftermath of the rule.

Finally, what about the funny extra vowel in *patted* and *horses*? Here again the change in sound is not some random act of vandalism. The vowel appears when *d* follows *t* or *d*, and when *z* follows *s* or *z*. The word endings that trigger the extra vowel are similar in pronunciation to the suffixes themselves, and that can't be a coincidence. Apparently a rule is trying to separate too-similar adjacent consonants by pushing a vowel between them: between *t* and *d*, *d* and *d*, *s* and *z*, *z* and *z*, *sh* and *z*, and so on. In many languages the rules of phonology do something when a rule of morphology leaves two identical or near-identical consonants in a row, presumably because there's no natural way to pronounce them. Some languages drop the second consonant, others merge the two into

one long consonant, and still others, like English, wedge a vowel between them. As with the rule that fiddles with voicing, the rule that inserts a vowel must live in a phonology module separate from rules that stick on the various suffixes, because the rule is oblivious to what kind of suffix it manipulates.

We even can deduce which of the two rules applies first, the one that changes the voicing setting or the one that inserts the vowel. The devoicing rule is triggered by adjacent consonants; the vowel rule breaks up adjacent consonants. If the voicing rule came first, it would convert *pat + d* to *pat + t*, and only then would the vowel be inserted, yielding *pātīt*:

Morphology:	<i>pāt + d</i>
↓	
Devoicing:	<i>pāt + t</i>
↓	
Vowel insertion:	<i>pāt + ī + t</i>

But that is not how we pronounce it; we say *pātīd*. This means that the vowel rule must have come first, creating *patted*; now the voicing rule is no longer compelled to do anything, because the *td* sequence that would trigger it has been broken up:

Morphology:	<i>pāt + d</i>
↓	
Vowel insertion:	<i>pāt + ī + d</i>
↓	
Devoicing:	not triggered

The ordering makes sense when you think about how the phonology module should be organized. It has some rules that edit the string of vowels and consonants composing a word (phonology proper), and other rules that convert the string into actual sounds or muscle movements (phonetics). The vowel-insertion rule makes a major change in the stuff that makes up a word, and belongs in the first subcomponent; the voicing rule does a last-minute adjustment of pronunciation for the benefit of the muscles, and belongs in the second.¹⁷

This completes the analysis of the three versions of the past-tense suffix. When we started, we needed forty-odd rules, each stipulating that some suffix be placed next to some word ending. We have ended up with just two rules. Best of all, what the rules do, why they do it, and in what order they do it all

make sense in the light of the sound pattern of English. Indeed, this kind of layering may be found in languages all over the world.

Incidentally, there is corroborating evidence of a completely different kind that shows that the three forms of *-ed* and *-s* are created on the fly by a phonological rule. Some psycholinguists keep a pad and pencil in their pockets and write down every slip of the tongue they hear. People make one or two such errors for every thousand words they say, and many of the errors consist in deleting, repeating, or switching around vowels or consonants.¹⁸ The last kind of error is called a Spoonerism, in honor of the Reverend William Spooner (1844–1930), warden of New College at Oxford, who came out with surprises such as *Our queer old dean, You have hissed all my mystery lessons and tasted the whole worm, and It is now kistomary to cuss the bride*. They sound too good to be true, but I have heard similar errors myself. After I spoke at a scientific symposium the chair wrapped up the session by saying *I would like to spank the speakers*, and when I asked a friend how he liked his new condominium, he said *It seats my nudes*.

Speech errors provide clues on how the speech system is organized. For example, when a person intends to say *grapefruits* but accidentally leaves out the *t*, how does he pronounce the plural? If there were a distinct plural suffix pronounced *-ss*, he would say *grapefrooss*, since this is what the *t* in the *grapefruit* entry would have demanded. In fact he says *grapefroofz*—pronouncing the plural as *z*, which is appropriate to words ending in a vowel.¹⁹ Similarly, a person may say *The infant tucks—touches the nipple*, not *tuck-is*, or may say *Did you buy enough breakfasis?*, not *breakfass*. The errors show that the form of the suffix must be computed after the vowels and consonants of the noun or verb were placed on the chute to the vocal tract.

English did not always have single-consonant suffixes and a rule that separates them from a too-similar word ending. Our current system is the result of a reorganization that began around the time of the origin of Modern English in the seventeenth century. Before that, *-ed* and *-s* suffixes were pronounced (and spelled) with vowels all the time, not just with words ending in *t* or *d* or in *s* or *z*. For centuries, English speakers had been concentrating stress on the first syllables of words, which shriveled the later syllables, and speakers began to leave out the vowels in the suffixes of many words. Writers called attention to the new, clipped pronunciations by spelling them phonetically with an apostrophe in place of the deleted vowel, as in Shakespeare's play about "a pair of star-cross'd lovers":

Death, that has suck'd the honey of thy breath,
Hath no power yet upon thy beauty:

Thou art not conquer'd; beauty's ensign yet
Is crimson in thy lips and in thy cheeks.

The guardians of the English language deplored the change, as they do all changes. In "A Proposal for Correcting, Improving, and Ascertaining the English Tongue," Jonathan Swift wrote:

What does your lordship think of the words "drudg'd," "disturb'd," "rebuk'd," "fledg'd," and a thousand others everywhere to be met with in prose as well as verse? Where, by leaving out a vowel to save a syllable, we form so jarring a sound, and so difficult to utter, that I have often wondered how it could ever obtain.

His contemporary, Samuel Johnson, who was standardizing the spellings of English words in a way that reflected the morphemes that composed them, recognized that *'d* and *-ed* were the same morpheme, and obliterated the distinction in their spelling, making *ed* the spelling for both.²⁰ It is unclear why he chose to leave the *e* in *-ed* across the board (*mapped* and *matted*), but opted to spell *-s* either with or without an *e*, depending on how it is pronounced (*maps* and *masses*).

Today the old syllabic suffix survives in a handful of adjectives: *accursed*, *aged*, *beloved*, *bended* (in the expression *on bended knees*), *blessed*, *crooked*, *cussed*, *dogged*, *jagged*, *learned*, *naked*, *ragged*, *wicked*, and *wretched*. (A few more survive in rural dialects, such as *forkèd*, *peakèd*, *streakèd*, and *stripèd*.)²¹ Many of them are archaic or poetic and are used mainly in self-conscious speech. The psychologist Melissa Bowerman, a researcher of child language, had this exchange with her four-year-old daughter about a class trip to a natural history museum:²²

MOTHER (*playfully*): Maybe you'll see something wingèd.
DAUGHTER: Maybe we'll see something snakèd!



We've seen why the syntax box, which builds phrases and sentences, has to be separated from the morphology box, which builds words. We also have seen why the phonology box, which massages words into a pronounceable stream of

sound, has to be separated from syntax, morphology, and the lexicon. But why do we need separate boxes for semantics (the thoughts expressed in language) and the lexicon? Could we reduce the difference between regular and irregular verbs to a difference in meaning between the two kinds of verbs, rather than putting one kind in the morphology box and the other in the lexicon? Do we even need to talk about an “entry in the mental lexicon,” the address in memory that holds a link to a sound and a link to a meaning? Or could we connect thoughts to sounds directly, eliminating the middleman? Here are some facts that suggest that we do need to credit the human mind with something like dictionary entries.

First, the English irregular verbs could not have arisen simply from a communal effort to optimize clarity. While irregular forms on average are harder to mistake for their base forms than regular forms are (*bring* doesn't sound like *brought*, nor *take* like *took*), many irregulars are *identical* to their base forms: *Today I hit, yesterday I hit; Today I put, yesterday I put*. A sentence such as *On Wednesday I cut the grass* could mean last Wednesday, next Wednesday, or every Wednesday. If *cut* were regular, the ambiguity would never arise: *On Wednesday I cutted the grass* would single out the preceding Wednesday. Despite the potential ambiguity, however, twenty-eight English verbs insist on remaining unchanged in the past tense.

Also, irregular forms do not correlate with any kind of meaning. Many verbs are similar in meaning but have completely different past-tense forms. For example, *hit*, *strike*, and *slap* all refer to hitting. *Hit* is an irregular verb that does not change in the past tense: *Today we hit golf balls; Yesterday we hit golf balls*. *Strike* is an irregular verb that changes its vowel, yielding *struck*. And *slap* is a regular verb, with past tense *slapped*.

Not only are there verbs with similar meanings and different past-tense forms, there are verbs with different meanings and the same past-tense forms. English has a class of verbs linguists call *light verbs*, such as *come*, *go*, *do*, *take*, *have*, *set*, *get*, *put*, and *stand*. Compared to ordinary verbs they are less filling; a light verb doesn't have a meaning that stays with it, but takes on dozens of meanings, especially in combination with particles such as *in*, *out*, *up*, *off*, *over*, and *around*:

come (move to here), *come around* (agree), *come in to* (inherit),
come (reach orgasm), *come off as* (appear), *come out* (divulge homosexuality), *come to* (awaken)

go (move to there), *go out with* (date), *go nuts* (dement), *go in for* (choose), *go off* (explode), *go off* (spoil)
do (act), *do in* (kill), *do up* (decorate), *do a number on* (overwhelm),
do lunch (eat together)
take (cause to go with), *take in* (swindle), *take off* (launch), *take in* (welcome), *take over* (usurp), *take up* (commence), *take a leak* (urinate), *take a bath* (lose money), *take a bath* (bathe), *take a walk* (walk), *take a look* (look)
have (possess), *have* (eat), *have* (seduce), *have a heart* (sympathize),
have over (entertain), *have a cow* (be angry)
get (retrieve), *get* (become), *get over* (survive), *get out* (divulge), *get off on* (enjoy), *get a life* (self-improve)
set (place), *set off* (ignite), *set up* (arrange), *set up* (trick), *set up* (introduce), *set right* (rectify), *set the stage* (prepare)
put (cause to be at), *put off* (procrastinate), *put off* (offend), *put one over on* (fool), *put down* (insult), *put down* (euthanize), *put in for* (request), *put out* (extinguish), *put out* (inconvenience), *put out* (consent to sex)
stand (rise), *stand out* (impress), *stand up for* (defend), *stand in* (replace), *stand off* (repel)

But in every instance they retain their irregular past tense forms in the extended meanings: *Barney came around*, *Barney came out*, *Barney came off as* (never comed); *Joan took him in*, *Joan took a bath*, *Joan took over* (never taked); and so on. All the meanings march in lockstep with the same irregular past-tense forms, no matter how tenuous the semantic thread that links them. The mind links an irregular sound such as *took* not with the meaning of a word directly but with the word's *root*—a unique address in the mental lexicon, like the boldfaced entry for a word in a dictionary, which can have several meanings listed under it.²³

An even more curious demonstration comes from families of words with the same stem and different prefixes. Words with prefixes keep the past-tense form of the stem: *eat*—*ate* becomes *overeat*—*overate*; *make*—*made* becomes *re-make*—*remade*. That is not surprising, because we all hear the *eat* inside *overeat*—overeating is, after all, a kind of eating, namely, eating too much. What is surprising is that the same thing happens when the meaning of the combination is opaque. Few people sense the meaning of the *stand* inside

understand, the *get* inside *forget*, or the *come* inside *become*. Nonetheless no one is tempted to say *understanded*, *forgot*, or *becomed*; the irregular forms persist, giving us *understood*, *forgot*, and *became*. Here are some examples:

come—came, become—became, overcome—overcame
go—went, undergo—underwent
get—got, forget—forgot
take—took, mistake—mistook, overtake—overtook, partake—partook,
undertake—undertook
set—set, beset—beset, upset—upset
stand—stood, understand—understood, withstand—withstood
draw—drew, withdraw—withdrew
hold—held, behold—beheld, uphold—upheld, withhold—withheld
give—gave, forgive—forgave

Irregular forms stick like glue to their verb roots, even when reduced to meaningless little tokens inside a bigger verb. Speakers of English seem to analyze *become* as *be-* + *come* and *understand* as *under-* + *stand*, even though the meaning of *become* is not computable from the meaning of *be-* and the meaning of *come*, and *understand* has nothing to do with standing. This is not something we have to learn in school. When we acquire language, our minds analyze sets of words, looking for their parts as if they were clues in a combinatorial puzzle. We mentally arrange them in a matrix according to overlap:

	be-	over-	under-	up-	with-
<i>come</i>	<i>become</i>	<i>overcome</i>			
<i>draw</i>					<i>withdraw</i>
<i>hold</i>	<i>behold</i>			<i>uphold</i>	<i>withhold</i>
<i>set</i>	<i>beset</i>			<i>upset</i>	
<i>stand</i>			<i>understand</i>		<i>withstand</i>
<i>take</i>		<i>overtake</i>	<i>undertake</i>		

and use the common denominators in the rows and columns to make incisions in the words, thinking of them thereafter as amalgams of parts: *become* = *be-* + *come*, *withdraw* = *with-* + *draw*, and so on.²⁴

Of course, it was English speakers of centuries past, our linguistic ancestors, who first analyzed *become* as *be-* + *come* and extended the *come—came* pattern to it, and it is possible that to them the words were as transparently

built out of parts as *overeat* or *remake* are to us today. Even so, it is unlikely that we have been stupidly memorizing *became*, *overcame*, *withdrew*, and so on as structureless strings of vowels and consonants. If we were to come across a new complex word, such as *undercome*, *bestand*, *overhold*, or *withset*, and were unaware of its meaning, we would almost certainly use the irregular forms of the words inside them: *undercome* (not *undercomed*), *bestood*, *overheld*, *withset*. Moreover, it is not pure sound that carries the irregular form: The past of *succumb* and *encumber* are *succumbed* and *encumbered*, not *succame* and *encamber*, because people don't perceive them as containing a prefix followed by the word *come*, only the sound *kūm*.

Clearly the perception of an embedded word comes from its spelling: *become* contains *c-o-m-e*; *succumb* doesn't. But spelling does not directly inform speakers how to form the past tense; it merely assigns a distinct visual signature to every root, and speakers choose the past-tense form that goes with the root. Samuel Johnson, who standardized the spellings of thousands of modern words, used people's perception of the anatomy of words as a rationale in his decisions, and that is one of the reasons that the spellings of English words notoriously do not always reflect their sounds; often they reflect morphological structure instead. We see this in the many words that sound alike but are not perceived as being the same word (that is, as having the same root), and are not given the same past-tense form:

<i>meet—met</i>	versus	<i>mete—meted</i>
<i>ring—rang</i>	versus	<i>wring—wring</i>
<i>bear—bore</i>	versus	<i>bare—bared</i>
<i>steal—stole</i>	versus	<i>steel—steeled</i>
<i>break—broke</i>	versus	<i>brake—braked</i>

In the last three cases the spellings divulge the presence of words that are recognizable in other guises—the adjective *bare*, the noun *steel*, the noun *brake*—and we will see in chapter 6 that this makes an especially big difference in how we compute their past tense forms.²⁵



The English system of inflection, we have seen, dissects cleanly into a few simple components. The past-tense rule belongs to a component, morphology, that builds things out of parts using rules. The rule itself is a masterpiece of

minimalism—“a word can be composed of a stem and a suffix”—with all other details distilled out and collected in the lexical entry for the suffix. The suffix itself is shared among several inflections (past tense, participle, and so on), and its variant pronunciations (*t*, *d*, *-id*) do not wastefully multiply listings but are computed automatically by two ubiquitous rules of phonology. The distinction between the lexicon (including irregular inflection) and grammar (including regular inflection) is a distinction between a list of entries and an algorithm for combining them, rather than a side effect of a general yearning to distinguish meanings.

That leaves the irregular words. Every irregular tells a story, and they are the topic of the next chapter.

3

BROKEN TELEPHONE

In the game known as Broken Telephone (or Chinese Whispers) a child whispers a phrase into the ear of a second child, who whispers it into the ear of a third child, and so on. Distortions accumulate, and when the last child announces the phrase, it is comically different from the original. The game works because each child does not merely degrade the phrase, which would culminate in a mumble, but *reanalyzes* it, making a best guess about the words the preceding child had in mind.

All languages change through the centuries.* We do not speak like Shakespeare (1564–1616), who did not speak like Chaucer (1343–1400), who did not speak like the author of *Beowulf* (around 750–800). As the changes take place, people feel the ground eroding under their feet and in every era have predicted the imminent demise of the language. Yet the twelve hundred years of changes since *Beowulf* have not left us grunting like Tarzan, and that is because language change is a game of Broken Telephone.

A generation of speakers uses their lexicon and grammar to produce sentences. The younger generation listens to the sentences and tries to infer the lexicon and grammar, the remarkable feat we call language acquisition. The transmission of a lexicon and grammar in language acquisition is fairly high in fidelity—you probably can communicate well with your parents and your chil-

*For a chart that summarizes the history, dates, and family affinities of the English language, see page 212.

dren—but it is never perfect. Words rise and fall in popularity as the needs of daily life change, and also as the hip try to sound different from the dweebs and graybeards. Speakers swallow or warp some sounds to save effort, and enunciate or shift others to make themselves understood. Immigrants or conquerors with regional or foreign accents may swamp the locals and change the pool of speech available to children.

Children, for their part, do not mimic sentences like parrots but try to make sense of them in terms of underlying words and rules. They may hear a mumbled consonant as no consonant at all, or a drawn-out or mispronounced vowel as a different vowel. They may fail to discern the rationale for a rule and simply memorize its outputs as a list. Or they may latch on to some habitual way of ordering words and hypothesize a new rule to make sense of it. The language of their generation will have changed, though it need not have deteriorated. Then the process is repeated with their children. Each change may be small, but as changes accumulate over centuries they reshape the language, just as erosion and sedimentation imperceptibly sculpt the earth.

That is how irregular forms, in particular, come down to us. Most of the forms were originally created by rules, but a later generation never grasped the rules and instead memorized the forms as words. They were words for every generation thereafter, and each irregular was free to accumulate its own quirks from subsequent distortions and reanalyses. Because irregulars originated from rules they are not a random grab-bag but rather display patterns, fossils of the long-dead rules. A. L. Kroeber, a founder of modern anthropology, reminisced that his “first remembered purely intellectual pleasure” was seeing patterns in English irregular verbs, a foretaste of his search for systematicity in culture more generally.¹

This chapter is a guided tour of the irregular nouns and verbs of English, with commentary on where they came from and where they are going. These words all will have their turns on stage throughout the book, so it’s helpful to get to know them individually. This is also a lively way to come to understand how language changes, including how it is changing today.

People often ask me how linguists know the way people pronounced things in centuries past. After all, Chaucer, unlike Nixon, did not secretly tape his conversations for the benefit of future historians. Old pronunciations can be painstakingly inferred from a diverse set of clues. One of them is spelling. Before Samuel Johnson standardized English orthography, people spelled more or less as they pleased, trying to capture the sounds of language as they heard them. Spellings were more phonetic, and changes in spelling give clues to

changes in pronunciation. For example, when writers started to spell Old English *bi-healfe* (behalf) as *behaf*, one can guess that people had stopped pronouncing the *l*. Other clues come from wordplay. For example, Shakespeare rhymed or punned *case* and *ease*, *hate* and *eate*, *say* and *sea*, and *shape* and *sheep*, suggesting that speakers of Early Modern English pronounced the vowels in each pair in the same way (clues from spelling suggest it was *ā*). A third kind of clue is found in the writings of language snobs who criticize or lampoon the speech of their contemporaries, inadvertently immortalizing it to the good fortune of modern linguists. Other clues exist as well, and together they can triangulate on the most common and most probable pronunciations.

We can never say for sure what *the* pronunciation of a given word at a given time actually was. Just as there are regional accents today (London, Boston, Texas, and so on), there were regional varieties of English centuries ago; indeed, many more of them, because people did not move around as much as we do, did not send their children to melting-pot schools, and had no dictionaries to consult. Also, the written record is haphazard. Most words and pronunciations were in use long before the first literate person chanced to write them down, and many others went to the grave along with their speakers. When word histories can be reconstructed, invariably they are convoluted, eye-glazing yarns. This is to warn you that the word histories presented here have been simplified to highlight the kinds of psychological processes that cause words to *have* histories.²



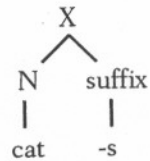
Words aren’t regular or irregular across the board. Words are regular or irregular only with respect to certain inflections, some more tolerant of irregularity than others.

The present progressive suffix *-ing*, as in *The joint is jumping*, is 100 percent regular. There isn’t a single exception to the rule, not even the rebellious *be*, which meekly submits and shows up as *being*. Why, when it comes to *-ing*, does no verb hear a different drummer? One reason is that the progressive construction came into English relatively recently, late in the Middle English period of 1100 to 1450. It borrowed the *-ing* suffix from the gerund (a construction that turns a verb into a noun, as in *the changing of the guard*), and the newly cloned *-ing* suffix had the progressive all to itself and did not have to compete with alternative forms hanging around from earlier periods. Another reason is that *-ing* is found in a separate syllable, which makes it easy for listeners to hear a word such as *breaking* as *break + ing*. That is an advantage over *-s* and *-ed*, which can sound as if they are part of a stem, like *act*, *box*, or

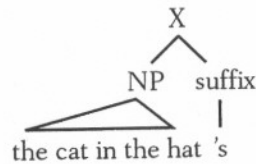
maze. As we shall see, the camouflage of *-s* and *-ed* can invite listeners to misanalyze a regularly inflected combination as a one-piece irregular word.

One other suffix is completely regular: the possessive *'s*. Any noun can take it, even the irregular nouns that cannot appear with an *s* sound when it is a plural suffix, such as *mouse* and *man*. We have no trouble saying *the man's hat*, *the mouse's mother*, or *the goose's egg*, even though we never say *the mans*, *the mouses*, or *the gooses*. Why no irregulars? The possessive is unusual because it attaches to a *phrase* rather than to a *word*. One can talk not just about *the cat's pajamas* but about *the cat in the hat's pajamas*, where the pajamas belong to the cat, not to the hat:

The plural *-s* attaches to a word:



The possessive *'s* attaches to a phrase:



A former student, Annie Senghas, once said to someone at a conference, "The woman sitting next to Steven Pinker's pants are like mine." I was fully clothed; the woman sitting next to me had pants like Annie's. Dave Barry's column-within-a-column "Ask Mr. Language Person" once had the following exchange:

Q: Recently, did your research assistant Judi Smith make a grammatically interesting statement regarding where her friend, Vickie, parks at the Miami Herald?

A: Yes. She said, quote, "She comes and parks in whoever's not here's space that day."

The word *here* is not even a noun! Since *'s* is perceived not to be attached to an adjacent noun, it cannot unite with that noun in people's minds, and therefore never evolves into an irregular word. The exceptions that prove the rule are the possessive pronouns *my*, *your*, *his*, *her*, *our*, and *their*, which are, in a sense, irregular replacements for *me's*, *you's*, *him's*, *her's*, *us's*, and *them's*. Pronouns are one-word phrases; in any sentence position where you can say *the man in the gray suit* you can also say *he* or *him*. A pronoun, being a phrase, is the only kind of word that *could* form a cohesive amalgam with *'s*, which in effect is what possessive pronouns are.

The third-person singular *-s*, as in *Dog bites man*, steps aside for irregular forms in only four verbs: *be-is* (not *be's*), *have-has*, *do-does* (pronounced *dūz*), and *say-says* (pronounced *sēz*). These, by the way, are the four most frequent verbs in the English language.³ In chapter 5 we will see that this is not a coincidence.

Nouns embrace several kinds of irregular plurals.⁴ Many nouns ordinarily don't take any plurals: mass nouns such as *mud*, *celery*, *furniture*, and *evidence* are treated as seamless stuff rather than countable things. (A former graduate student who is a Russian emigré was teased by fellow students for saying, "I hev three evidences for thees theory.") Of the count nouns, which do take plurals, exactly seven change their vowel instead of adding *-s*:

man-men, *woman-women* (pronounced *wīmīn*), *foot-feet*,
goose-geese, *tooth-teeth*, *mouse-mice*, *louse-lice*,

Why do we flip the vowels in these nouns? Originally they took plural suffixes, just like regular nouns, though the suffixes were different from today's *-s*. For example, *foot*, originally *fot*, had the plural *foti*. But as we saw in chapter 2, you can't just force a consonant or vowel onto the end of a word and hope that nothing else happens. People adjust their pronunciation of a sound in anticipation of the sounds to come. In many modern English dialects, for instance, speakers pronounce the *i* differently in *write* and *ride* and the *ou* differently in *shroud* and *about*. In *keep cool* the first *k* sound is pronounced toward the front of the mouth, the second one toward the back. In words like *find* and *sound* the *n* vanishes and the vowel reminds us of the vanished consonant by being sounded through the nose. Most of us are unaware that we make these adjustments and are puzzled when children spell *find* as *fid*, though it is an accurate transcription of the *n*-less word they hear. Some of these adjustments come from the way we control our muscles, but others get standardized into phonological rules, which define what we hear as an accent.

In the Germanic languages that were ancestral to English there was a phonological rule that changed the pronunciation of a vowel from the back of the mouth to the front of the mouth if the next syllable contained a vowel pronounced high and in front. The rule spared people from having to jerk their tongue backward and then forward while pronouncing the words. So in *foti*, the plural of *fot*, the back *o* was altered to a front *e*, harmonizing with the front

i in the suffix: roughly, *feti*. The process is called umlaut and it is still visible in our linguistic cousin, German, as the two little dots over some vowels: *die Kuh* (the cow), *die Kühe* (the cows).

In the Middle English period, speakers began to mumble the unstressed syllables at the ends of words and then began to drop them outright. At that point people must have been hearing the altered vowel in *feti* as a different vowel altogether, not as a tweaked *o*, because when the suffix was dropped, the speakers kept the altered vowel in the stem, even though nothing was there to tweak it anymore. The eventual result was *feet*. It reminds me of the explanation of why there is a basketball team in arid Los Angeles called the Lakers and a team in pious Utah called the Jazz. Originally the teams were based in Minneapolis, The Land of Lakes, and in New Orleans, The Birthplace of Jazz. When the teams moved, they kept their names, even though the names no longer made sense.

Another three irregular plurals take the old Anglo-Saxon suffix *-en* rather than *-s*:

child-children, ox-oxen, brother-brethren

Of the three, only *children* is part of the standard American vernacular (though the others are preserved in some nonstandard dialects, together with archaic plurals such as *eyen*, *shoon*, and *hosen*). Most Americans meet *oxen* mainly in writing, and commonly say *oxes* instead.⁵ Similarly, they perceive *brethren* as an inhorn term for monks and parishioners. As a result, the *-en* sounds archaic and lends itself to silly wordplay. Shortly after the appointment of Ruth Bader Ginsburg to the U.S. Supreme Court, where she joined fellow person of gender Sandra Day O'Connor, *Newsweek* reported, "The brethren—and now, two sistren—had to know that the swiftness and scope of their ruling would be viewed as a landmark victory for working women."⁶ In the argot of computer hackers, who try to outdo each other with logical extensions of irregular patterns, the plural of the computer called the VAX is *VAXen*, and there also have been sightings of *faxen*, *boxen*, *soxen*, and *Macintoshen*.⁷

Several names for gregarious animals that are hunted, gathered, or farmed are identical in the singular and plural:

fish, cod, flounder, herring, salmon, shrimp
deer, sheep, swine, antelope, bison, elk, moose
grouse, quail

These forms may have come from constructions in which the singular is used to refer to potential quarry in the aggregate, as in *We went hunting for duck*.

A fourth class of nouns takes the regular *-s* ending but changes its final consonant, usually *f* but sometimes *th* or *s*, from unvoiced to voiced:

calf-calves; also *elf, dwarf, half, hoof, knife, leaf, life, loaf, self, scarf, sheaf, shelf,*
thief, wife, wharf, wolf
mouth-mouths; also *truth, sheath, wreath, youth*
house-houses

Something familiar is going on here: A voiced consonant *z* is being shoved against an unvoiced consonant, and one of them bends to make the cluster consistent. We saw this happening in the regular nouns, where *-s* is pronounced differently in *dogs* and *cats*. But strangely, in these nouns the suffix *z* keeps its voicing, and the noun surrenders it—a right-to-left smearing that violates the usual left-to-right smearing of English phonology. Some linguists have posited a special rule, *regressive voicing*, to generate these examples. The rule, though, would have to be handcuffed to these two-dozen-odd words, because most nouns ending in *f* or *th* are regular and would have to be left untouched. The plural of *reef* is regular (*reefs*, not *reeves*), and the same is true for nouns such as these:

birth, booth, earth, faith, growth, hearth, length, month, tenth
belief, brief, chief, proof, safe, spoof, turf

Even many of the so-called irregular nouns are questionable; many speakers simply pronounce *hoofs*, *wharfs*, *oaths*, and *truths* in the ordinary way. I prefer a different theory: that some nouns have two stems, one for the singular, one for the plural, and that the plural stem is tagged as incomplete without a suffix: *knife-*, *loave-*, *wolve-*, and so on. After all, if *-ed* and *-ing* are tagged as suffixes that cannot be pronounced unless they are attached to a stem, why can't there be stems that cannot be pronounced unless they have a suffix attached to them? The regular suffix *-s* then applies, generating the plural form without further ado.⁸

Finally, there are nouns that take Latin or Greek plurals. As the singer Alan Sherman has pointed out, "One hippopotami / Cannot get on a bus. Because one hippopotami / Is two hippopotamus." Here are four families with Latin plurals:

alumnus—alumni; also *bacillus, cactus, focus, fungus, locus, nucleus, radius, stimulus*
genus—genera, corpus—corpora
alga—algae; also *alumna, antenna, formula, larva, nebula, vertebra*
addendum—addenda; also *bacterium, curriculum, datum, desideratum, erratum, maximum, medium, memorandum, millennium, moratorium, ovum, referendum, spectrum, stratum, symposium*
appendix—appendices; also *index, matrix, vortex*

And here are two families with Greek plurals:

analysis—analyses; also *axis, diagnosis, ellipsis, hypothesis, parenthesis, synopsis, synthesis, thesis*
criterion—criteria; also *automaton, ganglion, phenomenon*

These nouns come from science and academia, and the plurals were borrowed directly from Latin or Greek together with the singulars. They must be irregular forms that are memorized as a list, not the products of a rule attaching *-i* or *-ae*, because most nouns shun these plurals except in the speech of people with an attitude:

apparatus—apparatuses; also *bonus, campus, caucus, census, chorus, circus, impetus, prospectus, sinus, status, virus*
area—areas; also *arena, dilemma, diploma, drama, era, etc.*
album—albums; also *aquarium, chrysanthemum, forum, museum, premium, stadium, ultimatum*

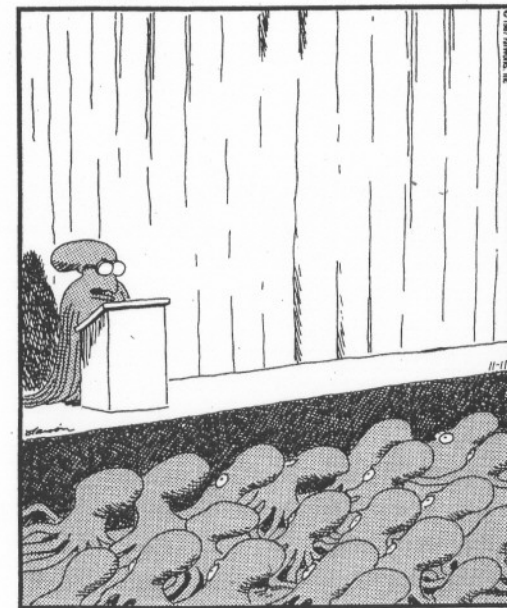
Latin- and Greek-inspired plurals in a sense are still not part of the English language. They are not acquired as part of the mother tongue in childhood, and are uncommon in everyday speech among nonacademic adults. Instead they are learned in school together with the Pythagorean theorem and the dates of the Peloponnesian War. Since they follow no living rule, and people couldn't have memorized them unless they went to the right schools and read the right books, they are shibboleths of membership in the educated elite and *gotcha!* material for pedants and know-it-alls (the kind of people who insist that the millennium begins January 1, 2001).

Admittedly, I cringe when I hear *this phenomena, those criterias, and the media is*, and I could barely contain myself during the speech from the president

of the alumni association who kept thanking *the alumnis*. I also get a perverse pleasure from correcting students who refer to *an important piece of data* or write that *this data is important*. (*Data* is the plural of *datum*, I tell them, so one ought to say, *The datum is important; The data are important*.) Yet by the same logic I ought to correct myself when I refer to *an agenda, two candelabras, this insignia, or that propaganda*, which are the plurals of *agendum, candelabrum, insignium, and propagandum*. And I refuse to hear a word about *genii, termini, aquaria, podia, lexica, fora, stadia, or apices*. In any case, whenever pedants correct, ordinary speakers hypercorrect. so the attempt to foist "proper" Greek and Latin plurals has bred pseudo-erudite horrors such as *axia* (more than one *axiom*), *peni, rhinoceri*, and this one:

THE FAR SIDE

By GARY LARSON



"Fellow octopi, or octopuses ... octopi? ... Dang, it's hard to start a speech with this crowd."

THE FAR SIDE © 1987 FARWORKS, INC. Used by permission. All rights reserved.

It should be "Fellow octopuses." The *-us* in *octopus* is not the Latin noun ending that switches to *-i* in the plural, but the Greek *pous* (foot). The etymologically defensible *octopodes* is not an improvement.

The flip side of plural pomposity is playful punning that deflates it, and for decades wags have seen the opening. In a *Peanuts* cartoon, Linus had to bring

eggshells to Miss Othmar's class so he could make *igli*. The comedian Shelley Berman has talked of *stewardi* wearing *blice*. Wayne and Schuster performed a skit in which Julius Caesar nibbled on a *spaghattus*. In Richard Lederer's "Foxen in the Henhice," Farmer Pluribus reached for some *Kleenices* while being serenaded by *tubae*, *harmonicae*, *accordia*, *fives*, and *dra*.⁹ Henry Beard and Roy McKie's *A Gardener's Dictionary* contains the following entry:¹⁰

Narcissus: wonderful, early-blooming flower with an unsatisfactory plural form. Botanists have been searching for a suitable ending for years, but their attempts—*narcissi* (1947), *narcissusses* (1954), *narcissus* for both singular and plural (1958), and *multinarcissus* and *polynarcissus* (1962, 1963)—haven't enjoyed any real acceptance, and thus, gardeners still prefer to plant the easily pluralized daffodil or jonquil.

This may seem silly and inconsequential, but the following story appeared in *The New Republic* on December 12, 1994: "In Las Vegas, The Flying Elvi sued The Flying Elvises for trademark theft. Both organizations leap from airplanes in Elvis Presley (late period) costumes and dance and pretend to sing upon landing."

The masterpiece in the underappreciated genre of irregular plural humor comes from the National Puzzlers' League, the association of twisted geniuses who devise impossibly clever word problems. One kind of puzzle, the *falsie*, begins by finding a pair of words that look as if they are related by a morphological rule:

False iteration: *bus-rebus*, *bozo-rebozo*, *ally-really*

False predecessor: *lope-antelope*

False feminine: *butter-buttress*, *car-caress*, *under-undress*

False comparative: *ling-linger*

False plural (from Hebrew): *inter-interim*

The puzzle itself takes the form of a poem (called a *flat*) that uses a pair of falsely related words. The words are deleted from the poem and their locations are marked with placeholders. The object of the puzzle is to guess the pair of words from the context of the poem. The following flat by the puzzler known as Trazom (in real life Joshua Kosman, the senior music critic of the *San Francisco Chronicle*) contains a seven-letter singular noun in the place marked ONE

and its six-letter false plural in the place marked MANY. Try it (the answer is in the notes).

False Plural (7, 6)

Turn over on your side, my dear,
And tuck your foot behind your ear;
And I, meanwhile, will crouch like this
And give your neck a tender kiss.
Let's see now—let your arms go slack
And clasp your hands behind my back;
I'll reach around and drape my knee
Across your shoulder—goodness me!
I must confess, this is a stretch,
But honeybunch, you mustn't kvetch.
I know it hurts, I know it smarts—
But these arcane erotic arts
Don't yield their secrets right at first;
And now, I think, we're past the worst.
So please don't throw a ONE, sweet miss—
The MANY says we'll soon reach bliss.¹¹

Now we come to the irregular verbs. A menagerie of nearly two hundred words coming in many shapes and sizes, they are a vivid demonstration of how the human mind, reacting to the events of history, reshapes a language over centuries and millennia.¹²

The verbs *be*, *have*, *do*, and *go* are irregular in many of the world's languages. They are the most commonly used verbs in most languages and often pitch in as auxiliaries: "helper" verbs that are drained of their own meanings so that they may combine with other verbs to express tense and other grammatical information, as in *He is jogging*, *He has jogged*, *He didn't jog*, *He is going to jog*. Many language scientists believe that the meanings of these verbs—existence, possession, action, motion—are at the core of the meanings of *all* verbs, if only metaphorically. For example, the mind treats *telling him a story* as causing the

story to go to him resulting in him having it, and it treats *dying* as going out of existence.¹³

In English we saw that *be* stands out from all other verbs with its eight-way conjugation. Its irregular past-tense form stands out too. Together with *go*, it is the only verb whose past tense is a completely unrelated word, a relation that linguists call suppletion:

be—*was/were*—*been*
go—*went*—*gone*; also *undergo*, *forgo*

Suppletion arises from a merger of two verbs. Old English, spoken from about 400 to 1100, had three verbs for *be*: *beon*, *esan*, and *wesan*. They probably differed in meaning, with *beon* referring to permanent states and the other *bes* to temporary ones. (The distinction is similar to the one in modern Spanish between *ser* and *estar*: *Yo soy Americano* [I am American], a long-term trait, contrasts with *Yo estoy contento* [I am happy], a temporary state.) Adding to the surfeit, different sets of *bes* were used in different parts of England. In the Middle English period (1100–1450) they merged into one verb. As in a corporate merger, in a linguistic merger the workers scramble to fill a smaller number of positions, because a verb generally permits only one form in every slot in its conjugation. *Beon* supplied the base form *be*; *esan* supplied *am*, *is*, and *are*; *wesan* supplied *was* and *were*.

For mysterious reasons, in the Middle English period the verb *go* usurped the past-tense form of another verb, *wend* (as in *to wend one's way*), namely, *went*. Today the verb *wend*, bereft of its old past-tense form, has the regular past *wended*, but its original form followed a pattern that can be seen today in other irregular verbs, such as *bend*—*bent*, *send*—*sent*, and *spend*—*spent*.

Have, also irregular in many languages, is one of two English verbs that drops its final consonant and replaces it with a *d*:

have—*had*, *make*—*made*

Originally these were *haved* and *maked*, but enough lazy speakers swallowed the consonants that at some point in the Middle English period speakers didn't hear them and assumed that they were not there at all.

The verb *do* does something slightly different—it takes on a *-d*, and changes its vowel: *do*—*did*—*done*. Its participle form *done* (as in *You've done it again*) is a contraction of the verb with an old suffix, *-en*. (The same thing happens in

be—*was*—*been* and *go*—*went*—*gone*.) The *-en* suffix is found in about fifteen English participles (such as *spoken*, *sworn*, *chosen*, *blown*, and *written*), but the suffix is not attached by a rule. New verbs, such as the neologisms *fax*, *Bork*, *spam*, and *mosh*, never get *-en* participles; no one says:

I've already faxen it.
 That's the third nominee the Republicans have Borken this session.
 The company has spammen its customers with ads once too often.
 Not tonight, dear; I'm sore from having moshen all night.

Putting aside weird *be*, what do all these verbs—*had*, *made*, *did*, and the *bent*—*sent*—*spent* family—have in common? They all end in *t* or *d*. These, of course, are the same consonants that make up the pronunciation of the regular suffix *-ed*. About half the irregulars end in *t* or *d*, because they originally took some version of the regular *-ed* suffix but then fell off the regular bandwagon for one reason or another. These lapsed regulars, together with the regulars themselves, were dubbed *weak* in 1819 by Jacob Grimm of Grimms' Fairy Tales fame; Grimm was also one of the first historians of the Germanic languages. Grimm called the verbs “weak” because they were too wimpy to hold on to their own unique past-tense forms. We will meet the more macho *strong* verbs later in the chapter.

Some version of the weak past-tense suffix *-ed* can be found in all the Germanic languages, including English, German, Dutch, and the Scandinavian languages. The suffix originated in an ur-language, Proto-Germanic, spoken by a tribe that occupied most of northern Europe in the first millennium B.C. Linguists call it the *dental suffix* because it was pronounced with the tongue against the gum ridge behind the teeth.

Why didn't the weak verbs make life simple and just stay regular? It is because combinatorial rules of grammar have a cost, as we saw in chapter 1: They blindly join things together without looking at what they are made of, and thus can create ungainly chimeras. Two strange things can happen when a verb finds itself with a suffix grafted onto its rear end. One of them is illustrated by the largest class of irregular verbs in English, the no-change verbs:

hit—*hit*; also *slit*, *split*, *quit*, *knit*, *fit*, *spit*, *shit*
rid, *bid*, *forbid*
shed, *spread*, *wed*
let, *bet*, *set*, *beset*, *upset*, *wet*

cut, shut, put
burst, cast, cost, thrust
hurt

(Some of these verbs have alternative past-tense forms: irregular *bid-bade*, *forbid-forbade/forbad*, *spit-spat*, the mainly British *shit-shat*, and regular *slitted*, *knitted*, *fitted*, *wetted*, and *thrusted*.)

Note that all twenty-eight verbs end in *t* or *d*. Most of them arose in Middle English and Early Modern English (1450–1700) when the regular ending was often *-de* or *-te*. Throughout the language, *es* at the ends of words, formerly pronounced, were dropping like flies; the “silent *e*” in the modern spelling of words such as *bake* is a souvenir of the earlier period. Thus a form such as *hitte* got reduced to *hit*. But why did speakers stand by as these past-tense forms shrank into confusing copies of their stems, rather than making the verbs regular, which would have given them the more distinctive *hitted*?

If I may be permitted to psychoanalyze speakers who have been dead for centuries, it probably came from a widespread human habit: We don’t like to put or keep a suffix on a word that looks like it already has the suffix.¹⁴ In this case, people don’t like to put a version of *-ed* on a verb that already ends in *t* or *d*. Psycholinguists have offered several explanations. Perhaps speakers develop a stereotype for “past-tense form,” namely, “ends with *t* or *d*,” and unconsciously think that a stem that fits the stereotype has already been inflected and stop themselves from adding the suffix again. Perhaps when the mind assembles past-tense forms, it gets confused between the *it* or *ed* or *ut* that is already at the end of the stem and the *-t* or *-d* it is trying to add and merges them into a single sound, like the girl who said, “I know how to spell *banana*, but I don’t know when to stop.” Perhaps the suffix *-d* is applied, and the unpronounceable result, *hitd*, is cleaned up, not by the ordinary phonological rule that inserts a neutral vowel between the *t* and *d*, but by a special rule that deletes the *d*. Perhaps several of these explanations are correct.

In any case the no-extra-suffix habit is alive and well in modern speakers. The psycholinguists who jot down speech errors have found that people are prone to leaving out *-ed* on regular verbs that end in *t* or *d*. For example, they say, *So we test ‘em on it*, intending to say *tested*, or *That’s what I need to do*, intending to say *needed*.¹⁵ The same thing happens when people are brought into the lab, given a list of verbs, and asked to say them aloud in the past tense as quickly as they can.¹⁶ Children, too, don’t like to add *-ed* to verbs ending in *t* or *d*—they make their signature error, *breaked*, less often with verbs that end

in *t* or *d*, such as *hitted*, *putted*, *builded*, and *meeted*, than with verbs with other endings, such as *bringed* and *buyed*.¹⁷ These habits are leaving their mark on English as it continues to evolve: Even in careful speech and writing, many people use no-change past and participle forms for verbs like *bust*, *pet*, *shred*, and *tread*, as in *She got the fleas when she pet the dog* and *This is an area where few psychologists have tread*.¹⁸

The phobia of adding a surplus suffix extends beyond the past tense. Gardening scribes often cannot bring themselves to write *crocuses*, *gladioluses*, and *narcissuses* (as we learned in the *Gardeners’ Dictionary* entry for *Narcissus*), and write headlines such as “Hardy Gladiolus Have Long Been a Favorite,” and “Dutch Crocus Herald the Arrival of Spring.” (No doubt these are symptoms of a Latin-conscious *-us/-i* anxiety as well.) I have seen an ad for a sprayer that *fits all hose* and another one for *the pantyhose that last*, and still another announcing *All fax on sale*. People treat the *sh* sound as similar to *s*, leading the *Boston Globe* handyman to write about adjusting *window sash*,¹⁹ and leaving every professor baffled as to how to refer to more than one *prefrosh* (pre-freshmen). Many people have trouble keeping up with the *Joneses* and instead merely try to keep up with the *Jones*. When it comes to the possessive *’s*, hardly anyone follows the advice in Strunk and White’s famous style manual to refer to *Charles’s hat* (charlziz) or *the Jones’s car*; it’s usually *Charles’ hat* and *the Jones’ car*, both in writing and speech. And what do you say to someone who has a *daddy-long-legs* climbing up each shoulder?

When a word has a verbatim replica of a suffix inside it, rather than just a reminder of one, the attempt to add the real suffix often results in clumsiness or unintelligibility. When there is rain or snow or hail or thunder coming down from the skies, it is said to be raining, snowing, hailing, and thundering. What about when there is lightning? Is it *lightninging*? Not very likely, and some speakers snip out an *-ing* and say *It is thundering and lightning*. Many adjectives can be turned into adverbs by adding *-ly*, such as *softly*, *surely*, and *happily*. What about those adjectives that already end in *-ly*, such as *ugly*, *friendly*, *heavenly*, or *leisurely*? *Uglily*? *Friendlyly*? *Heavenlily*? *Leisurelily*? Pthack. (The *Atlantic Monthly*, perhaps hoping no one would notice, once ran a story entitled “Friendily Yours.”) Sometimes brand names can be turned into colloquial verbs for traveling or sending:

We Chevy’d up and down Main Street.

I FedExed the package last night.

Down to their last thirty dollars, they Greyhounded home.

Because of his fear of flying he Amtrak'd to New York.

But even if your frequent flyer plan is with United Airlines, it is unlikely that you have ever *Uniteded* to San Francisco.

Sometimes people can get into trouble by speaking as if a word that appears to contain an affix really does contain it. An interstate trucking company must have lost the business of the literate when it proudly painted its trucks with the slogan "Faster than rail, regular than mail."²⁰ Former President George Bush used to tell reporters that he spent his vacation *bonefishing*, leading them to wonder what the best bait is for catching bones, and presumably his heart was in the right place when he explained, "I hope I stand for anti-bigotry, anti-Semitism, anti-racism."

Back to the verbs. Repeated-suffix phobia is also the explanation for the class that originally contained *wend-went*:

bend-bent; also *send, spend, lend, rend, build*

These verbs devoice their final consonant, *d*, into *t*. They began as *bend + de*, and the double *d* was fixed by trimming the final consonant of the stem, yielding past tense *ben + de*. The extra twist is that the phonological rule that today turns *-d* into *-t* in words like *walked* and *passed* used to be triggered by words ending in *l, m, n,* and *v* as well. *Bende* became *bente*, which then lost its *e* to give us *bent*. The overeager *-d → -t* rule can also be blamed for these verbs:

burn-burnt; also *learn, dwell, spell, smell, spill, spoil*

The irregular forms ending in *-t* show the English language changing before our eyes: Most of them are on their way out. American speakers mainly use *burnt* as an adjective, not a past-tense form—*The toast is burnt because Bernie burned it*—and would not be caught dead saying *learnt, dwelt, spelt, smelt, spilt, or spoilt*. *Rent* is used only for emotional resonance, as in *The Vietnam war rent the fabric of American society*, and *lent* is giving way in American English to *loaned*, the past of *to loan*. In "Childe Harold's Pilgrimage" (1812) Byron describes a battlefield using three verbs in the class that range from the moribund to the dead:

The thunder clouds close o'er it, which when rent,
The earth is covered thick with other clay

Which her own clay shall cover, heaped and pent,
Rider and horse—friend, foe, in one red burial blent!

Blend-blent, of course, has become completely regular, as have most of the other verbs with *-t* in their past-tense forms, such as *wend-went, pen-pent, gird-girt, geld-gelt, and gild-gilt*. Like many obsolete irregulars, *gilt* and *pent* have left relics among the adjectives: *a gilt-edged book, pent-up energy*.



Another reason that regular forms can go to seed is the Los Angeles Lakers effect that gave us irregular plurals such as *feet* and *mice*. Grafting a suffix onto a stem can trigger changes in the pronunciation of the stem, and sometimes the change can stay in the word long after the trigger has vanished.

Many languages distinguish a vowel sound pronounced quickly from the same sound drawn out; they are called short and long vowels. The vowels traditionally called "short" and "long" in English, such as the ones in *bet* and *beet*, used to differ in this way, as we see in their spellings: The long vowel was symbolized by writing two short vowels in a row, as if it took twice as long to pronounce.

Starting around the year 1000, English speakers shortened their pronunciation of a vowel when extra phonetic stuff (such as a consonant or syllable) was added, pushing new consonants into the syllable.²¹ Here are some examples that have survived in modern English:

bone-bonfire
break-breakfast
child-children
Christ-Christmas
deep-depth
five-fifth
know-knowledge
sheep-shepherd
wide-width
wise-wisdom

Shortening a vowel is a natural reaction when material is added to the end of the syllable. A syllable is a unit of timing, taking up a constant tick of the

speech clock. If material is added to the end of a syllable, the vowel is often shortened to maintain the rhythm.²² This habit of pronunciation could easily have turned into a full-fledged rule. In his April 5, 1997 column, the language maven William Safire ventured that the pronunciation of *seminal* as “SEM-uh-null” in place of “SEE-muh-null” was an instance of academic bowdlerization—prissy professors covering up the fact that the word *seminal* comes from the word *semen*. Safire’s theory, however, would have to go to lengths worthy of Oliver Stone to explain why those professors, presumably hatching plots in their SEEminars, have also changed the pronunciations of *vanity*, *sanity*, *cleanliness*, *brevity*, and *criminal* to hide the fact that they come from *vain*, *sane*, *clean*, *brief*, and *crime*. All, of course, are products of a phonological rule in English that shortens vowels at the beginning of many three-syllable words.

Take a verb with a long vowel like *keep*. Add the regular suffix and spell it phonetically: *keep*t. Shorten the vowel in response to the extra stuff at the end. We end up with something pronounced *kept*—one of a number of modern irregular past-tense forms that would be regular but for their shortened vowels:

keep–*kept*; also *creep*, *leap*, *sleep*, *sweep*, *weep*

Add some other habits of Middle English speakers that we have come across—using *-t* more widely, dropping suffixes—and you understand many other irregular verbs in modern English:

feel–*felt*; also *deal*, *kneel*, *dream*, *leave*

bleed–*bled*; also *breed*, *feed*, *lead*, *mislead*, *plead*, *read*, *speed*, *meet*

hide–*hid*; also *slide*, *bite*, *light*, *alight*

flee–*fled*, *say*–*said*, *hear*–*heard*, *lose*–*lost*, *shoot*–*shot*

sell–*sold*; also *tell*, *foretell*

do–*did*

(As before, some of these verbs allow regular past-tense forms, such as *kneeled*, *dreamed*, *speeded*, *lighted*, and especially, *pleaded*. For some of the verbs—particularly *sell*, *tell*, and *do*—the reasons for the vowel changes are a bit more complicated.)

Kept, of course, isn’t simply *keep*ed pronounced with a clipped vowel; neither is *hid* just a short version of *hide* nor *shot* a short version of *shoot*. The pairs of vowels traditionally called “long” and “short,” and spelled as if they are

double and single scoops of the same sound, are in fact very different vowels. How did that happen?

The perpetrator is a process of language change that is the opposite of the various slurrings and swallowings and cutting of corners that we have seen so far. All of those changes make it easier for the speaker to speak but do nothing for the listener, who would rather have the speaker enunciate clearly. Sometimes listeners do get their way; speakers *enhance* the difference between a pair of vowels by adding, exaggerating, or embroidering each in a different manner.²³

For many centuries speakers of Old and Middle English enhanced the difference between short and long vowels by making the long vowels tense: that is, the muscle at the root of the tongue is tensed up, changing its shape and making the vowel in *great* sound different, as well as longer, than the vowel in *get*. Enhancement went wild, however, during the dawn of Early Modern English in the fifteenth century, when the pronunciation of the long vowels was scrambled in a linguistic revolution called the Great Vowel Shift. Before the shift, *keep* had been pronounced something like *cape*, *hide* like *heed*, *boot* like *boat*. After the shift, the English spelling of the long vowels no longer made much sense, nor did the pairings of “short” and “long” vowels in siblings-like *keep* and *kept*. Since the children of Early Modern English could not have heard a relationship between the vowels, the past-tense forms struck them as a ragbag that just had to be memorized outright, and so they remained for subsequent generations. Thus verbs that entered the popular language after the Great Vowel Shift, such as *peep* (1460) and *seep* (1790), and verbs whose pronunciations eventually drifted into rhyming with the *keep* verbs, such as *reap* and *heap*, did not undergo a vowel change; they remained intact when they first submitted to *-ed*, giving us *peeped*, *seeped*, *reaped*, and *heaped*, not *pept*, *sept*, *reapt*, and *heapt*.

Here is a small mystery: What is the verb that goes with the past-tense form *wrought*, as in *The Watergate scandal wrought great changes in American politics*, and the participle form in Judges 23:23, *What hath God wrought!*, quoted by Samuel Morse in the first intercity telegram? According to the theory that irregulars are pairs of memorized words, an irregular past-tense form could, in principle, survive in memory without a corresponding stem. *Wrought* appears to be an example: Most people have no idea what the verb is. Many guess *wreak* (based on an analogy with *seek*–*sought*) or *wring* (based on an analogy with *bring*–*brought*), but both guesses are wrong. The answer is *work*: *Wrought iron* is worked iron, and a person who is *all wrought up* is a person who is all

worked up. (Old theater saying: “Plays are wrought, not written.”) *Wrought* belongs to a family of verbs that replace their rhyming parts with *ought* or *aught*:

buy–bought; also *beseech*, *bring*, *catch*, *fight*, *seek*, *teach*, *think*

How do you get *wrought* from *work* or *sought* from *seek*? The connection is less mysterious when we realize that the now silent *gh* used to be pronounced, somewhat like the *ch* of *Bach*, *loch*, and *Chanukah*. Start with *work* (actually *wyrkan*, but I will use modern spellings to make the changes clearer). Add the suffix *-t* to get *wor^{kt}*. Soften the *k* sound to *gh*, yielding *wor^{ght}*—an old phonological trick to avoid the strenuous *-kt*. A vowel and an adjacent *r* often switched places in the history of English, because *r* sounds a lot like a vowel, which makes its order with respect to a vowel hard to hear. Thus *brid* became *bird*, *thrid* became *third*, *hross* became *horse*, and *wor^{ght}* became *wro^{ght}*. We no longer pronounce the *gh*, and recall that many English vowels were shuffled during the Great Vowel Shift (the vowel spelled *ou* was once pronounced *ō*), and that vowels often get shortened when a suffix is added (so *ō* becomes *o*). The result is *wro^{ght}* and the mystery is solved.



In the 1980s the irascible *New York Times* book reviewer Anatole Broyard wrote that he doubted that English had “any life left in it, any flavor or idiosyncrasy.” His colleague Maggie Sullivan followed up in a column of her own:

Anatole Broyard is right to sound the alarm. We are losing this idiosyncrasy; as a language changes, strong verbs tend to become weak. For example: although once shepherds shorn their sheep, sheep are no longer shorn, they are sheared.

This issue should arouse lovers of the English language. Weakening the verbs can only weaken the language itself. To keep English from becoming a feeble tongue, we must reinforce our verbs. Fortunately, I have come up with a two-part plan. First, we must not allow new verbs to enter the language in a weak state. We must ensure, for example, that *to clone* is established as *clone*, *clewn*, *clown*, as in: *Future generations of booksellers may reproach us for not having clown Joyce Carol Oates and Isaac Asimov. . . .* And to *gentrify* as *gentrify*, *gentrifo*, *gentrifum*, as in: *The newcomers gentrifo one block and now the whole old neighborhood is gentrifum.*

Since new verbs are few and far between, I offer the second part of my plan—creating new strong verbs. English has some strong verbs with unique patterns for their principal parts, such as *go*, *went*, *gone*. Individuality makes them particularly vulnerable. Their patterns would hold up better if each pattern had more representatives. If we create allies for our unique strong verbs, we can buttress them and increase their number. Here are suggestions for new strong verbs:

Conceal, *console*, *consolen*: After the murder, Jake console the weapon.

Subdue, *subdid*, *subdone*: Nothing could have subdone him the way her violet eyes subdid him.

Fit, *fat*, *fat*: The vest fat Joe, whereas the jacket would have fat a thinner man.

Displease, *displpose*, *displpösen*: By the look on her face, I could tell she was displpösen.

Sullivan’s plan to “strengthen” the language captures two hallmarks of the second kind of irregular verb in English, the so-called strong verbs. They belong to alliances with similar sounds, and despite this solidarity, they have been dwindling for millennia.

The families of strong verbs have a history stretching back more than 5500 years. Most of the languages of Europe, Iran, and the northern half of India, and many current and extinct languages of Turkey, western Asia, and China, show similarities in vocabulary and grammar that suggest they are descendants of a single language spoken by an expansive and mysterious prehistoric tribe. The most popular theory is that they were a late-neolithic farming people with domesticated horses, wheeled vehicles, and a military leadership, who expanded from a homeland in southern Russia around 3500 B.C.²⁴ An alternative is that they were the people that first brought farming to Europe, beginning in 7000 B.C. from a homeland in eastern Turkey.²⁵ Though we don’t know who they were or where they came from, we know a lot about how they spoke. Their language, Proto-Indo-European, has largely been reconstructed by historical linguists working backward from the commonalities in the daughter languages.²⁶

Many Indo-European languages have echoes of the strong-verb patterns seen in English, such as *bear–bore*, *tear–tore*, and *sink–sank*, *drink–drank*. Some of these verbs and their past tense forms actually existed in the ancestral language, such as *bher–bhor-* and *senk^w–sonk^w*. Proto-Indo-European apparently had a set of rules for forming the past tense, not by adding a suffix as in modern English, but by changing the vowels, as in modern Hebrew—a kind of rule called gradation, apophony, or ablaut. There were probably seven ablaut

rules, more or less as follows: If the verb has *ei* followed by a consonant, change the *ei* to *a*. If the verb has *e* followed by a vowel-like consonant, change the *e* to *ā*—and so on for the other five classes.

When the Indo-Europeans started to spill out over Eurasia, the daughter tribelets lost touch, and games of Broken Telephone began in each one. Eventually the language radiated into the ancestors of our familiar languages and language families such as Germanic, Romance, Slavic, Celtic, Greek, Iranian, and Sanskrit. For example, the verb *werg-* (to do) ended up in Germanic as *werkam* (work), and in Greek as *erg-* (action) and *org-* (tool), which eventually crossed over into English as *energy*, *organ*, and *orgy*. When a word meaning “do” turns into a word meaning “orgy,” the changes wrought by the chain of whisperers must have been considerable. It is remarkable that the seven classes of Indo-European strong verbs came through, tattered but recognizable, in Proto-Germanic, then in the West Germanic language spoken by the Angles and Saxons, and then in Old English, Middle English, and Modern English. That is why the strong verbs fall into clusters of similar-sounding forms today.

The rules themselves, however, did not survive. Imagine a rule that replaced *ī* with *a*, and suppose that people started pronouncing *ī* as *ī* in some verbs, *e* in others, and *ɪ* in still others, depending on the consonants following the vowel and many other factors. Children would have a hard time making sense of the rule, and at some point they would stop trying and simply memorize the past-tense forms as a list. By the time of Old English, the Indo-European vowel-change rules were extinct and their products had been mangled in different ways by the slings and arrows of outrageous fortune. At least a fifth of the verbs no longer obeyed the rules of their original class, and in the following centuries so many verbs joined, left, or switched classes that today the classes no longer correspond very well to the organization of the verbs in the minds of modern speakers.

Here is one Old English class, Class I, that has come through in recognizable shape:

rise—rose—risen; also *arise*, *write*, *smite*, *ride*, *stride*, *dive*, *drive*, *shine*, *strive*, *thrive*

The list highlights a key feature of the strong verbs. While dictionaries happily list irregular forms such as *smite—smote—smitten*, *stride—strode—stridden*, *strive—strove—striven*, and heaven help us, *thrive—throve—thriven*, in the minds of real English speakers these forms are muzzy: People vaguely recognize them

from books but are uncomfortable using them in their own speech and are tempted to default to regular forms like *smited*, *strided*, *strived*, and *thrived*. Sometimes strong and weak forms live side by side in a person’s mind, forming doublets like *strove* and *strived* or *dove* and *dived*.²⁷

Doublets usually arise when an irregular form (such as *strove*) hovers in a twilight zone in memory and people are not sure whether they have heard the form or are confusing it with a similar form, like *drove*. Other doublets arise for the same reason that you say *tomayto* and I say *tomahto*: Britain and America are divided by a common language. The British prefer *dived*, the Americans prefer *dove*, and people who encounter both dialects, such as Canadians, are unsure. Often the members of a doublet will diverge in meaning, grammar, or formality, like twins who strive not to be confused. *Shone*, for example, is intransitive (without a direct object), as in *The stars shone in the sky*, and a touch poetic, whereas *shined* is an everyday form that may be used in transitive sentences such as *Melvin shined his shoes*. (It would sound silly to say *Melvin shone his shoes*.) For many people regular *hanged* means “suspended by the neck until dead,” irregular *hung* merely “suspended.” Sometimes a muzzy participle will enjoy full vigor as an adjective, often with its own meaning. For example, *smitten* is doing fine as an adjective that means “infatuated,” not literally “walloped” (though the original metaphor is clear enough, and visible in related metaphors such as *stunning* and *lovestruck*).

Some of the past-tense forms originally in this class became muzzier and muzzier until they faded out entirely and their verbs became regular. *Abode* used to be the past tense of *abide* and today survives only as a noun meaning “residence.” No speaker of modern standard English uses *chide—chode*, *glide—glode*, *gripe—grope*, or *writhe—wrothe*, though some examples, such as *climb—clomb*, cling to life in rural areas of Britain and America. Many of the wayward verbs did not fall into the arms of regularity but were attracted to other irregular patterns. For example, the short vowel *ɪ* is common in participles like *driven*, *risen*, and *written*, and in many weak verbs, and it inspired *bit* and *hid* in the standard dialect of English. In nonstandard dialects we find *clim*, *writ*, *strid*, *smit*, *div*, *driv*, and the forms immortalized in the Negro spiritual “Joshua fit the battle of Jericho” and in the doggerel “Spring has sprung / The grass is ris / I wonder where the boidies is.”²⁸

The pairing of *ī* and *ō* in *rise—rose*, *drive—drove*, and other descendants of Class I can be seen, with variations, throughout the strong verbs, where *ī*-like vowels are frequently replaced by *ō*-like vowels:

find–found; also *bind, grind, wind*
freeze–froze; also *speak, bespeak, steal, heave, weave*
wear–wore; also *bear, forbear, swear, forswear, tear*
take–took; also *mistake, partake, forsake, shake*
wake–woke; also *awake, break*

Forsook and *hove* are pretty *recherché* these days, with *hove* appearing mainly in nautical contexts such as *The ship hove to*; other uses, such as *Irving hove his lunch*, could only be said in jest. Like the other strong classes, the *swear–swore* class used to embrace more verbs, but many defected to the regular side:

But unburied whiten the bones of the crew;
 Ah! would that the widow and orphan but knew
 The place where their dirge by deep billows is sighed,
 The place where unheeded, unholpen, they died.²⁹

Some of the old irregular forms survive in rural dialects, such as *help–holp*, *tell–tole*, *melt–molt*, and *swell–swole*, and others survive in adjectives in the standard dialect such as *molten* and *swollen*.

If you shorten both vowels of the *e–o* pattern you get:

get–got; also *forget, beget, tread*

which also beget some muzziness. The participle *has got* is British, *has gotten* American. As with many differences between the dialects, it was the Mother Country that corrupted the mother tongue; *gotten* was the form used in England when the first colonists left in the seventeenth century, and the Americans preserved it while it vanished in the British Isles. *Trod* and *trodden* sound vaguely Winnie-the-Poohish to American ears, because Americans seldom use the verb *to tread*: Where the British say *tread on*, Americans say *step on* (notwithstanding one of the slogans of the American Revolutionary War, “Don’t Tread on Me”). When *tread* is used, it is regular: *He treaded water*; not *He trod water*. *Begot* suffers because of the familiarity of *begat* in the King James Bible and the countless satires based on it.

Strangely enough, three common verbs undergo these vowel changes in reverse:

come–came; also *become, overcome* (compare *wake–woke, take–took*)

fall–fell; also *befall* (compare *get–got*)
hold–held; also *behold* (compare *swear–swore*)

Came came from a very old irregular whose origins are obscure, but *hold* (and maybe *fall*) really did get reversed. Originally *to hold* was *to held* (actually, *healdan*) with past tense *hold* (*heold*). Similarly, *fall* used to have the forms *feallan–feoll*. Some ancient, influential, and confused group of speakers managed to mix up these verbs with their past-tense forms. This is not as addled as it may seem; today people occasionally confuse the parts of a verb when the past tense or participle is more commonly used than the stem:

Even as environmentalists speak of a seamless web of life, and the artery advocates speak of a seamless city, the designs on the drawing board still rent the land from the sea and undermine its urbanity [from *rend–rent*].³⁰

The videophone is the same size as a regular phone but includes a 3.3 inch color screen with a tiny camera and lens. The . . . company hopes to smitten prospective buyers by renting the phones for less than \$30 a day [from *smite–smote–smitten*].³¹

REEBOK KICKS ITSELF OVER NAME WITH BAD FIT . . . For a company that made its reputation by helping to shod the women’s aerobics movement, the Incubus name would definitely seem out [from *shoe–shod*; Reebok had named a women’s running shoe *Incubus*, not realizing that the word refers to an evil spirit that has sex with women while they are asleep].³²

Producer Harvey Weinstein hoves his boorish bulk up to the mike for his moment in the sun for the callow “Shakespeare in Love”—but is miraculously sent packing by the *deus ex machina* of the orchestra [from *heave–hove*].³³

Similarly, *hoist* was originally the past tense and participle of *hoise* (as in *For ’tis the sport to have the enginer Hoist with his own petar*, from *Hamlet*), but it has since been reanalyzed as the stem in *hoist–hoisted*.

The following family is a freeze-frame of the process by which neat classes can get messier over the centuries:

blow–blew; also *grow, know, throw, draw, withdraw, fly, slay*

What do they have in common? All end with a vowel, and all begin with a cluster of consonants except *know*. In fact even *know* begins with a consonant

cluster in its spelling, and that tells a story. Spellings usually reflect old pronunciations, and the *k* in *know* was originally spoken aloud; the word was pronounced *k'nawa*. So these verbs used to be *completely* consistent. Owing to the disappearance of *kn* and *gn* at the beginning of spoken English words, one member no longer fit the membership requirements and had to be kept in the class by sheer stipulation. In the history of languages many law-abiding classes become more and more ragged as general pronunciation shifts mangle their members, until eventually the criteria become indiscernible to children and the words are memorized individually.

With only one nonconformist member thus far, the *blow* class has not yet disappeared, though it has suffered losses. *Slay-slew* has a biblical feel and may be on the way out, if we are to judge by recent usages such as *Burr slayed Alexander Hamilton in a duel*.³⁴ *Crow-crew* survives in the bookworm expression *The cock crew*; even small changes in the expression, such as *The rooster crew*, sounds peculiar, and *Harvey crew over his victory* is unintelligible. Regional dialects have added or preserved a few more, such as *show-shew*, *saw-sew*, *sow-sew*, and *snow-snew*; in 1942 the *Chicago Sun* wrote of the weather, *It blew and snew and then it thew*. These forms are rarely heard today, however, and the trend is in the opposite direction: attrition into the regular class. Children make errors such as *blowed* and *knowed* more often than for any other kind of irregular verb.³⁵ The journalist H. L. Mencken was an assiduous student of the vernacular speech of the United States and documented many common nonstandard past-tense forms in his magisterial volumes *The American Language*. Among them are *blowed*, *knowed*, *throwed*, *drawed*, and one made famous by a character in Harriet Beecher Stowe's *Uncle Tom's Cabin*. Theodore Bernstein, in *The Careful Writer*, comments on her oft-quoted words:

TOPSY

"In the absence of such reorganization, the city's court structure as a whole has just 'growed,' like Topsy"; "Like Topsy, that Government-held surplus of farm commodities 'just keeps growin'.'" Once and for all, Topsy's exact words, punctuated variously in different editions and in different books of quotations, were: "I 'spect I grow'd." No "just," no "jes'," no "growin'," no nuffin'. Anyway, Topsy, Queen of the Clichés, should drop dead. See Clichés.³⁶

A few verbs besides *came* take an *ā* in the past tense:

eat-ate; also *give*, *forgive*, *bid*, *forbid*, *lie*

Bade is a somewhat stilted past tense form of *bid* in the sense of "ask" or "command to," though not in the sense of poker, bridge, or defense contracts—no one says *He bade three clubs*. *Lie-lay* is a trap seemingly designed to lure speakers into errors and to provide material for the lamentations of language lovers (including me, in private moments). A recent article by Cullen Murphy in the *Atlantic Monthly*, "The Lay of the Language," was devoted to the verb,³⁷ and even the Muppets have been dragged into the controversy. In 1999 the talking doll Sing and Snore Ernie had to be reprogrammed after purists objected to his statement, "It feels good to lay down"³⁸ (the biggest hooaha over a talking doll since Barbie set back the cause of gender equality by whining, "Math is hard"). What's wrong with *lay*? Officially, it belongs to two verbs. One is an intransitive irregular verb, *lie-lay-lain*, meaning "recline":

Stem: *Please lie down and tell me about your childhood.*

Past tense: *He lay down on the couch.*

Participle: *He has lain down on the couch.*

The other is a transitive regular verb, *lay-laid-laid*, meaning "set down":

Stem: *Lay your cards on the table.*

Past tense: *He laid his cards on the table.*

Participle: *He has laid his cards on the table.*

Like Ernie, many casual speakers use *lay* for both—as in *I'm going to lay down*—and who can blame them? As if the sharing of *lay* in the two conjugations weren't confusing enough, the two verbs *ought* to be one, according to the grammatical logic of English. *Lay* means "cause to lie," and is one of a handful of verbs meaning "cause to X" that differ by a vowel from a related verb meaning "to X." The others are *sit-set*, *rise-raise*, *fall-fell* (as in *to fell a tree*), and believe it or not, *drink-drench*. In most other cases, the verb that means "to X" and the verb that means "cause to X" sound the same:

The stick leaned against the house.

I leaned the stick against the house.

The planter stood on the deck.

I stood the planter on the deck.

The baby sat on the bed.
I sat the baby on the bed.

The “ungrammatical” intransitive *lay* follows the pattern of *lean*, *stand*, and *sit* perfectly. Many purists believe that intransitive *lay* is a recent corruption, disseminated by rock lyrics such as Bob Dylan’s *Lay Lady Lay* and Eric Clapton’s *Lay Down, Sally*. But a rule of thumb in language is that any so-called corruption that occurs frequently enough for the guardians to notice it will turn out to have been common in the language for a century or more. Intransitive *lay* was unexceptionable in the seventeenth, eighteenth, and nineteenth centuries; for example, in 1812 Byron wrote, “There let him lay” in “Childe Harold’s Pilgrimage.”³⁹ The historical linguists Thomas Pyles and John Algeo report:

The brothers H. W. and F. G. Fowler (1931, p. 49) cite with apparently delighted disapproval “I suspected him of having laid in wait for the purpose” from the writing of Richard Grant White, the eminent nineteenth-century American purist—for purists love above all to catch other purists in some supposed sin against English grammar.”⁴⁰

Another long-term trend reshaping the English language is most apparent in our final class of irregular verbs, illustrated in the greeting card by Suzy Becker on the opposite page.

The *ing-ang-ung* pattern often is generalized in dialects and in affectations of dialects, as in the jocular *Who would have thunk?* In 1998 the Texan columnist Molly Ivins entitled a book *You’ve Got to Dance with Them What Brung You*, allegedly a backwoods aphorism though more likely an urbanite’s attempt at hick-chic. The baseball pitcher and sportscaster Dizzy Dean was said to have narrated a play as follows:

The pitcher wound up and flang the ball at the batter. The batter swang and missed. The pitcher flang the ball again and this time the batter connected. He hit a high fly right to the center fielder. The center fielder was all set to catch the ball, but at the last minute his eyes were blound by the sun and he dropped it!⁴¹

Dave Barry, defending himself against enraged Neil Diamond fans after making a joke at the singer’s expense in a prior column, describes the results of a reader survey:

Unfortunately, a lot of survey voters are not so crazy about Neil’s work, especially the part of “Play Me” where he sings, “. . . song she sang to me, song she brang to me . . .” Of course I think those lyrics are brilliant; however, they brang out a lot of hostility in the readers.

The *ing-ang-ung* pattern came down to us from another class of strong verbs in Old English, Class III, which included *singan-sang-sungen*. Many modern verbs follow it to varying degrees:

ring-rang-rung; also *sing*, *spring*, *drink*, *shrink*, *sink*, *stink*, *swim*, *begin*
cling-clung; also *fling*, *sling*, *sting*, *string*, *swing*, *wring*, *slink*, *stick*,
dig, *spin*, *win*
run-ran-run
hang-hung
strike-struck, *sneak-snuck*
sit-sat, *spit-spat*

Most of the *ing-ang-ung* verbs end in *-ing* or *-ink*. Two of the others deserve comment.



Begin has the distinction of being the only common irregular verb that is neither monosyllabic nor built around a monosyllabic root. (The common, Anglo-Saxon words we use every day tend to be monosyllables, and the irregular verbs are no exception.) *Begin* is formed with the prefix *be-*, as in the similar irregulars *become*, *befall*, *beget*, *behold*, *beset*, and *bespeak*. In *begin*'s case, however, the residue, *-gin*, is not an English word; it came from a now-defunct Proto-Germanic verb meaning "open." (There are two other irregular past-tense forms, both somewhat unusual, whose stems cannot stand alone as verbs: *forsake*–*forsook* and *besech*–*besought*.)

Snuck has the distinction of being the most recent irregular to enter the standard language, with a first citation in the *Oxford English Dictionary* from 1887. According to a recent survey, most younger Americans have no problem with *snuck*, though most older Americans frown on it.⁴² William Safire quotes a letter from Doris Asmundsson, a professor emerita of English: "Words like *creak*, *critique*, *eke*, *freak*, *leak*, and *tweak* do not, in the past tense, become *cruck*, *crituck*, *uck*, *fruck*, *luck*, and *twuck*. Why then *snuck*? Eventually a *sneaker* might turn into a *snucker*."⁴³ According to one theory, *snuck* sneaked into English via sound symbolism. Its connotation of quickness, furtiveness, and mild disreputability brought to mind the sound pattern of *slunk* and *suck*, especially since all three end in a suitably crisp *k*.⁴⁴ A less far-fetched explanation is that *sneak* is close in pronunciation to *sting*, *strike*, *dig*, and especially *stick*—an *ī* is just a lax, short *ē*, and *n* is basically *t* or *d* pronounced through the nose, as any cold-sufferer can tell you. The failure to rhyme with *creak* and *tweak* was no impediment, because similarity in the gestures of articulation matter more than similarity in sound, and that makes it tempting to analogize *stick*–*stuck* to *sneak*–*snuck*.

Many dialectal past-tense forms that don't rhyme exactly with *cling* or *slink* still take the *ū* vowel in the past tense. Mencken and others report *climb*–*clumb*, *shake*–*shuck*, *take*–*tuck*, *dive*–*duv*, and *drive*–*druv*, also heard in the English proverb "Sussex won't be druv." One speaker described what they used to do to endangered species in the olden days as follows: *They killed 'em and skun 'em out*. Dizzy Dean was famous for saying *He slud into second*, and some baseball fans say, "If Dykstra hadn't dropped the ball, the runner wouldn't have tug" (tagged).⁴⁵ On the following page is another common example.

The *ing*–*ang*–*ung* verbs are a bellwether of a millennium-old and still ongoing trend in the English language. In the fourteenth century the egalitarian preacher John Ball roused the rabble with the slogan "When Adam delved and Eve span / Who was then a gentleman?" *Span* was the past tense of *spin*, fol-

ARLO & JANIS® by Jimmy Johnson



ARLO & JANIS reprinted by permission of Newspaper Enterprise Association, Inc.

lowing the *i*–*a*–*u* pattern of verbs such as *sing*, *swim*, and *begin*. But eventually the participle *spun* usurped the past-tense slot, relegating *span* to the dustbin of history. These takeovers are still going on. When the Walt Disney corporation released a film called *Honey, I Shrank the Kids*, English teachers were up in arms: It should be *Honey, I Shrank the Kids*, they said. Nonetheless, most people say *shrunk*, *sprung*, *sunk*, and *stunk*, not *shrank*, *sprang*, *sank*, and *stank*. (Some go the other way: At an infamous moment in the O. J. Simpson murder trial, the prosecutor Christopher Darden said hopefully, "The gloves appear to have shrank somewhat.")

The teachers are fighting a losing battle because even the language mavens are losing their grip on the distinction. William Safire got an earful from the Gotcha! Gang and the Uofallpeople Club when he wrote, "*Trivialize* had its moment in the vogue-verb sun, until the usage of this older verb shrunk to the very occasional."⁴⁶ The *Boston Globe*'s language maven, Jan Freeman, wrote that she once did a double-take upon hearing *They sort of sprang it on me*, momentarily thinking it was incorrect.⁴⁷

Shrank, together with the other *ank* and *ang* words, is under assault from two directions—from its own past participle *shrunk*, and from the many *ing* verbs that have already lost their *angs* and really do take the *ung* form in the past tense as well as in the participle: *He slung* (not *slang*) *the hash*; *They strung* (not *strang*) *him up with a rope*, *He flung* (not *flang*) *the ball at the batter*. Surely and steadily, *ing*–*ung*–*ung* is displacing *ing*–*ang*–*ung*, part of a larger erosion of the distinction between participles and past tense forms throughout the verb system.

Regular verbs fail to distinguish pasts and participles at all—I *walk*, I *walked*, I *have walked*—and fewer than half of our irregular verbs continue to distinguish them; most are like *mean*–*meant*–*has meant* or *find*–*found*–*has found*. In nonstandard dialects the distinction is even feebler. *I seen it* and *A man come into the bar* are absolutely standard outside the upper and middle

classes, even in urban areas, and *He begun to cry*, *She done it*, and *They gone home* are also common. In the early decades of the twentieth century, Mencken reported the past-tense forms *div*, *driv*, *riz*, *swole*, *taken*, *thrown*, and *writ*, and the participle forms (*has*) *ate*, *blew*, *broke*, *did*, *drank*, *drive*, *froze*, *gave*, *rode*, *rose*, *ran*, *stole*, *swam*, *took*, *tore*, *woke*, *wore*, and *wrote*. Nonetheless, many people assume that the erosion is a recent development:

Dear Ann Landers:

Have Americans forgotten there is such a thing as verb tense? I am shocked when I hear people say “woulda came,” “coulda went,” “shoulda did,” “woulda took,” “had went,” “hadn’t came,” and so on.

Don’t they realize “woulda” and “coulda” are slang versions of “would’ve” and “could’ve”—which are contractions for “would have” and “could have”?

I heard a narrator say, “I seen” in a political commercial, and a TV reporter say, “We haven’t spoke.” . . . A TV anchorwoman said, “had threw it” and “between you and I.”

I am a secretary for almost 50 years and am thankful that, with only a high-school education, my English is impeccable. You will do a lot of folks a big favor if you print this letter and bring it to their attention.

E. E.
Wood Ridge, N. J.

Dear E. E.:

Thanks for taking the time and trouble to write. I shoulda thunk to tell them off myself.⁴⁸

Confusions of past and participle forms are easy to explain. Some may originate in mishearings. As E. E. pointed out, the auxiliaries *has* and *have* that signal the perfect construction are often contracted to *he’s*, *we’ve*, *could’ve*, *should’ve*, and *would’ve*, or even *coulda*, *shoulda*, and *woulda*. (Anyone who has graded student papers or dipped into internet discussion groups is also familiar with *could of*, *should of*, and *would of*.) That makes the *haves* easy to miss in rapid speech; *He’s seen it*, in particular, is easily reanalyzed as *He seen it*.

Yet the main reason for the decline of the *ang-ung* distinction is that *all* distinctions in English inflection have been declining for the past thousand years;

syntax has been shouldering the load formerly borne by morphology. Old English and Middle English had separate verb forms not only for present, past, and participle, but also for different persons (*I*, *you*, *he/she*) and numbers (singular and plural) within the past tense. The past forms for *sing*, for example, would have been:

I sang	We sungon
Thou sunge	You sungon
He/she sang	They sungon

When the person and number distinctions collapsed, every verb had to end up with a single past-tense form, and a game of musical chairs broke out, with the different stems competing for the remaining chair. With some verbs the singular won, such as *sing-sang-sung*; with others the plural won, such as *sling-slung-slung* (the past plural was usually similar to the participle—the phenomenon of syncretism we met in chapter 2). Another free-for-all took place among the participles of the verbs that kept their *-en* and had to grab a stem from the collapsing conjugation. Some took the stem of the base form, such as *take-took-taken*, others took the stem of the past form, such as *break-broke-broken*, and still others kept their own stem, such as *swell-swelled-swollen*. Some participles can jump ship to another pattern: If you apply the *break-broke-broken* pattern to *shake* and *take*, you get the somewhat cutesy *shooken* and *tooken*. I have been advised that *tooken* has become standard in Generation X circles, but if it is, do not blame it on their ethos of ironic detachment; it was used as early as 1946 in “Put That Kiss Back Where You Found It,” a song recorded by Benny Goodman: “Took it when I wasn’t lookin’/ And my heart you’ve also tooken.”⁴⁹ The steady erosion of distinctions in English inflection helps us understand why we continue to be confused by verbs such as *shrink* and *spring* in the second millennium after the end of Old English.

For the sake of completeness, here are the remaining irregular forms. *Shorn* and *swollen* belong to a small group of verbs that are regular except for their participles:

swell-swelled-swollen, *shear-sheared-shorn*
show-showed-shown; also *sow*, *sew*, *prove*, *strew*

(A few other irregular participles were orphaned from their verbs and survive only as adjectives, most of them somewhat unusual: *bereft*, *unbidden*, *clad*,

cleft, cloven, drunken, forlorn, girt, gilt, misbegotten, hewn, beholden, laden, molten, mown, pent, misshapen, clean-shaven, shod, sodden.) I couldn't figure out where to put these:

beat—beat—beaten
choose—chose—chosen
see—saw—seen
stand—stood; also understand, withstand

The *stand—stood* pattern is heard in the oft-cited plaint of the fight promoter Joe Jacobs, "I should have stood in bed," and in the modal auxiliary verb

can—could

which retains a present-past contrast in usages like *I can't polka now, but I could before I broke my leg*. Other pairs of modal auxiliaries—*may—might*, *will—would*, and *shall—should*—began life as different tenses of the same verb, but the couples divorced long ago and *might*, *would*, and *should* are no longer past-tense forms.

Exactly how many irregular verbs are there in the English language today? If we don't double-count prefixed families such as *get* and *forget*, don't count dialectal form such as *drug* and *brung*, do count verbs that are irregular either in standard American or standard British English, and do count the muzzy but widely recognizable forms, we end up with 164 modern irregular verbs: 81 weak (ending in *t* or *d*), 83 strong. Compare this to Old English, with 325 strong verbs alone, and it is clear that English is indeed becoming "weaker." In later chapters we will see whether the surviving but endangered irregular verbs are sustainable.



We have seen how the weak past-tense forms can be traced to Proto-Germanic about 2000 years ago, and the strong forms can be traced back to Proto-Indo-European at least 5500 years ago. But where did they come from? They certainly were not designed by a committee, and presumably did not arise from divine revelation. No one knows the answer, but a few brave linguists have speculated.

The dental suffix in Proto-Germanic, the ancestor of our *-ed*, may have come from a reduced form of the verb *to do*.⁵⁰ Many languages use an empty

verb like *do* as an auxiliary verb that carries information about the statement as a whole, such as tense, degree of completion, and negation. Indeed, Modern English uses *do* for that purpose in *yes-no* questions (*Do you want to dance?*) and in negations (*Alice doesn't live here anymore*). In the history of a language, prefixes and suffixes often arise from the erosion of verbs such as *do*, *take*, *be*, and *have*, a process called grammaticalization.⁵¹

If the dental suffix came from *do*, it would explain why it has the sound *d* or *t*. In Proto-Germanic, *do* could come after a noun or another verb, very roughly like *He hammer-did* or *She walk-did*. The *do* could have eroded to the stub *d* and attached itself to the verb, giving us the ancestor of *-ed*.

The theory also explains why *-ed* has become the *regular* suffix, applying freely to any new or strange verb. The phrase containing *do* and a verb would have been created by the rules of syntax, the combinatorial system par excellence, which allows almost anything to combine with anything else. A promiscuous auxiliary verb would have been a natural ancestor of a promiscuous suffix: Just as a verb like *do* can combine with any verb at all (*He did abandon*, *He did abate*, *He did abbreviate*, and so on), so its descendant *-ed* might have retained this habit, allowing it to combine with any verb at all (*abandoned*, *abated*, *abbreviated*, and so on).

The Indo-European ablaut or vowel-change patterns, the ancestors of our strong verb forms, change an *e* (a sound between *Ed* and *aid*) or a neutral vowel to *a* (as in *father*), or to *ō* (as in *hoe* or *horse*). The *e* is pronounced with the tongue hump toward the front of the mouth, the *a* and *ō* with the tongue low and toward the back. This contrast, between a higher front vowel and a lower back vowel, survives in the majority of modern English irregular verbs. The base forms have sounds like *ě* and *ē* and *ĭ* and *ī* and *ā*, and the past-tense forms have sounds like *ǎ* and *ǒ* and *ō* and *ŭ* and *ǔ*.

That may not be a coincidence. Three of the great linguists of the middle decades of the twentieth century, Roman Jakobson, Jerzy Kuryłowicz, and Morris Swadesh, noticed that in many languages the vowels pronounced with the tongue high and at the front of the mouth tend to be used for the basic forms of nouns and verbs (such as the singular form of a noun and the infinitive of a verb), whereas the vowels pronounced with the tongue lower and farther back tend to be used for the specially marked forms (such as plural nouns and tensed verbs).⁵² Moreover, the higher and farther front vowels have different connotations from the lower and farther back vowels in pairs of contrasting words. The high front vowels come first in expressions such as *pitter-patter* and *dribs and drabs*; we don't say *patter-pitter* or *drabs and dribs*. And in pairs such

as *this* and *that*, *here* and *there*, and *me* and *you*, the higher and farther-to-the-front vowels are found in the word that means “self” or “near the self,” the lower and farther-to-the-back word means “other” or “far from the self.” That is true not only in English but in many families of languages.⁵³

Perhaps this ubiquitous vowel contrast is a case of sound symbolism. The linguist Roger Wescott has pointed out that high front vowels are pronounced with a constricted mouth cavity and the tongue close to the visible part of the vocal tract, whereas low back and central vowels are pronounced with a large mouth cavity and the tongue buried from view. That may call to mind the conceptual distinction between presentness and pastness. Pastness may remind people of a cavity or space, because a past event is separated by an interval of time from the present moment, and metaphorically speaking time equals space. It may also remind people of remoteness or distance, because metaphorically speaking long ago equals far away. Perhaps as Indo-European was developing, speakers vaguely felt that lower and farther back vowels fit better with the concept of an event separated in time from the present, and that higher and farther front vowels fit better with an event in the here and now.⁵⁴ Of course, the Indo-Europeans had to pick *some* vowel contrast if they were to mark tense with a vowel, and for all we know they could just have easily gone the other way. But the fact that the vowel contrast appears in many unrelated languages with similar roles, and was preserved and embellished in our own 5500-year game of Broken Telephone, hints that it might have some semantic resonance for human minds.

IN SINGLE COMBAT

Now that we know all about regular and irregular verbs, how well does the words-and-rules theory hold up? In some ways quite well; in other ways not so well. According to the theory, the ingredients of language are a list of memorized words, each an arbitrary pairing between a sound and a meaning, and a set of productive rules that assemble words into combinations. Regular and irregular forms exemplify the two ingredients: Regular forms are generated by rule, irregular forms are memorized by rote.

The dissection of language in chapter 2 showed that the organs that secrete regular forms are elegant combinatorial systems, just as the theory had led us to expect. But the excavation in chapter 3 showed that the depository for irregular verbs is not disorganized and inert, which is what the theory had led us to expect. The irregular verbs are shot through with patterns:

blow–blew, grow–grew, know–knew, throw–threw
bind–bound, find–found, grind–ground, wind–wound
drink–drank, shrink–shrank, sink–sank, stink–stank
bear–bore, swear–swore, tear–tore, wear–wore

This is not what we would expect if they were a laundry list of arbitrary items. Suppletive pairs such as *go–went* and *be–was* are the exception, whereas if the irregulars were truly acquired one by one they could just as easily have been the rule.