Pinker, Steven (1999): Words and Rules. The Ingredients of Language. New York: Basic Books.

7

KIDS SAY THE DARNEDEST THINGS

Moving its translucent mass through the watery
shadows of the dock and then, past the dock (some-
thing so real which now is not), the jellyfish
swam in its slow float while we (I and my daughter,
then just three) ran back and forth predicting that limp
pink gleam and each embodiment it would seem.

"A jello umbrello!" she began and turned
to me expectantly. Censoring (an after-
birth, broken veins, or Medusa's myth, the monstrous
queen made mortal and mother), I stood in silence
until it ended with a shout: the jelly-
fish glided out. Now months have passed, but surprise!

"The jellyfish was in my eyes!" Caroline calls
while caught between depth and surface of a dream.
"It bleeded and it singed!" Her conjugations
soon will exact simple irregularities
and tensing will be not verbs, but time's tentacles
untangling her parachute, waving at me.[1]

—*Susan Kinsolving, "The Jellyfish," 1999*

Grammatical errors like bleeded and singed have long epitomized the innocence and freshness of children's minds. The errors are acts of creation, in

which children lift a pattern from their brief experience and apply it with impeccable logic to new words, unaware that the adult world treats them as arbitrary exceptions. In A Dark-Adapted Eye, the novelist Barbara Vine introduces an unlikable child by remarking, "He would refer to 'adults' instead of 'grown-ups,' for instance, and get all his past tenses right, never saying 'rided' for 'rode' or 'eated' for 'ate.'"[2]

Children's errors with irregular verbs also have been prominent in debates on the nature of language and mind. The neurologist Eric Lenneberg pointed to the errors when he and Noam Chomsky first argued that language was innate; the psychologists David Rumelhart and James McClelland set them as a benchmark when they first argued that language could be acquired by generic neural networks. Psychology textbooks cite the errors to rhapsodize that children are lovers of cognitive tidiness and simplicity; researchers who study learning in adults cite the errors as a paradigm case of the human habit of overgeneralizing rules to exceptional cases.[3]

Nothing is more important to the theory of words and rules than an explanation of how children acquire rules and apply them—indeed overapply them—to words. The simplicity of these errors is deceptive. As we shall see, it is not easy to explain why children start making them, and it's even harder to explain why they stop.

~

Overgeneralization errors are a symptom of the open-ended productivity of language, which children indulge in as soon as they begin to put words together. At around eighteen months children start to utter two-word microsentences like *See baby* and *More cereal*.[4] Some are simply telegraphic renditions of their parents' speech, but many are original productions. "More outside!" says a tot who wants to play in the park. "Allgone sticky!" says another after his mother has washed jam off his fingers. My favorites in the data from my own lab are "Small loud" after someone had turned down the stereo, and "Circle toast!" shouted repeatedly to uncomprehending parents who couldn't figure out that the child wanted a bagel.[5]

By their twos, children produce longer and more complicated sentences, and begin to supply grammatical morphemes such as *-ing*, *-ed*, *-s*, and the auxiliaries.[6] Sometime between the end of the second year and the end of the third year, children begin to overgeneralize *-ed* to irregular verbs. All children do it, though parents don't always notice it. My sister told me that her son Carl never made this kind of error, and as if to contradict her, he said *sticked* in my presence a minute later. When children are old enough to sit still in experiments, they pass the *wug*-test: After hearing that a man knows how to *rick* or *bing*, they say that yesterday he *ricked* or *binged*.[7]

Children regularize almost anything they can. They put *-ed* not only on irregular stems, as in *breaked* and *eated*, but on irregular past-tense forms, as in *broked* and *ated*. They put it on their own neologisms, such as *poonked*, *lightninged*, and *spidered*. They put it on past-tense forms that already *have* a suffix, as in *sweepened*, *presseded*, and *My brother got sick and pukeded*.[8]

The past tense is not the only source of irregularity in English, and it is not the only regular pattern children overgeneralize. Alongside past tense errors such as *breaked* and *putted* we find plural errors such as *mans*, *foots*, *tooths*, and *mouses*.[9] Three English verbs are visibly irregular in the third-person singular present tense, and children overgeneralize *-s* to all three:

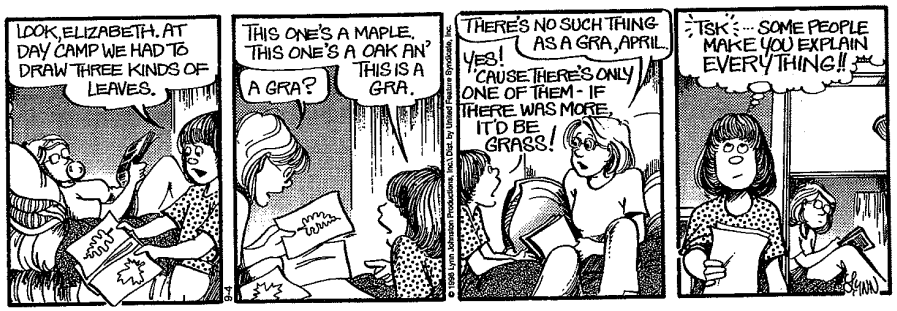
> He just haves a cold.
> She do's what her mother tells her.
> No, she be's bad, then she be's good, OK?[10]

The suffixes *-er* and *-est* turn many adjectives into comparative or superlative forms. It's easy to forget that the rule has exceptions until we hear children adding suffixes to them. They overgeneralize the suffixes to polysyllabic adjectives, as in *specialer* and *powerfullest*, and to a handful of suppletive irregulars:[11]

THE FAMILY CIRCUS. Reprinted with special permission of King Features Syndicate.

Children often generalize from *fourth, fifth*, and *sixth* to *oneth, twoth*, and *threeth*, or sometimes *firstth, secondth*, and *thirdth*. They leap from *myself, yourself*, and *herself* to *hisself*, and from *ourselves* and *yourselves* to *theirselves*. I have heard of one child who used *its* rather than *them* as the plural of the pronoun *it*, and another who liked drawing *rectangles, triangles*, and *cirtangles* (circles).[12]

Children are overzealous grammarians not only in applying inflections in their own speech but also in analyzing them in the speech of others. They have little choice. Children are never given grammar lessons presenting *-ed* or *-s* with lists of stems to conjugate or decline; they must mentally snip the suffixes out of the full, inflected words they hear in conversation. As they are figuring it out, they occasionally snip too eagerly and come out with strange back-formations:



FOR BETTER OR WORSE reprinted by permission of United Feature Syndicate, Inc.

I suspect that comic strips showing a child making a speech error are usually based on real-life instances known to the cartoonist; in almost every example I have seen, similar errors have been documented in the scientific literature. Alan Prince studied a girl who, like April, was delighted by her discovery that *eats* and *cats* were really *eat* + *-s* and *cat* + *-s*. She used her new suffix snipper to derive *mik* (mix), *upstair, downstair, clo* (clothes), *len* (lens), *sentent* (sentence), *bok* (box), *brefek* (from *brefeks*, her word for breakfast), *trappy* (trapeze), even *Santa Claw*.[13] Another child, overhearing his mother say they had booze in the house, asked what a "boo" was. One seven-year-old said of a sports match, "I don't care who they're going to verse," from expressions like *the Red Sox versus the Yankees*.[14]

We laugh, but adults do the same thing, or at least our ancestors did. *Cherry* is a back-formation from *cerise*, and *pea* is the invented singular of the mass noun *pease*, as in the nursery rhyme "Pease porridge hot, pease porridge cold." (Perhaps someday a grain of rice will be known as a *rouse*.) Many people have

to be reminded that there is no such thing as a *kudo*: The noun *kudos* is singular, from the Greek word for glory.

A striking feature of children's past-tense errors is that they appear, sometimes suddenly, after long stretches in which the children use the past tense correctly when they use it at all. A child might say *sang, went*, and *heard* for many months before coming out with *singed, goed*, and *heared*.[15] In a sense, the child gets worse before getting better; if the percentage of past-tense forms of irregular verbs that are correct is plotted over time, the shape of the graph looks a bit like a U. "U-shaped development" fascinates child psychologists because with almost anything else you measure, children get better as they get older.[16] No one considers childhood to be a period of decline (that comes later), so the newly appearing errors are taken as a sign of a *reorganization* in the child's mind. A laundry list of disconnected items suddenly reveals itself as having a pattern, and the child extracts the pattern and applies it across the board.

In the case of the past tense, children have a smattering of regular forms such as *played* and *used* before they make their first error with an irregular, and they use them correctly to talk about events in the past.[17] Presumably they have memorized those forms as indigestible chunks and use them like any other word, with the "pastness" simply being part of their meanings.

At a certain point a child notices (not consciously, of course) that many words come in ever-so-slightly different versions: *walk* and *walked*, *use* and *used*, *play* and *played*, *push* and *pushed*. Logically speaking, these could be interpreted as meaningless variations in pronunciation or speaking style, yet something impels children to seek a principle behind the variation. By subtracting *walk* from *walked*, *push* from *pushed*, and so on, a child can isolate *-ed*. By correlating its use with its meaning—that is, noticing that Mom and Dad use *-ed* when describing events that are over and done with—the child can infer that *-ed* means "past tense." This synopsis brushes aside many complexities, such as how the child knows to look out for "present–past" instead of "hot–cold," "indoor–outdoor," "good mood–bad mood," and hundreds of other interesting distinctions. It also sweeps aside how a child deduces that the rule is obligatory: You can't say *I already eat breakfast this morning*, even though the meaning would be clear. Yet children do succeed, and once the rule has been discovered they can feed any verb into it, regular or irregular. They now can say *goed* and *heared* and *bleeded* and *singed* in situations where earlier they might have said *went* and *heard* and *bled* and *sang*.[18]

~

Unfortunately, the rule-epiphany theory by itself cannot explain why children make errors like *bleeded* and *singed*. I have said that children start saying *bleeded* and *singed* because they have acquired an "Add *-ed*" rule. But adults have an "Add *-ed*" rule too, and we don't say *singed*. (If we did, we wouldn't call the child's form an error.) Something important is missing: the difference between children and adults, and how children overcome the difference as they grow up.

A first guess is that children become adults because language development is driven by communication: Children improve their language in directions that allow them to communicate their wishes more effectively. Wrong. There is nothing unclear about the meaning of *bleeded* or *singed*. In fact, as long as children make these errors, their language is more communicative than adults'. English has about twenty-five irregular verbs that don't change their forms in the past tense, such as *cut*, *set*, and *put*. These verbs are ambiguous between past and nonpast: *On Tuesday I put the trash out* could mean last Tuesday, next Tuesday, or every Tuesday. The childlike form *On Tuesday I putted the trash out* could mean only a preceding Tuesday. A language is certainly a powerful tool for communication, but children could not acquire its details by figuring out which ones help in communication; they learn the whole language, with all its strengths and weaknesses, because they just can't help it.

A second guess is that we adults don't say *bleeded* and *singed* because we never hear other adults say them. Wrong again. Adults say lots of things they never hear other adults say. New verbs constantly enter the language—*to diss, to snarf, to fax, to mung, to wild, to flame, to mosh*—and an adult who learns *diss* in the present tense does not have to wait to hear someone say *dissed* before using it in the past tense. If adults say *dissed* even though they have never heard it, they should be willing to say *singed* even though they have never heard it.

The reason adults avoid making regularization errors is not that the error has never been heard; it's that the irregular counterpart *has* been heard. There must be a component of adult psychology that causes the experience of hearing an irregular form such as *sang* to inhibit the application of the *-ed* rule to that item. As noted in chapter 5, this component is called blocking: A specific form in the mental lexicon blocks the application of a general rule that would express the same grammatical notion (past tense, plural, and so on), perhaps through an inhibitory link from the lexicon to the rule.[19] Thus *sang*, listed as a past tense of *sing*, blocks the past tense rule, preempting *singed*; *geese*, listed as the plural of *goose*, blocks *gooses*; *better*, listed as a comparative of *good*, blocks *gooder*.

Perhaps, then, children lack the blocking principle and have to learn it. But how? To learn the blocking principle children first would have to know that forms like *singed* are ungrammatical. Remember, not hearing other people say *singed* isn't enough, because other people don't say *wugged* either, and may not say *munged* or *flamed*, but people do not avoid the unheard past-tense forms.

The only way for children to know that *singed* is ungrammatical is to use it and get a negative feedback signal from their parents—a correction, a frown, a puzzled look, or a non sequitur as a response. Information about what is *not* in a language is called *negative evidence*, and it is one solution to what linguists call "the logical problem of language acquisition": how a child could, in principle, learn an entire infinite language from a finite sample of the behavior of its speakers.[20]

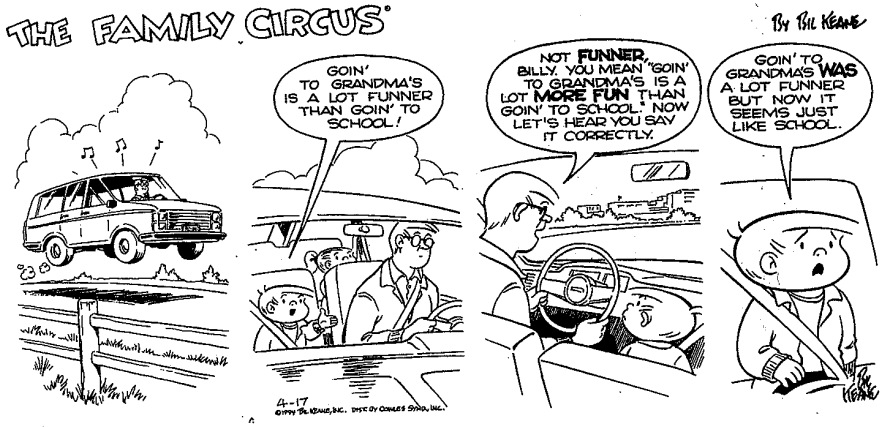
Children almost certainly do not solve the language acquisition problem by depending on negative feedback from parents. For one thing, parents could not very well correct or disapprove of their young children every time they err. Most of toddlers' sentences are ungrammatical in some way, so parents would be chiding them all day long. Parents focus on the content of their children's sentences, not their form, and let most errors slip by:

FATHER: Where is that big piece of paper I gave you yesterday?
ABE: Remember? I writed on it.
FATHER: Oh that's right don't you have any paper down here buddy?[21]

What happens when parents do correct their children? The cartoonist Bill Keane shows two of the results:

"Mommy, Dolly hitted me."
"Dolly HIT me."
"You too?! Boy, she's in trouble!"

THE FAMILY CIRCUS Bil Keane, Inc. Dist. By Cowles Synd., Inc.



THE FAMILY CIRCUS Bil Keane, Inc. Dist. By Cowles Synd., Inc.

Keane has a fine ear for children's language, and the dialogues are not fanciful. Here is a real one, transcribed by the psychologist Courtney Cazden:[22]

CHILD: My teacher holded the baby rabbits and we patted them.
ADULT: Did you say your teacher held the baby rabbits?
CHILD: Yes.
ADULT: What did you say she did?
CHILD: She holded the baby rabbits and we patted them.
ADULT: Did you say she held them tightly?
CHILD: No, she holded them loosely.

Systematic studies bear out the anecdotes. The linguist Arnold Zwicky, observing his daughter's overgeneralization of participle endings, reported that "six subsequent months of frequent corrections by her parents had no noticeable effect."[23] The psychologists James Morgan and Lisa Travis looked at transcribed speech of three children and their parents, sampled biweekly for several years. They wanted to see whether the children's errors elicited any consistent pattern from their parents—not only overt corrections, but partial or full repetitions, requests for clarification, questions, attempts to move the conversation on, or silence. No consistent pattern was found. In a follow-up study, Morgan and Travis found a different kind of grammatical error, in which parents do sometimes recast a child's sentence in correct English. But they found that the recasting had no effect—if anything, it had an adverse effect—on the child's subsequent improvement.[24]

The psychologist Karin Stromswold has a particularly dramatic demonstration that parental feedback cannot be crucial. She studied a child who, for unknown neurological reasons, was unable to talk, but who was an avid listener and understood complex sentences. When the boy was four, Stromswold tested his knowledge of past-tense forms by asking the boy to teach a dog puppet to talk. She asked him to give the dog a bone when it spoke correctly and a rock when it made an error. The boy awarded bones for *heated*, *baked*, *showed*, and *sewed*, and rocks for *eated*, *taked*, and *knowed*. He made just one error, awarding a bone for *goed*, similar to the performance of normal children. Somehow the boy, and presumably other children, can come to recognize that overgeneralized forms are ungrammatical without first having to make the errors and note their parents' response.[25]

Children must solve the logical problem of language acquisition in a different way. Perhaps, rather than learning the blocking principle from evidence that *singed* is not English, they *begin* with the blocking principle and use it to *deduce* that *singed* is not English. That is, blocking might be built in to the circuitry that drives language acquisition—what Chomsky calls Universal Grammar and what I have called the language instinct. As with all sane proposals about innate structure, such an instinct would not be an alternative to learning but rather an explanation of how learning works. In this case, because children hear parents say *sang* in the course of ordinary conversation, they retain *sang* in memory, and the blocking mechanism represses their tendency to say *singed*, turning them into adults.[26]

We need one more assumption to get the theory to work. If children already have blocking, and all else is the same, they should never say *singed* to begin with! Having heard their parents say *sang* even once should be enough to block the rule from applying to it. Fortunately, the extra assumption is as parsimonious as a theory in child psychology can be.

What is the simplest conceivable hypothesis of how children differ from adults? Answer: They have not lived as long. That is what being a child *means*. Now, among the experiences we accumulate as we live our lives is hearing the past-tense forms of irregular verbs. Human memory profits from repetition. If children have heard *sang* less often than adults have, their memory trace for it will be weaker and their ability to retrieve it will be less reliable. Sometimes, when they are trying to express the thought "singing in the past," *sang* will not pop into mind (or at least not quickly enough to get put into the sentence). Before children acquire the *-ed* rule, when they fail to retrieve *sang* they have no choice but to use the bare stem *sing*, even for events that happened in the past. But once they have acquired the rule, they can apply it to *sing*, creating

singed, thereby satisfying the syntactic constraint that tense be marked in every sentence.

This minimalist theory combines a simple idea from linguistics (blocking) with a simple idea from psychology (memory improves with repetition). It explains why children get worse before they get better, and solves the logical problem of how they exorcise their errors without parental feedback. Correct forms such as *sang* that a child used early on do not go anywhere once the child has acquired the rule, nor are they incapable of blocking errors: They simply must be retrieved from memory to do the blocking, and they are not always retrieved. The cure for overgeneralization is living longer, hearing irregulars more often, and consolidating them in memory, improving their retrievability.

Indeed this account, which posits that the mind of a child and the mind of an adult work the same way, is deducible from the very logic of irregularity, augmented only by the fact that memory is fallible. What is the past-tense form of the verb *to shend*, meaning "to shame"? If you answered *shended* then you have overgeneralized; the correct form is *shent*. This "error," of course, is to be expected. Irregular forms, by definition, are not predictable, so the only way you could have produced *shent* is if you had previously heard and remembered it. But you have heard it zero times and can't have remembered it. If in two years you were asked the question and erred once more, it still would not be surprising, because you would have heard it only once. Now put yourself in the child's shoes. Many verbs will be like *shent* for you: never heard, or not heard enough times to be recallable on demand. The mystery of why children say *singed* and *bleeded* has been solved.

~

When children say *singed*, are they simply little adults with bad memories? Gary Marcus and I combed through computer files with transcripts of the spontaneous speech of 83 children and extracted 11,500 sentences with irregular past-tense forms. We wanted to figure out when and why children start making errors, how often they do it, and with which verbs. Most of what we found fit the simple theory.[27]

First we looked at the error rate. If a child's language system is basically like an adult's, it should be designed to suppress the regularization of verbs that the child remembers are irregular. The suppression cannot be perfect because memory is not perfect, but children's memory for words ought to be fairly good; the child is, after all, using thousands of words and acquiring a new one, on average, every two hours. Overgeneralization errors should be the exception, not the rule, coming from the occasional breakdown of a system built to prevent the error. In fact the average error rate across children is only 4 percent. More than 95 percent of the time a child utters the past-tense form of an irregular verb it is a correct form like *sang*, not an error like *singed*. (Adults tend to overestimate the error rate because they remember the errors, which stick out like sore thumbs, and fail to notice the boring correct forms.) Once children begin to make the errors in their third year, they continue at this low rate until well into the school-age years.

No verb is immune to the errors, not even those a child used correctly before the error-making began. Nor is any verb consistently erred on. A child might use *felt* when young, then both *felt* and *feeled* when somewhat older. The errors are haphazard; children sometimes use correct and incorrect versions in quick succession, like this: "Daddy comed and said 'hey, what are you doing laying down?' And then a doctor came. . . ."[28] The hit-or-miss nature of the errors suggests that children are not ignorant of the correct forms; they are fallible at retrieving them. Some verbs are more error-prone than others, and the simple theory predicts that these should be the verbs that the child has heard less often. So we counted how often the children's parents used each irregular verb in the past tense. If a parent used *told* and *brought* more often than, say, *froze* and *won*, the child should have a stronger memory trace for *told* and *brought* than for *froze* and *won*, and should say *telled* and *bringed* less often than *freezed* and *winned*. We examined ninety irregular verbs and found that with every child, the more often the child's parents used a verb in the past tense, the less often the child regularized it.

Could children on some level really know that their errors are errors? Sometimes they do. The psycholinguists Dan Slobin and Tom Bever tried using their children's errors in their own speech, just for fun.[29] The children were not amused:

TOM: Where's Mommy?
CHILD: Mommy goed to the store.
TOM: Mommy goed to the store?
CHILD: NO! (*annoyed*) Daddy, *I* say it that way, not you.

CHILD: You readed some of it too . . . she readed all the rest.
DAN: She read the whole thing to you, huh?
CHILD: Nu-uh, you read some.
DAN: Oh, that's right, yeah. I readed the beginning of it.

CHILD: Readed? (*annoyed surprise*) Read! (*pronounced* rĕd)
DAN: Oh, yeah, read.
CHILD: Will you stop that, Papa?

In more controlled studies children are asked to judge the past-tense forms of a language-impaired puppet. They let many errors slip by, but they object to errors more often than to correct forms. And when asked to choose, the children, on average, prefer the correct forms.[30] All this suggests that children really do know irregular past-tense forms like *went* and *read*; their errors must be slip-ups in which they cannot slot an irregular form into a sentence in real time.

If overgeneralizing children are not qualitatively different from adults, we should see adults making the errors, and indeed they do, approximately once in every 25,000 sentences in which they use an irregular past-tense form.[31] This figure is about a thousand times less frequent than children's errors, but the estimate includes common verbs like *came* and *went* and *told* that have been drilled into our heads tens or hundreds of thousands of times. With the less common irregulars adults make "errors" quite often. It's hard to say *how* often, because we adults get to say what counts as "correct," and if we regularize an irregular often enough, we simply declare by fiat that it is not an error! These muzzy alternatives—*dreamed* and *dreamt*, *pleaded* and *pled*, *leaped* and *leapt*, *strided* and *strode*—are lower in frequency than pure irregular verbs like *went* and *came*, much as children's errors such as *winned* tend to occur with the verbs they hear less often. Even among pure irregular verbs, those used with lower frequency like *slew* and *strove* are judged to be somewhat unnatural, and their regular counterparts are judged to be relatively unobjectionable.[32]

Over the long run this psychology changes the composition of a language. Say you have heard *strode* only a few times in your life—more often than *shent*, but far less often than *held*. You would have a weak memory trace for *strode*, just strong enough for you to recognize it and for a little voice in your mind's ear to whisper, "strode!", but not strong enough to block the regular rule from applying. You may very well say *strided*, just as a child would say *hided*. If your neighbors are similarly ambivalent, the language community may be divided, with some people saying *strided*, others saying *strode*, and still others, hearing their neighbors using both forms without rhyme or reason, memorizing both and using them interchangeably.

With rarer verbs adults' "errors" create a vicious circle: They use an irregular form less and less, so their children and neighbors hear it slightly less often,

causing their memory traces in turn to be weaker, causing them to use it less (and regularize it more), in turn causing *their* children and neighbors to hear it less, and so on. An irregular form that falls below a critical frequency could disappear outright after a few generations. As we saw in chapter 3, that is exactly what has occurred in the history of English: The irregular forms of less common verbs such as *chide–chid*, *cleave–clove*, and *geld–gelt* became extinct.[33] Verbs, like all bits of culture, can rise or fall in popularity, and one can imagine a time when the verb *to geld* had slipped so far that a majority of adults lived their lives without having heard *gelt*. When pressed, they would have used *gelded*; the verb had become regular for them, and for all subsequent generations. That is why irregular verbs tend to be high in frequency; the list has been filtered repeatedly through the minds of children and adults, both of whom regularize uncommon irregular verbs.

~

What launches the transformation from regurgitating correct forms to creating incorrect ones? Why does a child wake up one morning and start to say *bleeded* and *singed*?

The simplest theory is that that is precisely the point at which the child has acquired the past tense rule, a result of the process described on page 193. The rule must be acquired at some point; it could not possibly be innate, because some languages don't mark tense on their verbs, and those that do don't use the English *-ed*. Prior to learning the rule, a child with an irregular form stuck on the tip of her tongue could do no better than to utter the bare stem *sing*; with the rule in hand she can fill the vacuum with *singed*.

One way to confirm the theory is to watch what happens to *regular* verbs when the child makes the first error with an irregular. Before the first error children leave regular verbs unmarked most of the time; they say *Yesterday we walk*. Then they begin to mark these verbs most of the time, as in *Yesterday we walked*. It is during this transition that the first error with an irregular form, like *singed* or *heared*, appears. We can interpret the tandem development of *walked* and *singed* as two signs of a single underlying process, the acquisition of the *-ed* rule: correct performance where the rule is called for, and errors where it is not.[34]

This idea that children add a rule onto their list of words is simpler than the suggestion that children radically reorganize their language, abandoning the list in favor of an imperialistic rule system and then slowly reacquiring the list. The simpler idea also fits the facts better. When children begin to make

these errors, where do the four percentage points of errors come from? Are they produced in situations where previously the child would have produced a correct form such as *sang*? Or are they produced in situations where the child would have produced a bare stem like *sing*? That is, do the errors drive out correct forms, a mysterious step backward on the road to adult language, or is one kind of error driving out a different kind? The data say that the errors of commission (*singed*) are driving out other errors, errors of omission (*sing*), not correct forms (*sang*). For example, before making his first overt error, one boy we studied used correct forms 74 percent of the time and produced the bare verb 26 percent of the time. When the errors began, at a rate of 2 percent, did they come out of the 74 percent of his verb usages that were correct, driving performance down to 72 percent? No—the correct forms *increased*, to 89 percent; the two percentage points of new errors came at the expense of the errors of omission, which dropped to 9 percent. Children don't backslide; when they replace *sang* with *singed*, they take a step forward, because the syntax of the sentence, which demands a past-tense form, is satisfied more often.[35]

What triggers the "Eureka!" moment, when a child first discovers a rule? Why does it dawn on some children in their late ones, but on others not until their late twos? I suspect we will never understand what triggers the very first error. Two children we studied made no errors for seven or eight months, popped out a single error (*feeled* or *heared*) just before turning three, and then went another five months before doing it again. Why the false start? What were the children thinking in the months when they failed to act on their epiphany? One possibility is that the gap is an illusion of sampling. Perhaps a newborn rule is wobbly and unreliable, and there are only so many times a child has the urge to use an irregular verb in the past tense, fails to retrieve its stored form, runs the rule to completion, and the tape recorder is running. A steady low probability in the mind of the child may surface as sputtering fits and starts in a record of the child's speech.

Another possibility is that language development at times really is chaotic, because the child is trying to make sense of the language with a changing brain. Synapses, the connections between brain cells, sprout and die in large numbers in the first few years of life, and the churning may temporarily swamp or wash away the newly laid down trace of a rule. Also, countless random events affect the microscopic structure of the growing brain. The human genome does not have nearly enough information to specify the wiring of the brain down to the last connection. We see this in identical twins, who share all

their genes and most of their experiences, and who have similar, but not identical, brains, intellects, and personalities.[36]

Jennifer Ganger and I suspected that at least some of the timing of language development, including the past tense rule, is controlled by a maturational clock. Children may begin to acquire a rule at a certain age for the same reason they grow hair or teeth or breasts at certain ages. If the clock is partly under the control of the genes, then identical twins should develop language in tighter synchrony than fraternal twins, who share only half their genes. We have enlisted the help of hundreds of mothers of twins who send us daily lists of their children's new words and word combinations. The checklists show that vocabulary growth, the first word combinations, and the rate of making past-tense errors are all in tighter lockstep in identical twins than in fraternal twins. The results tell us that at least some of the mental events that make a child say *singed* are heritable. The very first past-tense error, though, is not. When one twin makes an error like *singed* for the first time, an identical twin is no quicker to follow suit than a fraternal twin. These gaps—an average of thirty-four days between the first past-tense errors of two children with the same genes exposed to the same speech—are a reminder of the importance of sheer chance in children's development.[37]

I have explained children's creative errors by crediting them with a rule, but there is an alternative: Children might *analogize* from words they already know. They might say *holded* because *hold* sounds like *fold*, *mold*, and *scold*, whose past-tense forms are *folded*, *molded*, and *scolded*. Even with verbs like *sing* and *ring*, which do not rhyme exactly with any common regular verb, children could be reminded of bits and pieces of similar verbs like *sipped*, *banged*, *rimmed*, and *rigged*, and cobble together analogous *singed* and *ringed*.

That of course is the basis of the pattern associator memories developed by Rumelhart, McClelland, and their connectionist followers. Rumelhart and McClelland's model acquired hundreds of regular and irregular verbs, generalized to dozens of new verbs, and strikingly, appeared to go through a U-shaped sequence, first producing correct past-tense forms for irregular verbs and later overgeneralizing *-ed* to them. Yet the model had nothing that looks like a word, a rule, or a distinction between regular and irregular systems. How did they get a memory model to learn in a way that everyone has always taken to be a hallmark of rules?[38]

They had an ingenious idea. Rumelhart and McClelland figured that children acquire common verbs first, rarer verbs later. Since common verbs tend to be irregular, and rare verbs regular, the mixture of irregular and regular verbs in children's vocabularies should shift toward the regulars as their vocabulary grows and they begin to run out of irregulars and encounter more and more regulars. Moreover, children's vocabulary growth shows a big spurt several months after they learn their first words. That spurt could cause a sudden influx of regular verbs.

Pattern associator memories are highly sensitive to changes in the statistics of their input. When given a small number of oddball items, they memorize their patterns individually; when given a torrent of items sharing a pattern, they go with the numbers, extract the pattern, and run roughshod over the individual items, gradually reacquiring them over many subsequent bouts of training. That sounds a lot like children.

Rumelhart and McClelland imagined an extreme case: A child first learns a few common verbs, mostly irregular, followed by a spurt of hundreds of verbs, mostly regular. They consulted a list of word frequencies in English, found the dividing line that gave the strongest contrast in the mixture of regulars and irregulars, and trained their network in two stages accordingly. First the model was fed the 10 most frequent verbs in the list, 80 percent of which were irregular, 10 times apiece; then it was fed the next 410 verbs, 80 percent of which were *regular*, 190 times apiece. The network learned the first ten easily. Then, when bombarded by regular verbs, it strengthened thousands of connections to *-ed*, which overwhelmed the connections to the irregular pasts and led the model to make errors such as *breaked*. Connectionist modelers following in their footsteps used more sophisticated networks, but they also induced child-like behavior by changing the models' diet of regular and irregular verbs over time.[39]

Did the computer really "mimic the brain," as the headlines put it? It all depends on whether children begin to say *breaked* in response to an influx of regular verbs. Michelle Hollander and I checked the transcripts of the speech of three children over several years to see whether parents at some point start using more regular verbs when talking to their children. They did not. The proportion of regular verbs in parents' speech—about 25 to 30 percent—is the same When their children are two as when they are five. At first that may seem odd: When children are young, parents should favor the common, irregular verbs such as *make* and *do*; only when their children are older should they dip into the lower frequencies and use regular verbs such as *abate*, *abbreviate*,

and *abhor*. The reason that scenario doesn't appear in the statistics is that common irregulars like *make*, *do*, and *hold* are indispensable, general-purpose verbs that people of all ages depend on in every conversation. *Abate*, *abbreviate*, and so on compete *with one another* for air time, so even when the number of different regular verbs rotating in and out of conversation increases, the proportion of conversation filled with regular verbs remains constant.

Perhaps then we should be looking not at the number of *times* the verbs are used but rather at the number of *verbs* in the child's vocabulary, each counted once. There the proportion of regular verbs *must* increase, because there are only so many irregular verbs in the language, and when they begin to run out, the child has to acquire more and more regulars. That's how Rumelhart and McClelland derived their prediction. But counting vocabulary items is a bit odd if you think about the actual events that make up language acquisition. Children presumably learn as they listen to the speech coming out of the mouths of their parents, not by scanning their own mental dictionaries and feeding each verb into their network once per pass. We wanted to be charitable, though, so we checked the transcripts to see whether there really is a vocabulary spurt, and thus a richer mixture of regular verbs, when children begin to overapply *-ed*.

There was not. Children's vocabularies spurt in the mid-to-late ones, about a year too soon to trigger their past-tense errors, which begin in the mid-to-late twos. In the years in which children make the errors, regulars are coming in at a *lower* rate than they were earlier, when the children were using the irregulars correctly. The timing is not what we would need to get a pattern associator to overgeneralize after an early stage of correct performance.

The general problem is that Rumelhart and McClelland balanced their model on a knife-edge of assumptions about the statistics of the speech input to the child. But language acquisition is a robust process that does not live or die by the nuances of parental speech statistics. Throughout the world's cultures, children must learn the combinatorial tools of their language across a wide range of input mixtures, as we will see in the next chapter. Closer to home, even the English plural shows statistics unlike those of the past tense. The handful of irregular nouns known to children (*men, children, feet, teeth*) never could dominate their noun vocabularies the way that irregular verbs, at least in theory, could dominate their verb vocabularies. Yet children show the same U-shaped development with plurals as with past-tense forms. When they begin to speak, all of their plurals are correct, and then they begin to overgeneralize at a low rate for several years.[40]

Michael Ullman and I gave the pattern associator model two more chances to prove that it mimics the human brain. If children, like the model, learn by analogy, their irregular verbs should be lured into error by similar-sounding regular verbs and protected from error by similar-sounding irregular verbs. If *holded* is an analogy from soundalike *folded*, then the more soundalike regulars a verb has, and the more frequent they are, the more likely the verb will be regularized. *Holded* might be more common than *singed*, for example, because *holded* is strongly attracted to frequent *folded* and to a lesser extent *scolded* and *molded*, whereas *singed* is weakly attracted to low-frequency *blinked* and not much else. But when we correlated the number of potential seducers of a verb with its error rate in children's speech, we found little to no effect.[41]

The model did mimic the brain in one way. If *drank* owes its survival to similar *irregular* forms in memory such as *sank* and *rang*, then verbs with more irregular allies, and more common ones, should be erred on less often. And indeed they are. This difference—irregular forms need similar irregulars, but regular forms do not need similar regulars—parallels the findings from adults discussed in earlier chapters. It reinforces the compromise conclusion that pattern associators capture something about irregular forms and the memory in which they are stored, but fail to capture the nature of regular forms and the system in which they are computed.

~

What have children actually acquired when we say they have acquired a past tense rule? Is it just one more noise they can make, or is it the powerful combinatorial tool that, in conjunction with the rest of grammar, gives rise to the vast expressive range of a language and the elegant logic behind its apparent quirks?

Children's past-tense and plural rules really do seem like wobbly versions of the adult's, with their sweeping power to inflect any verb or noun. Children apply their past tense rule to almost all their irregular verbs, despite the strong associations to irregular past-tense forms. They apply it to unusual-sounding verbs of their own creation, such as *lightninged*, *smunched*, and *poonked*. They apply the rules to words built out of phrases, such as *eat lunched*, *cut-upped egg*, and *There is two Empire Strikes Backs*. Bilingual children sometimes apply a rule to words from their other language, as in *perachs* and *sefers*, Hebrew for flower and book.[42]

Children also apply the rule to rootless and headless words, the ones that lead to such curiosities as *lowlifes* and *flied out*, explained in the preceding

chapter. Kim, Marcus, and I gave children a *wug*-test with a twist. Half the new verbs were identical in sound to irregulars but obviously were based on nouns, like *to fly* meaning "to cover a piece of paper with flies" and *to ring* meaning "to put a ring on something." These are precisely the circumstances that for adults turn irregular-sounding verbs into regulars—*flied out to right field*; *high-sticked the goalie*; *ringed the city with artillery*—because a verb based on a noun lacks a root or head and cannot tap into the system of irregular roots stored in memory. Children as young as four work the same way. They regularize verbs based on nouns (as in *She flied the paper*) more often than they regularize verbs with verb roots (as in *They are flying down the road*).[43]

In a similar experiment children saw objects labeled with irregular nouns. Some were simple noun roots, such as a *fuzzy mouse* and a *little goose*; some were based on names, such as a *Mickey Mouse* and a *Mother Goose*; and some were bahuvrihi compounds, such as a *snaggletooth* (a walruslike creature) and a *bigfoot*. When asked to describe collections of these toys, the children used regular plurals for names and headless compounds (*Mickey Mouses, Mother Gooses, snaggletooths, bigfoots*) more often than for the simple noun roots (*fuzzy mouses, little gooses*). Children, like adults, don't just listen to a word's sound when they compute its inflection; they also analyze its grammatical structure.

Children are also sensitive to the other curiosity of irregular nouns discussed in the preceding chapter: the contrast between *mice-infested*, where an irregular behaves like any other word and can be inserted into a compound, and *rat-infested*, where the regular plural *rats* is computed too late to be inserted into a compound. I know of one child who insisted to his father that a building with mice was a *mice building*, and another child who said, "These aren't only handcuffs; they can be feetcuffs." I have never heard of a child say anything like *rats building* or *handscuffs*, though of course we need an experiment to show the difference is real.

Peter Gordon introduced a set of three- to five-year-olds to Cookie Monster and asked them, "Here is a monster who likes to eat X. What would you call him?" while varying the X. First he trained them on mass nouns like *mud*, which don't take a plural, until the children would say *mud-eater*. That introduced them to the compound construction without biasing their subsequent answers. Then he asked the children what they would call a monster who likes to eat rats. The children virtually always said *rat-eater*, not *rats-eater*, even though they had just heard the experimenter say *rats*. In contrast, they often called a monster who likes to eat mice a *mice-eater*—and those children who

occasionally said *mouses* never used it in compounds, as in *mouses-eater*. The avoidance of regular plurals was not simply an aversion to the sound *-s* inside a compound. As with adults, when the children were asked about pluralia tantum nouns such as *pants* and *clothes*, which sound regular but have to be stored like irregulars, they were happy to call the monster a *pants-eater* or a *clothes-eater*.[44]

Gordon then tested whether children could have learned the distinction by noticing irregular plural-containing compounds such as *teethmarks* in their parents' speech, while noticing the absence of regular-plural-containing compounds such as *clawsmarks*. He examined all the compounds in standard frequency counts and discovered that *neither* kind of plural-containing compound is common; virtually all commonly used compounds take a singular first noun, such as *toothbrush* and *mousetrap*. Most of the children walked into the lab never having heard a compound containing a plural, but the first time they faced the temptation they used irregular plurals and avoided using regular plurals. Children's sensitivity to the distinction between *mice-infested* and *rats-infested*, Gordon concluded, is a product of the innately specified architecture of their language system, not a product of tabulating forms in parental speech.

Kim and I asked the same question of children's ability to distinguish *flied* meaning "covered with flies" from *flew* meaning "soared." Children hear plenty of verbs-from-nouns, such as *to fish, to plug, to rain*, and *to screw in*. We discovered, however, that they do not hear any verbs-from-nouns that sound like an irregular verb, such as *flied out* or *high-sticked*.[45] That means that prior experience could not have told them what to do when a verb's sound calls for one past-tense form and its structure calls for another; they tend toward the correct answer on their own.

Of course, the speech heard by young children must contain information that tells them that an inflection is regular to begin with. What is that information? It cannot be simply the presence of added material on some words, because that would not distinguish the regular *-ed* in *pat–patted* from the irregular *-en* in *shake–shaken*. Nor can it be the sheer number of words bearing added material, because we saw that children's use of a rule has nothing to do with the proportion of regular verbs in their parents' speech or in their own vocabularies.

How can children recognize a regular inflection when they hear one? Suppose that children's language systems are prepared for words and for rules, and are always on the lookout for examples of each in parents' speech. Children listen for stretches of sound that fit the canonical pattern for a word in their language and that are arbitrarily paired with a meaning. Those are grabbed by the word system and stored as roots. Children also listen for words that might be modified versions of some other word after a rule has had its way with it. Those are snipped into stems and suffixes and the suffix is stored separately as rule-ready material. To distinguish words from rules, then, children must have antennae for signs that a pattern of vowels and consonants has been *added* to a word rather than being *part* of that word.

What might those signs be? They could be the kinds of information that linguists themselves use to determine whether an inflection is regular, the kind of evidence we explored in the preceding two chapters. If children hear a suffixed version of a verb that falls into a family of similar irregular verbs, such as *blinked* and *showed* (which sound like they should belong to the *drink-shrink-sink* family and the *blow-grow-throw* family), they can infer that the words have been modified by something strong enough to have nullified the pull toward the family. If they notice suffixed verbs based on nouns, such as *combed* and *fished*, or on onomatopoeia, such as *cracked* and *squeaked*, they could hear the noun or the environmental sound inside the word, and assume that the residue must have been added by a rule that is free to apply to words that aren't verb roots. If children hear suffixed verbs with nonbasic sounds such as *attached* and *exercised* (which are polysyllabic), they can guess that these forms are unlikely to be roots linked in memory to other roots; the extra bit is likely to have been added by a rule that doesn't care about sound. We don't know whether children rely on these telltale signs, but we do know that the signs are available if children knew what to listen for. All four kinds of verbs may be found in the vocabularies of young children before the stage at which they clearly apply rules in errors such as *singed*.[46] So children *could* use these signs of wordhood and rulehood if they had the mental apparatus of words and rules to interpret what they hear. Once a suffix has been identified as a rule product using the audible cues, it would be available for productive combination with new verbs.

~

Children's speech errors, which make such engaging anecdotes in poetry, novels, television features, and web sites for parents, may help us untangle one of the thickest knots in science, nature and nurture. When a child says *It bleeded and it singed*, the fingerprints of learning are all over the sentence. Every bit of every word has been learned, including the past tense suffix *-ed*. The very existence of the error comes from a process of learning that is still incomplete: mastery of the irregular forms *bled* and *sang*.

But learning is impossible without innately organized circuitry to do the learning, and these errors give us hints of how it works. Children are born to attend to minor differences in the pronunciation of words, such as *walk* and *walked*. They seek a systematic basis for the difference in the meaning or form of the sentence, rather than dismissing it as haphazard variation in speech styles. They dichotomize time into past and nonpast, and correlate half the time line with the evanescent word ending. They must have a built-in tendency to block the rule when a competing form is found in memory, because there is no way they could learn the blocking principle in the absence of usable feedback from their parents. Their use of the rule (though perhaps not the moment when they first use it) is partly guided by their genes. They spontaneously deploy their new rule to a wide range of words coined by an experimenter or by themselves, and to verbs whose irregular forms are too faint to retrieve. Children fit the rule into its proper place in the logic of their grammatical system, keeping regular forms out of certain word structures and irregular forms out of others.

I suspect that in other parts of our psychology the interaction of nature and nurture has a similar flavor: Every bit of content is learned, but the system doing the learning works by a logic innately specified. Charles Darwin captured the interaction when he called human language "an instinctive tendency to acquire an art." "It certainly is not a true instinct," he noted, "for every language has to be learned. It differs, however, widely from all ordinary arts, for man has an instinctive tendency to speak, as we see in the babble of our young children; while no child has an instinctive tendency to brew, bake, or write."[47]

8

THE HORRORS OF THE GERMAN LANGUAGE

Though it is sometimes easy for Americans to forget, English is not the only language spoken in the world. Humans babble in some six thousand languages falling into thirty-odd families. For many reasons, those mother tongues are a motherlode for the understanding of language and mind.

First, no one is biologically disposed to speak a particular language. The experiments called immigration and conquest, in which children master languages unknown to their ancestors, settled that question long ago. This means that if some feature of language is the handiwork of a fundamental mechanism of the human language faculty, it ought to be visible anywhere from Lapland to Lesotho, from Peru to Papua New Guinea.

Also, to understand language we have to test hypotheses about cause and effect, but linguists don't have the luxury of synthesizing a language in a test tube and seeing how it is spoken, learned, and changed. The differences among languages already out there make up the only laboratory apparatus that allows a linguist to vary one factor and see how it affects another.

Finally, no one supposes that language evolved six thousand times. We find different languages because people move apart and lose touch, or split into factions that hate each other's guts. People always tinker with the way they talk, and as the tinkerings accumulate on different sides of the river, mountain range, or no-man's-land, the original language slowly splits in two. To compare