**Stanford Presidential Lectures
in the Humanities and Arts**

Calendar

Lecturers

**Douglas
Hofstadter**

Excerpts

Bibliography

Hofstadteriana

Schedule

Archives

**Stanford
Humanities
Center**

Analogy as the Core of Cognition

by Douglas R. Hofstadter

Once upon a time, I was invited to speak at an analogy workshop in the legendary city of Sofia in the far-off land of Bulgaria. Having accepted but wavering as to what to say, I finally chose to eschew technicalities and instead to convey a personal perspective on the importance and centrality of analogy-making in cognition. One way I could suggest this perspective is to rechant a refrain that I've chanted quite oft in the past, to wit:

One should not think of analogy-making as a special variety of *reasoning* (as in the dull and uninspiring phrase "analogical reasoning and problem-solving," a long-standing cliché in the cognitive-science world), for that is to do analogy a terrible disservice. After all, reasoning and problem-solving have (at least I dearly hope!) been at long last recognized as lying far indeed from the core of human thought. If analogy were merely a special variety of something that in itself lies way out on the peripheries, then it would be but an itty-bitty blip in the broad blue sky of cognition. To me, however, analogy is anything but a bitty blip — rather, it's the very blue that fills the whole sky of cognition — analogy is *everything,* or very nearly so, in my view.

End of oft-chanted refrain. If you don't like it, you won't like what follows.

The thrust of my chapter is to persuade readers of this unorthodox viewpoint, or failing that, at

least to give them a strong whiff of it. In that sense, then, my article shares with Richard Dawkins's eye-opening book *The Selfish Gene* (Dawkins 1976) the quality of trying to make a scientific contribution mostly by suggesting to readers a shift of viewpoint — a new take on familiar phenomena. For Dawkins, the shift was to turn causality on its head, so that the old quip "a chicken is an egg's way of making another egg" might be taken not as a joke but quite seriously. In my case, the shift is to suggest that every concept we have is essentially nothing but a tightly packaged bundle of analogies, and to suggest that all we do when we think is to move fluidly from concept to concept — in other words, to leap from one analogy-bundle to another — and to suggest, lastly, that such concept-to-concept leaps are themselves made via analogical connection, to boot.

This viewpoint may be overly ambitious, and may even — horrors! -- be somewhat wrong, but I have observed that many good ideas start out by claiming too much territory for themselves, and eventually, when they have received their fair share of attention and respect, the air clears and it emerges that, though still grand, they are not quite so grand and all-encompassing as their proponents first thought. But that's all right. As for me, I just hope that my view finds a few sympathetic readers. That would be a fine start.

Two Riddles

We begin with a couple of simple queries about familiar phenomena: "Why do babies not remember events that happen to them?" and "Why does each new year seem to pass faster than the one before?"

I wouldn't swear that I have the final answer to either one of these queries, but I do have a hunch, and I will here speculate on the basis of that hunch. And thus: the answer to both is basically the same, I would argue, and it has to do with the relentless, lifelong process of chunking — taking "small" concepts and putting them together into bigger and bigger ones, thus recursively building up a giant repertoire of concepts in the mind.

How, then, might chunking provide the clue to these riddles? Well, babies' concepts are simply too small. They have no way of framing entire events whatsoever in terms of their novice concepts. It is as if babies were looking at life through a randomly drifting keyhole, and at each

moment could make out only the most local aspects of scenes before them. It would be hopeless to try to figure out how a whole *room* is organized, for instance, given just a keyhole view, even a randomly drifting keyhole view.

Or, to trot out another analogy, life is like a chess game, and babies are like beginners looking at a complex scene on a board, not having the faintest idea how to organize it into higher-level structures. As has been well known for decades, experienced chess players chunk the setup of pieces on the board nearly instantaneously into small dynamic groupings defined by their strategic meanings, and thanks to this automatic, intuitive chunking, they can make good moves nearly instantaneously and also can remember complex chess situations for very long times. Much the same holds for bridge players, who effortlessly remember every bid and every play in a game, and months later can still recite entire games at the drop of a hat.

All of this is due to chunking, and I speculate that babies are to life as novice players are to the games they are learning — they simply lack the experience that allows understanding (or even perceiving) of large structures, and so nothing above a rather low level of abstraction gets perceived at all, let alone remembered in later years. As one grows older, however, one's chunks grow in size and in number, and consequently one automatically starts to perceive and to frame ever larger events and constellations of events; by the time one is nearing one's teen years, complex fragments from life's stream are routinely stored as high-level wholes — and chunks just keep on accreting and becoming more numerous as one lives. Events that a baby or young child could not have possibly perceived as such — events that stretch out over many minutes, hours, days, or even weeks — are effortlessly perceived and stored away as single structures with much internal detail (varying amounts of which can be pulled up and contemplated in retrospect, depending on context). Babies do not have large chunks and simply cannot put things together coherently. Claims by some people that they remember complex events from when they were but a few months old (some even claim to remember being born!) strike me as nothing more than highly deluded wishful thinking.

So much for question number one. As for number two, the answer, or so I would claim, is very similar. The more we live, the larger our repertoire of concepts becomes, which allows us to gobble up ever larger coherent stretches of life in single mental chunks. As we start seeing life's patterns on higher and higher levels, the lower levels nearly vanish from our perception.

This effectively means that seconds, once so salient to our baby selves, nearly vanish from sight, and then minutes go the way of seconds, and soon so do hours, and then days, and then weeks...

"Boy, this year sure went by fast!" is so tempting to say because each year is perceived in terms of chunks at a higher, grander, larger level than any year preceding it, and therefore each passing year contains fewer top-level chunks than any year preceding it, and so, psychologically, each year seems sparser than any of its predecessors. One might, somewhat facetiously, symbolize the ever-rapider passage of time by citing the famous harmonic series:

$$1 + 1/2 + 1/3 + 1/4 + 1/5 + 1/6 + 1/7 + 1/8 +...$$

by which I mean to suggest that one's nth year feels subjectively n times as short as one's first year, or $n/5$ times as short as one's fifth year, and so on. Thus when one is an adult, the years seem to go by about at roughly a constant rate, because — for instance — $(1/35)/(1/36)$ is very nearly 1. Nonetheless, according to this theory, year 70 would still shoot by twice as fast as year 35 did, and seven times as fast as year 10 did.

But the exact numerical values shown above are not what matter; I just put them in for entertainment value. The more central and more serious idea is simply that relentless mental chunking makes life seem to pass ever faster as one ages, and there is nothing one can do about it. So much for our two riddles.

Analogy, Abstract Categories, and High-level Perception

Before I go any further, I would like to relate all this to analogy, for to some the connection may seem tenuous, if not nonexistent. And yet to me, by contrast, analogy does not just lurk darkly here, but is right up there, front and center. I begin with the mundane observation that vision takes an input of millions of retinal dots and gives an output of concepts — often words or phrases, such as "duck," "Victorian house," "funky chair," "Joyce Carol Oates hairdo," or "looks sort of like President Eisenhower." The (visual) perceptual process, in other words, can be thought of as the triggering of mental categories — often standard lexical items — by scenes. Of course, high-level perception can take place through other sensory modalities: we

can hear a low rumbling noise and say "helicopter," can sniff something and remark "doctor's office," can taste something and find the words "okra curry" jumping to our tongue, and so on.

In fact, I should stress that the upper echelons of high-level perception totally transcend the normal flavor of the word "perception," for at the highest levels, input modality plays essentially no role. Let me explain. Suppose I read a newspaper article about the violent expulsion of one group of people by another group from some geographical region, and the phrase "ethnic cleansing," nowhere present in the article, pops into my head. What has happened here is a quintessential example of high-level perception — but what was the input medium? Someone might say it was vision, since I used my eyes to read the newspaper. But really, was I perceiving ethnic cleansing visually? Hardly. Indeed, I might have heard the newspaper article read aloud to me and had the same exact thought pop to mind. Would that mean that I had aurally perceived ethnic cleansing? Or else I might be blind and have read the article in Braille — in other words, with my fingertips, not my eyes or ears. Would that mean that I had tactilely perceived ethnic cleansing? The suggestion is absurd.

The sensory input modality of a complex story is totally irrelevant; all that matters is how it jointly activates a host of interrelated concepts, in such a way that further concepts (e.g., "ethnic cleansing") are automatically accessed and brought up to center stage. Thus "high-level perception" is a kind of misnomer when it reaches the most abstract levels, but I don't know what else to call it, because I see no sharp line separating it from cases of recognizing "French impressionism" in a piece of music heard on the radio or thinking "Art Deco" when looking at a typeface in an advertisement.

The triggering of prior mental categories by some kind of input — whether sensory or more abstract — is, I insist, an act of analogy-making. Why is this? Because whenever a set of incoming stimuli activates one or more mental categories, some amount of slippage must occur (no instance of a category ever being precisely identical to a prior instance). Categories are quintessentially fluid entities; they adapt to a set of incoming stimuli and try to align themselves with it. The process of inexact matching between prior categories and new things being perceived (whether those "things" are physical objects or bite-size events or grand sagas) is analogy-making par excellence. How could anyone deny this? After all, it is the mental mapping onto each other of two entities — one old and sound asleep in the recesses of long-

term memory, the other new and gaily dancing on the mind's center stage — that in fact differ from each other in a myriad of ways.

The Mental Lexicon: A Vast Storehouse of Triggerable Analogies

We humans begin life as rather austere analogy-makers — our set of categories is terribly sparse, and each category itself is hardly well-honed. Categories grow sharper and sharper and ever more flexible and subtle as we age, and of course fantastically more numerous. Many of our categories, though by no means all, are named by words or standard phrases shared with other people, and for the time being I will concentrate on those categories — categories that are named by so-called lexical items. The public labels of such categories — the lexical items themselves — come in many grades, ranging more or less as follows:

- *Simple words:* chair, clock, cork, cannon, crash, clown, clue, cloak, climber...
- *Compound words:* armchair, alarm clock, corkscrew, cannonball, skyscraper, station wagon, sexpot, salad dressing, school bus, jukebox, picket line, horror movie, wheeler-dealer...
- *Short phrases:* musical chairs, out of order, Christmas tree ornament, nonprofit organization, business hours, foregone conclusion, rush-hour traffic, country-Western music, welcome home, tell me about it, give me a break, and his lovely wife, second rate, swallow your pride...
- *Longer phrases:* stranded on a desert island; damned if you do, damned if you don't; praise the Lord and pass the ammunition; not in the foreseeable future; to the best of my knowledge; and they lived happily ever after; if it were up to me; haven't seen her since she was knee-high to a grasshopper; you could have knocked me over with a feather; thank you for not smoking; handed to him on a silver platter...

Such lists go on and on virtually forever, and yet the amazing fact is that few people have any inkling of the vastness of their mental lexicons (I owe a major debt here to Joe Becker — see Becker 1975). To be sure, most adults use their vast mental lexicons with great virtuosity, but they have stunningly little explicit awareness of what they are doing.

It was Roger Schank, I believe, who pointed out that we often use proverbs as what I would

call "situation labels," by which I mean that when we perceive a situation, what often springs to mind, totally unbidden, is some proverb tucked away in our unconscious, and if we are talking to someone, we will quote that proverb, and our listener will in all likelihood understand very clearly how the proverb "fits" the situation-in other words, will effortlessly make the mapping (the analogy, to stress what it is that we are talking about here) between the phrase's meaning and the situation. Thus the following kinds of phrases can easily be used as situation labels:

- That's the pot calling the kettle black if I ever saw it!
- It just went in one ear and out the other...
- Speak of the devil!
- When the cat's away the mice will play!

The Common Core behind a Lexical Item

I now make an observation that, though banal and obvious, needs to be made explicitly nonetheless — namely, things "out there" (objects, situations, whatever) that are labeled by the same lexical item have something, some core, in common; also, whatever it is that those things "out there" share is shared with the abstract mental structure that lurks behind the label used for them. Getting to the core of things is, after all, what categories are for. In fact, I would go somewhat further and claim that getting to the core of things is what thinking itself is for-thus once again placing high-level perception front and center in the definition of cognition.

The noun "shadow" offers a good example of the complexity and subtlety of structure that lurks behind not just some lexical items, but behind every single one. Note, first of all, the subtle difference between "shadow" and "shade": we do not speak of cattle seeking shadow on a hot day, but shade. Many languages do not make this distinction, and thus they offer their native speakers a set of categories that is tuned slightly differently.

In many parts of the world, there are arid zones that lie just to the east of mountain ranges (e.g., the desert in Oregon just to the east of the Cascade mountains); these regions are standardly referred to as the mountain chain's "rain
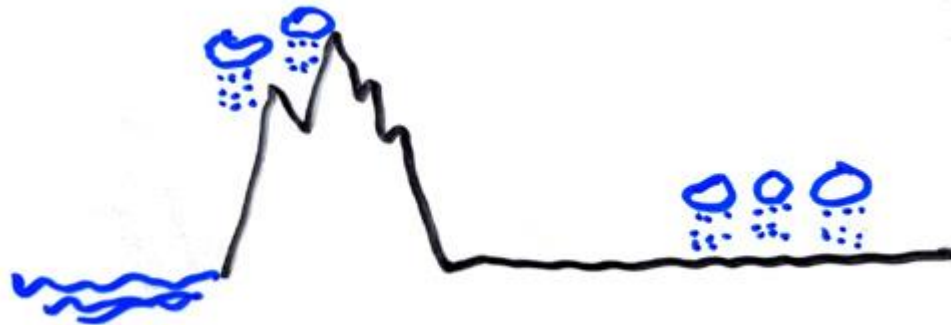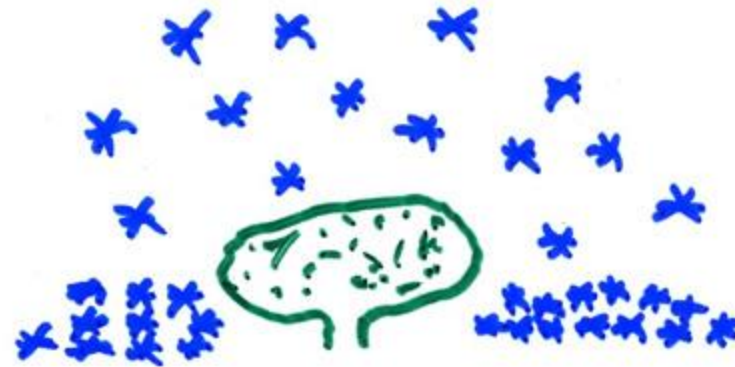

shadow."

What does one call the roughly circular patch of green seen underneath a tree after a snowfall? It could clearly be called a "snow shadow" — the region where snow failed to fall, having been blocked by an object.

A young woman who aspires to join her high-school swimming team, but whose mother was an Olympic swimmer, can be said to be "in the shadow of her mother." In fact, if she joins the team and competes, she might even be said to be "swimming in the shadow of her mother." And if she performs less well than her mother did, she will be said to be "overshadowed" by her mother.

One might say about a man who has had a bout with cancer but has recovered and is now feeling more secure about his health, "He is finally feeling more or less out of the shadow of his cancer." Along similar lines, many countries in Europe have recovered, to a large extent, from the ravages of World War II, but some might still be said to lie "in the shadow of World War II."

Another type of shadow cast by World War II (or by any war) lies in the skewed population distribution of any decimated group; that is, one imagines the human population as constituting a kind of flow of myriad tiny entities (individual people) down through the years (like that of photons or snowflakes through space), but long after the war's end, there are certain "regions" of humanity (e.g., certain ethnic groups) where the flow of births has been greatly reduced, much as if by an "obstacle" (namely, the millions of deaths in prior generations, whose effect continues to reverberate for many decades before gradually fading away, as a group's population replenishes itself).

There is of course no sharp line between cases where a word like "shadow" is used conventionally and cases where it is used in a novel manner; although "rain shadow" is something of a standard phrase, "snow shadow" (even though it is far easier to see) is less common. And notions like that of "population shadow" mentioned at the end are probably

novel to most readers of this article, even though a closely related notion like "in the shadow of the war" is probably not new.

In short, the domain of the word "shadow" is a blurry region in semantic space, as is any human category, and — here I hark back to my initial refrain — that blur is due to the subtleties of mapping situations onto other situations-due, in other words, to the human facility of making analogies. The point is, a concept is a package of analogies.

Complex Lexical Items as Names of Complex Categories

Over the next few pages I will present a potpourri of mental categories (via proxies — namely, their English-language lexical-item representations); I invite you to think, as you consider each item, just what it is that very different exemplars of the category in question tend to have in common. Thus:

- dog
- backlog
- probably
- probab-*lee!*

I interrupt the list momentarily to comment on the last two entries, which of course are not nouns. (Who says nouns are the only mental categories? Obviously, verbs represent categories as well-but the same holds true, no less, for adjectives, adverbs, and so forth.) Some situations call forth the word "probably"; most do not. To some situations, the concept behind the word "probably" simply fits, while to most, it does not fit. We learn how to use the word "probably" over the course of years in childhood, until it becomes so ingrained that it never crosses our mind that "probably" is the name that English speakers give to a certain category of situations; it simply is evoked effortlessly and rapidly by those situations, and it is uttered without any conscious thought as to how it applies. It just "seems right" or "sounds right."

What, then, about the word below it: "probab-*lee*"*?* This, too, is a lexical item in the minds of most native speakers of contemporary American English — perhaps not often used, perhaps more commonly heard than uttered by readers of this article, but nonetheless, we native

speakers of American English all relate to hearing the word "probably" accented on its final rather than its initial syllable, and we all somehow realize the connotations hidden therein, though they may be terribly hard to articulate. I won't try to articulate them myself, but I would merely point out that this phonetic variant of the word "probably" fits only certain situations and not others (where the "situation" includes, needless to say, not just what is being talked about but also the mood of the speaker, and the speaker's assessment of the mood of the listener as well). Example: "Are our stupid leaders ever going to learn their lesson?" "Who knows? Maybe they're doomed to keep on repeating the mistakes of the past." "Mmm... Probab-*lee*..."

My point, with all the phrases cited above, is to bring to your conscious awareness the fact that there are certain situations that one could call "probab-*lee*! situations" no less than there are certain situations that are "musical-chairs situations" or "speak-of-the-devil situations." In short, lexical items can be very abstract categories evoked by special classes of situations and not by others. This applies to adjective, adverbs, prepositions, interjections, short and long phrases, and so on. Thus let me continue my list.

- Come on!
- Go for it!
- It's about time!
- Well, excuuuuuuuuuuse me!
- Let's not stand on ceremony!
- without batting an eyelash
- ain't

Lest the lowest item above seem puzzling, let me point out that the notorious contraction "ain't," although it is in a certain sense ungrammatical and improper, is nonetheless used very precisely, like pinpoint bombing, by politicians, reporters, university presidents, and the like, who carefully and deliberately insert it into their speech at well-timed moments when they know their audience almost expects it-it fits the context perfectly. For example, a general trying to justify a bombing raid might say, in describing the series of deadly skirmishes that provoked it, "I'm sorry, but a Sunday picnic it just ain't." This is just one of many types of "ain't" situations. We native speakers know them when we hear them, and we likewise have a

keen ear for improper uses of the word "ain't" by educated people, even if we ain't capable of putting our finger on what makes them inappropriate. (Curiously enough, shortly after drafting this paragraph, I came across an article in the *New York Times* about the failure of a test missile to hit its target, and a perfectly straight photo caption started out, "Two out of four goals ain't bad..." As I said above, even the most highly placed sources will use this "ungrammatical" word without batting an eyelash.)

"Suggestions" Imparted on the Soccer Field

As a nonnative speaker of Italian watching the 1998 Soccer World Cup on Italian television, I was struck by the repeated occurrence of a certain term in the rapid-fire speech of all the commentators: the word *suggerimento*

It would be naive to imagine that each lexical item defines a perfectly "spherical" region in conceptual space, as pristine as an atomic nucleus surrounded by a spherical electron cloud whose density gradually attenuates with increasing distance from the core. Although the single-nucleus spherical-cloud image has some truth to it, a more accurate image of what lies behind a typical lexical item might be that of a molecule with two, three, or more nuclei that share an irregularly shaped electron cloud. .

Suggerimento provides a perfect example of such a molecule, with one of its constituent atoms being the notion of a verbal piece of advice, another the notion of prompting on a theater stage, yet a third being the notion of a certain type of downfield soccer pass, and so forth. There is something in common, of course, that these all share, but they are nonetheless distinguishable regions in conceptual space.

Often native speakers of a language have a hard time realizing that two notions labeled identically in their language are seen as highly distinct concepts by speakers of other languages. Thus, native speakers of English feel the verb "to know" as a monolithic concept, and are sometimes surprised to find out that in other languages, one verb is used for knowing facts, a different verb for knowing people, and there may even be a third verb for knowing how to do things. When they are first told this, they are able to see the distinction, although it may seem highly finicky and pointless; with practice, however, they build up more refined categories until a moment may come when what once seemed an unnatural and gratuitous division of mental space now seems to offer a useful contrast between rather distinct notions. And conversely, speakers of a language where all three of these notions are represented by distinct lexical items may find it revelatory, fascinating, and perhaps even elegant to see how they are all subsumed under one umbrella-word in English.

My main point in bringing this up is simply to make explicit the fact that words and concepts are far from being regularly shaped convex regions in mental space; polysemy (the possession of multiple meanings) and metaphor make the regions complex and idiosyncratic. The simplest concepts are like isolated islands in a sea; the next-simplest are like pairs of islands joined by a narrow isthmus; then there are trios with two or three isthmuses having various widths; and so on. Caveat: When I say "simplest concepts," I do not mean those concepts that we pick up earliest in life, but in fact quite the contrary. After all, the majority of concepts planted in

earliest childhood grow and grow over a lifetime and turn into the most frequently encountered concepts, whose elaborate ramifications and tendrils constitute the highest degree of twistiness! What I mean by "simplest concept" is merely "concept with maximally simple shape"; such a "simple" concept would most likely owe its simplicity precisely to its low frequency, and thus would seem like a sophisticated adult concept, such as "photosynthesis" or "hyperbola."

Conceptual Families and Lexical Rivalry

Walking down the corridors of a building in Italy in which I have worked over several summers, I have been faced innumerable times with an interesting problem in high-level perception that has to be solved in real time — in a couple of seconds at most, usually. That is, how do I greet each person who I recognize as we approach each other in the hall, and then pass? Here are five sample levels of greeting (there are dozens more, needless to say):

- *Buon giorno!* ("Hello!" or perhaps "Morning.")
- *Salve!* ("Howdy!" or perhaps "How are you.")
- *Buondì!* (Perhaps "Top o' the mornin'!" or "How ya doin'?")
- *Ciao!* ("Hi!" or "Hi there!")
- *Come stai?* ("How are you doing?" or perhaps "What's up?")

Each of them conveys a particular level of mutual acquaintance and a particular position along the formality/informality spectrum. And of course it frequently happens that I recognize someone but can't even remember how often I've met them before (let alone remember what their name is or what their role is), and so I have to make a decision that somehow will allow me to cover at least two different levels of friendliness (since I'm really not sure how friendly we are!). The choice is incredibly subtle and depends on dozens if not hundreds of variables, all unconsciously felt and all slightly contributing to a "vote" among my neurons, which then allow just one of these terms (or some other term) to come bubbling up out of my dormant Italian mental lexicon.

Consider the following spectrum of phrases all having in a certain sense "the same meaning," but ranging from very vulgar to somewhat incensed to quite restrained to utterly bland:

- He didn't give a flying f*** .
- He didn't give a good God damn.
- He didn't give a tinker's damn.
- He didn't give a damn.
- He didn't give a darn.
- He didn't give a hoot.
- He didn't care at all.
- He didn't mind.
- He was indifferent.

For many native speakers, there are situations that correspond to each of these levels of intensity. To be sure, some speakers might be loath to utter certain of these phrases, but true native-level mastery nonetheless entails a keen awareness of when each of them might be called for in, say, a movie, or simply coming out of the mouth of someone else. After all, a large part of native mastery of a language is deeply knowing how other people use the language, regardless of whether one oneself uses certain phrases. And thus, to reiterate our theme, there are "He-didn't-give-a-good-God-damn situations" and there are situations of a very different sort, which could be called "He-didn't-care-at-all situations," and so forth. Each of the above expressions, then, can be thought of as the name of a particular type of situation, but since these categories are much closer to each other than just randomly chosen categories, they constitute potential rivalries that may take place during the ultra-fast high-level perceptual act that underlies speech.

Lexical Blends as a Window onto the Mind

Lexical blends, which are astonishingly common though very seldom noticed by speakers or by listeners, reveal precisely this type of unconscious competition among close relatives in the mental lexicon. A lexical blend occurs when a situation evokes two or more lexical items at once and fragments of the various evoked competitors wind up getting magically, sometimes seamlessly, spliced together into the vocalized output stream (see, for example, Hofstadter and Moser 1989). Occasionally the speaker catches such an error on its way out and corrects it, though just as often it goes totally unheard by all parties. Thus people make blends of the following sorts:

- Word-level blends: mop/broom => brop
- Phrase-level blends: easygoing/happy-go-lucky => easy-go-lucky
- Sentence-level blends: We'll leave no stone unturned/We'll pull out all the stops => We'll pull no stops unturned.

Blends reveal how much goes on beneath the surface as our brains try to figure out how to label simpler and more complex situations. In a way, what is amazing is that blends are not more common. Somehow, through some kind of cerebral magic, speakers light most of the time upon just one lexical label despite the existence of many potential ones, rather than coming out with a mishmosh of several — much as when a good pianist plays the piano, it is very seldom that two keys are struck at once, even though it might seem, a priori, that striking two neighboring keys at once ought to happen very often.

A Lexical Item as One Side of a Perceptual Analogy

At the risk of boring some readers, I shall now continue with my rather arbitrary sampler of lexical items, just to drive the point home that every lexical item that we possess is a mental category, and hence, restating what I earlier claimed, every lexical item, when used in speech (whether received or transmitted), constitutes one side of an analogy being made in real time in the speaker's/listener's mind. I thus urge readers to try on for size the mindset that equates a lexical item with the "name" of a certain blurry set of situations centered on some core. Though this sounds quite orthodox for nouns, it is less so for verbs, and when applied to many of the following linguistic expressions, it is highly unorthodox:

- slippery slope
- safety net
- shades of...
- Been there, done that.
- Forget it!
- It was touch-and-go.
- take a turn for the worse
- Be my guest!
- Make my day!

- Fancy that!
- Put your money where your mouth is!
- I mean,...
- Well,...
- Don't tell me that...
- It's fine to [do X] and all, but...
- kind of [+ adj.]
- when it comes to the crunch...
- You can't have it both ways!
- ...that's for sure!
- the flip side [of the coin] is...
- You had to be there.
- It's high time that...
- Whatever!

Consider the teenager's favorite rejoinder, "Whatever!" If one were to try to capture its meaning — its range of applicability — one might paraphrase it somewhat along these lines: "You think such and so, and I disagree, but let's just agree to disagree and move on..." It takes a good number of years before one has acquired the various pieces of cognitive equipment that underpin the proper usage of such a phrase (which again ties in with the fact that one cannot remember events from one's babyhood).

High-Level Mental Chunks That Lack Labels

Although long stock phrases like "Put your money where your mouth is!" might seem to stretch the notion of mental chunking to the limit, that's hardly the case. Indeed, such phrases lie closer to the beginning than to the end of the story, for each one of use also remembers many thousands of events in our personal lives that are so large and so idiosyncratic that no one has ever given them a name and no one ever will, and yet they nonetheless are sharp memories and are revealed for the mental categories they are by the fact that they are summoned up cleanly and clearly by certain situations that take place later, often many years later. Thus take this one sample mental chunk, from my own personal usually dormant repertoire:

that time I spent an hour or two hoping that my old friend Robert, who I hadn't seen in two years but who was supposed to arrive from Germany by train sometime during that summer day in the little Danish fishing village of Frederikssund (which in a series of letters he and I had mutually picked out on maps, and in which I had just arrived early that morning after driving all night from Stockholm), might spot me as I lurked way out at the furthest tip of the very long pier, rather than merely bumping into me at random as we both walked around exploring the stores and streets and parks of this unknown hamlet

As its length suggests, this is a very detailed personal memory from many years ago (and indeed, I have merely sketched it for readers here I could write pages about it), and might at first seem to be nothing at all like a mental category. And yet, how else can one explain the fact that the image of myself standing at pier's end tingling with unrealistic hope jumped instantly to mind some fifteen years later as I was idly seeking to rearrange the eight letters in the last name of Janet Kolodner, a new acquaintance, in such a way that they would spell a genuine English word? Without success, I had tried dozens of fairly "obvious" pathways, such as "rendlook," "leodronk," and "ondorkle," when out of the blue it occurred to me that the initial consonant cluster "kn", with its cleverly silent "k," might be the key to success, and I started excitedly trying this "brilliant idea." However, after exploring this strategy for a while, I realized, to my chagrin, that no matter how lovely it would be if the silent "k" were to yield a solution, the probabilities for such a clever coup were rapidly diminishing. And at the precise instant that this realization hit, the Frederikssund-pier image came swooshing up out of memory, an image to which I had devoted not even a split second of thought for many years.

There was, of course, a perfectly logical reason behind this sudden resurfacing — namely, a strong and rich analogy in which the mundane idea of merely walking around the fishing village mapped onto the mundane exploration of "rendlook" and cousins, in which the "romantic" idea of lingering way out at the tip of the pier mapped onto the "romantic" hope for an anagram beginning with the tricky "kn" cluster, and in which the growing recognition of the likelihood of failure of the more unlikely, more "romantic" strategies was the common core that bound the two otherwise remote events together.

The Central Cognitive Loop

Abstract reminders of this sort have been noted here and there in the cognitive-science literature, and some attempts have been made to explain them (e.g., Roger Schank's *Dynamic*

Memory [1982]), but their starring role in the phenomenon of cognition has not, to my knowledge, been claimed. It is my purpose to stake that claim.

To make the claim more explicit, I must posit that such a large-scale memory chunk can be thought of as being stored in long-term memory as a "node" — that is, something that can be retrieved as a relatively discrete and separable whole, or to put it metaphorically, something that can be pulled like a fish out of the deep, dark brine of dormant memory. Once this "fish" has been pulled out, it is thrown in the "bucket" of short-term memory (often calling "working memory"), where it is available for scrutiny.

Scrutiny consists in the act of "unpacking" the node to some degree, which means that inside it are found other nodes linked together by some fabric of relationships, and this process of unpacking can then be continued recursively, given that the contents of unpacked nodes themselves are placed in short-term memory as well, and hence are themselves subject to more detailed scrutiny, if so desired. (I suppose one could extend the fishing analogy by imagining that smaller fish are found in the stomach of the first fish caught, as it is "cleaned" — and so forth, recursively. But that fanciful and somewhat gory image is not crucial to my story.)

Thus, if it is placed under scrutiny, inside the "Frederikssund pier" node can be found nodes for the exchange of letters that preceded Robert's and my Danish reunion, for Frederikssund itself, for my Stockholm drive, for Robert's train trip, for a few of the town's streets and shops, for the pier, for my growing disappointment, and so on. Not all of these will be placed into short-term memory each time the event as a whole is recalled, nor will the inner structure of those nodes that are placed there necessarily be looked into, although it is quite possible that some of their inner structure will be examined.

Thus the unpacking process of this kind of high-level unlabeled node (such as the "Frederikssund pier" node or the "Kolodner anagram" node) can fill short-term memory with a large number of interrelated structures. It must be stressed, however, that the unpacking process is highly context-dependent (i.e., sensitive to what concepts have been recently activated), and hence will yield a somewhat different filling-up of short-term memory on each occasion that the same high-level node is pulled up out of the ocean of long-term memory.

Once there are structures in short-term memory, then the perceptual process can be directed at any of them (this is, in fact, the kind of high-level perception that forms the core of the Copycat and Tabletop models of analogy-making — see Hofstadter and FARG 1995), the upshot of which will be the activation — thanks to analogy — of further nodes in long-term memory, which in turn causes new "fish" to be pulled out of that brine and placed into short-term memory's bucket. What we have described is, in short, the following central cognitive loop:

A long-term memory node is accessed, transferred to short-term memory and there unpacked to some degree, which yields new structures to be perceived, and the high-level perceptual act activates yet further nodes, which are then in turn accessed, transferred, unpacked, etc., etc.

An Illustration of the Central Cognitive Loop in Action

The foregoing may seem too abstract and vague, and so to make the ideas more concrete, I now will present a dialogue most of which actually took place, but some of which has been added on, so as to make some points emerge a little more clearly. The fact, however, that it all sounds perfectly normal is what matters — it certainly could pass for spontaneous cognition in the minds of two speakers. The dialogue exemplifies all the processes so far described, and — at least to my mind — shows how these processes are what drives thought. So here is the dialogue.

A and B are walking by a church when A looks up and notices that on the steeple, there are some objects that look like emergency-warning sirens attached to the base of the cross.

A: Hey, fancy that! Shades of "Praise the Lord and pass the ammunition!"

B: What do you mean?

A: Well, it's kind of amusing to me. On the one hand, the cross implies a belief in protection by the Lord, but on the other hand, the sirens suggest the need for a backup system, some kind of safety net. I mean, it's fine to believe in divine protection and all, but when it really comes to the crunch, religious people's true colors emerge...

B: Well, sooner safe than sorry, no?

A: Sure, but isn't a cross wrapped in danger sirens kind of hypocritical? I mean, why don't religious people put their money where their mouth is? If they really believe in God's benevolence, if they really have the courage of their own convictions, then how come it doesn't suffice to speak softly — why do they need to carry a big stick as well? Put it this way: Either you're a believer, or you ain't.

B: That's a bit black-and-white, isn't it?

A: Of course! As it should be! You can't have it both ways. Somehow this reminds me of when I had to leave my bags in a hotel in Italy for a few days, and the hotel people stored them in a tiny little chapel that was part of the hotel. A friend joked, "Well, this way they'll be protected." But why is such a remark so clearly a joke, even to religious people? Aren't churches houses of God? Shouldn't a sacred place be a safer place?

B: Yes, but being sacred doesn't make churches immune to disaster. We've all heard so often of churches whose roofs collapse on the assembled parishioners...

A: Exactly. And then pious people always say, "The Lord works in mysterious ways... It's beyond our comprehension." Well, how they can continue to believe after such an event is beyond my comprehension, that's for sure.

B: You're talking about people who claim to believe but in some sense act as if they don't really believe, deep down. But then there's the flip side of the coin: people who claim not to believe but act in a way as if they do. The reverse type of hypocrite, in sort.

A: Do you have an example in mind?

B: Yes — Niels Bohr, the great Danish physicist. I once read that in his house there was a horseshoe hanging over one door, and someone asked him, "What's this all about?" Bohr answered, "Well, horseshoes are supposed to bring good luck, so we put it up there." The friend then said, "Come now — surely you don't believe it brings good luck, do you?" Bohr

laughed and said, "Of course not!" And then he added, "But they say it works even if you don't believe in it."

A: I see your point — in a way Bohr's remark is the flip side of "Praise the Lord and pass the ammunition." In the trench-warfare case, you have a believer whose actions reveal deep doubts about their proclaimed belief, and in the Bohr case, you have a skeptic whose actions reveal that he may doubt his own skepticism. But that cross with the sirens — I just can't believe that they would wrap them around the cross, of all things — that's the height of irony! I mean, it's like some priest who's going into a dangerous area of town and doesn't just carry a handgun along in case of need, but in fact a cross that doubles as a handgun.

B: You've made the irony rather clear, I agree. But tell me — would you propose that the pope, simply because he's a big-time believer in God, should travel through the world's cities without any protection? Would you propose that true believers, if they are to be self-consistent, shouldn't put locks on their churches?

A: Well, won't God take care of his flock? Especially the pope?

B: It's not that simple.

A: Come on — if God doesn't look after the pope, who does he look after?

B: Come on, yourself! They crucified Jesus, didn't they? If anyone should have had divine immunity, it was Jesus — but he didn't. And yet that in itself doesn't mean that Jesus wasn't God's son.

This exchange illustrates all of the themes so far presented. In the first place, it shows A and B using ordinary words — bite-size lexical items such as "cross," "sirens," "bags," "hotel," "when," "people," and dozens more — nouns, verbs, adjectives, adverbs, prepositions, and so forth. Nothing unusual here, of course, except that readers are being exhorted to picture each of these words as the tip of an iceberg that hides a myriad hidden analogies — namely, the analogies that collectively allowed the category to come into being in the first place in the speaker's or listener's or reader's mind.

In the second place, the dialogue shows a good number of the shorter phrases cited in lists above being used in realistic situations — smallish stock phrases such as "Fancy that!" "kind of," "I mean," "ain't," "that's for sure," and many more. These phrases are used by speakers because they meet the rhetorical needs of the particular context, and when perceived by listeners they activate familiar rhetorical-context categories.

In the third place, the dialogue illustrates high-level perception — the retrieval of high-level labels for perceptions — such as A's opening statement, in which the lexical item "Praise and Lord and pass the ammunition" is the effortlessly evoked label for a church cross with warning sirens attached to it. In fact, all through the dialogue, the participants use large lexical items to label situations that are being categorized in real time in their minds. Thus we hear "backup system," "safety net," "when it really comes to the crunch," "flip side of the coin," "put your money where your mouth is," "sooner safe than sorry," "speak softly and carry a big stick," "black-and-white," and many more.

In the fourth place, we have large-scale remindings. First there is the shift from the cross-wrapped-in-sirens scene to the suitcases-left-in-hotel-chapel situation, then the shift to the collapsing-churches scenario, after which comes the shift, mediated by a kind of conceptual reversal, to Niels Bohr's horseshoe-that-works-despite-skepticism (probably an apocryphal story, by the way). Following that image comes a different kind of shift — an analogy where a given, known scenario is compared with a spontaneously concocted hypothetical scenario — thus, for instance, the cross-wrapped-in-sirens scene is compared with a hypothetical cross/handgun blend. This is swiftly followed by a trio of further concocted analogues: first the pope traveling without protection, then churches that are left unlocked, and finally God not even taking care of his own son.

The Central Cognitive Loop in Isolation and in Interaction

The broad-stroke pathway meandering through the limitless space of potential ideas during the hypothetical conversation of A and B is due to various actual scenes or imagined scenarios being reperceived, in light of recently activated concepts, in novel fashions and thereby triggering dormant memories, which are then fished up from dormancy to center stage (i.e., short-term memory), where, partially unpacked, they are in turn subjected to the exact same

context-dependent reperception process. Around and around in such a loop, alternating between fishing in long-term memory and unpacking and reperceiving in short-term memory, rolls the process of cognition.

Note that what I have just described is not problem-solving, which has traditionally played such a large role in modeling of thought and been tightly linked with "analogical reasoning"; no, everyday thought is not problem-solving or anything that resembles it at all; rather, it is a nonrandom stroll through long-term memory, mediated by high-level perception (which is simply, to echo myself, another name for analogy-making).

To be sure, thought does not generally take place in a sealed-off vat or an isolation chamber; most of the time, external events are constantly impinging on us. Therefore the purely self-driven flow that the "central loop" would suggest is just half of the story — it is the contribution from within one's private cognitive system. The other half — the contribution from outside — comes from inanimate objects impinging on one's senses (skyscrapers and sunsets and splashes, for instance), from animate agents seen mostly as objects (mosquitos that one swats at, people that one tries not to bang into as one hastens down a crowded sidewalk), or from other cognitive agents (conversations with friends, articles read in the paper, email messages, scenes in movies, and so on).

This buzzing, booming confusion in which one is immersed most of the time tends to obscure the constant running of the private inner loop — but when one retreats into solitude, when one starts to ponder or daydream, when one tries to close oneself off from these external impingements and to be internally driven, that is when the above-posited "central loop of cognition" assumes the dominant role.

Goal-Drivenness and the Central Loop

Where do goals enter this picture? How does the deeply goal-driven nature of human thought emerge from what might seem to be the randomness of the posited central loop? The answer resides in the enormously biased nature of each individual's perception.

Each person, as life progresses, develops a set of high-level concepts that they tend to favor,

and their perception is continually seeking to cast the world in terms of those concepts. The perceptual process is thus far from neutral or random, but rather it seeks, whenever possible, to employ high-level concepts that one is used to, that one believes in, that one is comfortable with, that are one's pet themes. If the current perception of a situation leads one into a state of cognitive dissonance, then one goes back and searches for a new way to perceive it. Thus the avoidance of mental discomfort — the avoidance of cognitive dissonance — constitutes a powerful internal force that helps to channel the central loop in what amounts to a strongly goal-driven manner.

The Sapir-Whorf Hypothesis: Language and the Central Loop

The viewpoint I have been proposing here — in most ways quite unrevolutionary! — can be rephrased in terms of "perceptual attractors," which are long-term mental loci that are zoomed into when situations are encountered (see Kanerva 1988). We all have many thousands of such attractors in our dormant memories, only a tiny fraction of which are accessed when we encounter a new situation. Where do such attractors come from? How public are they? Do they have explicit labels? Here I list three main types:

- Standard lexical items (words, names, phrases, proverbs, etc.) provided to a vast public through a shared linguistic environment
- Shared vicarious experiences provided to a vast public through the media (e.g., places, personages, and events of small and large scale in books, movies, television shows, and so on), the smaller of which have explicit linguistic labels, the more complex of which have none
- Unique personal memories, lacking any fixed linguistic labels (such chunks are generally very large. and complex, like the Frederikssund memory discussed above, or even far larger events, such as a favorite high-school class, a year spent in a special city, a protracted divorce, and so on)

Since a sizable fraction of one's personal repertoire of perceptual chunks is provided from without, by one's language and culture, this means that inevitably language and culture exert powerful, even irresistible, channeling influences on how one frames events. (This position is related to the "meme's-eye view" of the nature of thought, as put forth in numerous venues,

most recently in Blackmore 1999.)

Consider, for instance, such words as "backlog," "burnout," "micromanaging," and "underachiever," all of which are commonplace in today's America. I chose these particular words because I suspect that what they designate can be found not only here and now, but as well in distant cultures and epochs, quite in contrast to such culturally and temporally bound terms as "soap opera," "mini-series," "couch potato," "news anchor," "hit-and-run driver," and so forth, which owe their existence to recent technological developments. So consider the first set of words. We Americans living at the millennium's cusp perceive backlogs of all sorts permeating our lives — but we do so because the word is there, warmly inviting us to see them. But back in, say, Johann Sebastian Bach's day, were there backlogs — or more precisely, were backlogs perceived? For that matter, did Bach ever experience burnout? Well, most likely he did — but did he know that he did? Or did some of his Latin pupils strike him as being underachievers? Could he see this quality without being given the label? Or, moving further afield, do Australian aborigines resent it when their relatives micromanage their lives? Of course, I could have chosen hundreds of other terms that have arisen only recently in our century, yet that designate aspects of life that were always around to be perceived but, for one reason or another, aroused little interest, and hence were neglected or overlooked.

My point is simple: we are prepared to see, and we see easily, things for which our language and culture hand us ready-made labels. When those labels are lacking, even though the phenomena may be all around us, we may quite easily fail to see them at all. The perceptual attractors that we each possess (some coming from without, some coming from within, some on the scale of mere words, some on a much grander scale) are the filters through which we scan and sort reality, and thereby they determine what we perceive on high and low levels.

Although this sounds like an obvious tautology, that part of it that concerns words is in fact a nontrivial proposition, which, under the controversial banner of "Sapir-Whorf hypothesis," has been heatedly debated, and to a large extent rejected, over the course of the twentieth century. I myself was once most disdainful of this hypothesis, but over time came to realize how deeply human thought — even my own! — is channeled by habit and thus, in the last accounting, by the repertoire of mental chunks (i.e., perceptual attractors) that are available to the thinker. I now think that it is high time for the Sapir-Whorf hypothesis to be reinstated, at least in its

milder forms.

Language, Brains, and "Just Adding Water"

The usual goal of communication is, of course, to set up "the same thought" in the receiver's brain as is currently taking place in the sender's brain. The mode by which such replication is attempted is essentially a drastic compression of the complex symbolic dance occurring in the sender's brain into a temporal chain of sounds or a string of visual signs, which are then absorbed by the receiver's brain, where, by something like the reverse of said compression — a process that I will here term "just adding water" — a new symbolic dance is launched in the second brain. The human brain at one end drains the water out to produce "powdered food for thought," and the one at the other end adds the water back, to produce full-fledged food for thought.

Take, for instance, the paragraph given a few pages back:

that time I spent an hour or two hoping that my old friend Robert, who I hadn't seen in two years but who was supposed to arrive from Germany by train sometime during that summer day in the little Danish fishing village of Frederikssund (which in a series of letters he and I had mutually picked out on maps, and in which I had just arrived early that morning after driving all night from Stockholm), might spot me as I lurked way out at the furthest tip of the very long pier, rather than merely bumping into me at random as we both walked around exploring the stores and streets and parks of this unknown hamlet

Obviously, this set of black marks on a white background is not similar to the time I spent in Frederikssund, nor is any part of it similar to a pier, a drive from Stockholm, a body of water, or dashed hopes. And yet these marks triggered in your brain a symbolic dance so vivid that you saw, in your mind's eye, a fishing village, two young friends, their joyful anticipation of a semirandom reunion, a pier stretching far out into a gulf, a barely visible person anxiously pacing at its tip, and so on. A never-before-danced dance inside your brain, launched by a unique set of squiggly shapes, makes you feel almost as if you had been there; had I spelled it out with another page or two of intricate black-on-white patterns, it would feel all the more vivid. This is a wonderful kind of transportation of ideas between totally different media — uprooting ideas from one garden and replanting them in a garden never even imagined before, where they flourish beautifully.

Transportation

In his book *The poetics of translation* (Barnstone 1993), poet and translator Willis Barnstone has a section called "The Parable of the Greek Moving Van," where he points out that on the side of all Greek moving vans is written the word μεταφορά (phonetically "metafora" and semantically "transportation"). He then observes:

To come to Greece and find that even the moving vans run around under the sun and smog of greater Athens with advertisements for transportation, for metaphor, and ultimately with signs for translation should convince us that every motor truck hauling goods from one place to another, every perceived metamorphosis of a word or phrase within or between languages, every decipherment and interpretation of a text, every role by each actor in the cast, every adaptation of a script by a director of opera, film, theater, ballet, pantomime, indeed every perception of movement and change, in the street or on our tongues, on the page or in our ears, leads us directly to the art and activity of translation.

I pack my mental goods down into tight, neat bundles, I load them as carefully as I can into the metafora truck of language, it drives from my brain to yours, and then you unpack. What a metaphor for communication! And yet it has often been said that all communication, all language, is metaphorical. Since I believe that metaphor and analogy are the same phenomenon, it would follow that I believe that all communication is via analogy. Indeed, I would describe communication this way: taking an intricate dance that can be danced in one and only one medium, and then, despite the intimacy of the marriage of that dance to that medium, making a radically new dance that is intimately married to a radically different medium, and in just the same way as the first dance was to its medium.

Trans-Sportation

To make this all a little more concrete, let us consider taking a complex dance done in the medium of the sport of basketball and trans-sporting that dance into the rather different medium of the sport of soccer. Indeed, imagine taking the most enthralling basketball game you ever watched — perhaps a championship game you saw on television — and giving a videotape of that game to a "soccer choreographer," who will now stage all the details of an

artificial soccer game that is in some sense analogous to your basketball game. Of course this could be done in many ways, some conservative and some daring.

Some choreographers, citing irreconcilable differences between the two sports (for instance, the difference in the number of players per team, the lack of any counterpart to a goalie in basketball, the low frequency of scoring in soccer relative to basketball, and on and on), might severely bend the rules of soccer, creating a game with only five players on a team, taking away the goalies, vastly reducing the size of the field (and the goals), and so forth, thus effectively creating a hybrid soccer-basketball game that looks very much like basketball, only it is played on grass and involves propelling the ball with the lower rather than the upper limbs. When one watched the reenactment of one's favorite basketball game in this artificial medium, one would not have the sense of watching a soccer game but of watching a very distorted basketball game.

Other choreographers, more willing to go out on a limb, would retain the normal rules of soccer but would attempt to stage a game whose every play felt like a particular play of the original basketball game, even though eleven players were retained on a side, even though the goals remained huge compared to baskets, even though there were still goalies, even though the goals might be coming a little too thick and fast, and so forth. There would be plays that would be essentially like slam-dunks while at the same time looking every bit like normal soccer plays. In such a case, one would feel one was watching a genuine soccer game — perhaps a peculiar one in some ways, but nonetheless genuine. In the ideal case, one could have the two counterpart games running on side-by-side television screens, and a "neutral" commentator using only terms that apply to both sports could be effectively heard as describing either of the games.

Anything in between these two extreme philosophies of "trans-sportation" can also be imagined — and just such a bizarre scenario is what I think everyday communication is actually like. Two brains are, in general, far more unalike than are the sports of soccer and basketball — and yet our society is predicated on mutual comprehensibility mediated by language.

Translation

It is astonishing to me how often people — even linguistically sophisticated people, such as philosophers, writers, linguists, translators, and cognitive scientists — will speak as if communication among members of a single language community were total and perfect, with serious communication gaps only taking place at the interface between different languages — as if translation were needed only between languages, never within a language community. Thus it is taken as obvious and indisputable that Russians all read, say, a novel by Pushkin in one and the same way, but that no one who reads an anglicized version of that novel could possibly get anything like "that same experience" (as if the reading of that novel engendered just one experience in the vast world of all different Russian speakers). My retort would be that what matters is not the dried linguistic powder that is used to transport the dance between brains — what matters is the dance set up inside a brain by whatever dried powder is used for the transport. Linguists (I exempt those in the very recent cognitive-linguistics movement) concentrate so hard on the overt dried powder that they wind up largely ignoring the covert dances that engender it, and that it engenders. As an ironic consequence, the standard model of language that has been built up this century by linguists is hugely impoverished.

Most people's (and most linguists') model of translation is as dry as the powder that carries dehydrated ideas from brain to brain; indeed, they conceive of translation as a mapping from one purely dehydrated chain of symbols to another dehydrated chain of symbols, without any need for "adding water" at any stage of the process. The whole process happens purely at the level of the dry symbols. Translation would thus be an activity for drones — and hence ideal for computers to carry out. Here — courtesy of my Sofia hotel — is an example of the "drone" theory of translation:

> *Please*   *disturb*
>   ↕   ↕   ↕
> *Bitte*   *nicht*   *stören*

Or, as the early machine-translation pioneer Warren Weaver once wrote (Weaver 1955), "When I look at an article in Russian, I say, 'This is really written in English, but it has been coded in some strange symbols. I shall now proceed to decode.'"

Since translation is but the challenge of communication rendered crystal-clear, and since

communication is but metaphor, and since metaphor is but analogy, I shall spend the rest of this article on analogy focusing on translation and showing how at its core translation is analogy, and indeed, is analogy at its most sublime and enchanting.

Evgeniĭ Onegin

When, a few paragraphs back, I wrote the phrase "a novel by Pushkin," my choice was not as flippant as I tried to make it seem. Indeed, as a recent translator of Alexander Pushkin's novel in verse *Eugene Onegin,* I have been totally absorbed over the past year or so in the delicious but daunting task of reincarnating Pushkin's sparkling poetry in the medium of contemporary English — or rather, contemporary American. It has not, needless to say, been a process that looked anything like the Sofia-hotel model, with the two-headed vertical arrows connecting words. In order to give a sense of what was involved, I must first describe the building blocks of the novel, usually called "Onegin stanzas." Each sonnet (of which there are nearly four hundred) is a "crystal" — a pattern to transplant from one medium to another. What is the nature of these crystals?

To begin with, each one consists of fourteen lines of strict iambic tetrameter (which means — at least in Russian — that stresses never fall on odd-numbered syllables). The rhyming pattern is always as follows:

A B A B C C D D E F F E G G

and within this framework, the "A," "C," and "E" line-pairs have the special property of being feminine rhymes, while the "B," "D," "F," and "G" line-pairs are masculine. The distinction is as follows: "return/discern" is a masculine rhyme, because the final syllables not only rhyme but are stressed, whereas "returning/discerning" is a feminine rhyme, because the penultimate syllables rhyme and are stressed, while the final syllables are not only unstressed but identical. In other words, in feminine rhymes, the "rhyming action" takes place before the line's final syllable (which is unstressed), whereas in masculine rhymes, the rhyming action takes place on the final syllable (which is stressed).

As a consequence of this intricate design, an Onegin stanza's lines have varying numbers of

syllables, depending on whether they are feminine or masculine. The six "A," "C," and "E" lines have nine syllables apiece, while all others have eight, as follows:

9 8 9 8 9 9 8 8 9 8 8 9 8 8

All four hundred crystals in the original Russian have this property, and thus all four hundred crystals in the counterpart work in English should — should they not? — have this same property. The crucial question is, of course, what kind of compromises should be made in the transportation of Pushkin's virtuoso game into the new medium. One type of translator (Nabokov 1964) might insist on retaining the most literal possible rendering of each word and even much of the Russian word order, in which case all rhyming and rhythmic properties would have to be sacrificed. This would seem rather akin to the word-for-word Sofia-hotel model, and quite uninspired as a translation philosophy.

Another type of translator would insist on retaining the medium-message marriage that well-wrought poetry inevitably is, and thus on looking behind the scenes, looking beyond the dry dust on the paper, looking at the sparkling mental dance to which the dry powder gives rise, once water is added to it. To such a translator, what matters is that each semantic chunk of the original poetry (whether contained within a single line or spread across several) gives rise to a scene in the mind's eye of a reader (not to mention that of Pushkin), and this type of translator, having tried to envision that scene as clearly, fully, and. faithfully as possible, then uses it as a source for English words and phrases that can be used in lines of English poetry that obey. the formal constraints. Such a translator, in short, is inspired by the inner dance and not merely by the dry powder.

Since the scene conjured up by a line or two of the original goes far beyond the literal words in those lines (i.e., since "just adding water" adds such richness!), there is much more to draw on as potential material for a new poem in English, and so one is enormously freed up. There remain, of course, all the rhythmic and rhyming hoops to jump through, but by adding water, one has at least given oneself a fighting chance at finding a solution satisfying all the relevant constraints.

Of course satisfying those constraints is not a simple task, nor is it by any means a black-and-

white matter to judge whether (or to what degree) the constraints have actually been met. There are many pressures vying with each other, and by no means are they all explicit, although some are. One might cite the following sets of pressures under which a translator must work:

- Content: the image evoked by the words and phrases in a semantic chunk
- Structural pattern: the above-described features that define the phonetic nature of an Onegin stanza
- Tone: an intangible brew of subliminally felt qualities suggested by the following oppositions: humorous vs. serious; straightforward vs. ironic; heavy vs. light; old-fashioned vs. modern; meditative vs. peppy; sweet vs. sad; resigned vs. delighted; highbrow vs. lowbrow; etc.

The only one of these constraints that has a sharp, black-and-white feel to it is that of the structural pattern, since it is generally fairly objective whether two words rhyme or not, how many syllables are in a word, where stress should fall, and whether a given two-syllable chunk is an iamb or not (although, in truth, these matters are surprisingly often quite blurry — does "midnight" make a true feminine rhyme with "slid right"? is "finally" bisyllabic or trisyllabic?).

The other constraints are anything but sharp, since the content of any lexical item is (as has been the thrust of this paper) determined by a host of prior analogies, and hence is tremendously blurry, and since tone is not only vague but also highly multidimensional, allowing for any conceivable combination of degree of irony, degree of modernity, degree of sadness, and so on, ad infinitum.

Given the complexity of this range of competing pressures, it is hardly surprising that there will occur, in the translation of nearly every single line, smaller or larger creative slippages, typified by, but by no means limited to, the following list:

- The "perfect" literal translation of a word is abandoned in favor of a slightly less perfect choice, because of (say) phonetic constraints
- A syntactic reversal, slightly unusual in English, is resorted to for (say) reasons of metric purity

- An idea or image is shifted from one line to another because English grammar works that way
- An alliterative pattern is dropped in one stanza but is introduced out of the blue in another, in order to replicate fairly accurately the overall density of alliteration in the original
- A modern-seeming word is used in a passage that has an older tone, or vice versa, because of (say) certain extra connotations that are gained thereby
- A word strongly evocative of something linked tightly to the target culture but not the original culture (e.g., "jive") is used, even if the effect creates a very short-lived subliminal shift of venue from source to target culture
- A perfect rhyme is sacrificed for a near-rhyme, in order to gain an extra set of connotations or to conjure up a precise image that would otherwise not be attainable
- A word is used in a highly metaphorical manner, stretching it even beyond its normal degree of plasticity
- A metaphor is dropped or is replaced by a different metaphor, because the original metaphor makes no sense in the target culture
- A metaphor is introduced out of the blue, perhaps because it is implicit in a stock phrase or proverb that fits aptly and that also rhymes very strongly;
- Etc., etc., etc.

The amusing fact about the result of all these kinds of creative slippage is that what emerges can often be so powerfully evocative of the original that it seems — at least on some levels — perfectly plausible to refer to the English-language Onegin stanza thereby produced as being "by Alexander Pushkin," and therefore to write those three words on the front cover and spine and title page of the book, perhaps even relegating the translator's name to nothing more than a line in fine print on the copyright page.

We shall now take a look at the results of all these kinds of slippages caused by multiple rival pressures in the minds of different translators with different philosophies of translation. I have selected one stanza, the 29th from chapter II, to illustrate what kinds of things can occur. (See also chapters 8, 9, and 13 of Hofstadter 1997.)

First I display Pushkin's original Russian and, next to it, a literal translation by Vladimir

Nabokov; thereafter, in order, stanzas by the following translators (in the chronological order of publication of their translations): Babette Deutsch, Oliver Elton, Walter Arndt, Charles Johnston, James Falen, and Douglas Hofstadter.

Александр Пушкин (1825)

Ей рано нравились романы;
Они ей заменяли все;
Она влюблялася в обманы
И Ричардсона и Руссо.
Отец ее был добрый малый,
В прошедшем веке запоздалый;
Но в книгах не видал вреда;
Он, не читая никогда,
Их почитал пустой игрушкой
И не заботился о том,
Какой у дочки тайный том
Дремал до утра под подушкой.
Жена ж его была сама
От Ричардсона без ума.

Vladimir Nabokov (1964)

She early had been fond of novels;
for her they replaced all;
she grew enamored with the fictions
of Richardson and of Rousseau.
Her father was a kindly fellow
who lagged in the precedent age
but saw no harm in reading books;
he, never reading,
deemed them an empty toy,
nor did he care
what secret tome his daughter had
dozing till morn under her pillow.
As to his wife, she was herself
mad upon Richardson.

Babette Deutsch (1936)

She found in a romantic story
All one might care to be or know;
S
In Richardson as in Rousseau.
Her father saw no harm in reading
(He was a decent chap, conceding
He lived in quite another age);
But then he never read a page.
He did not know that books could say things

Oliver Elton (1937)

Romances were her only passion,
And all the world to her; and so
She fell in love, the dupe of fashion,
With Richardson, and with Rousseau.
Quite a good fellow was her father,
Of the last age, belated rather;
He saw no mischief in a book,
Though in one he would never look;
Thought it a toy, and held it lightly,

To move you even while you slept;
He thought the tomes his daughter kept
Beneath her pillow, empty playthings;
While, on the other hand, his wife
Held Richardson as dear as life.

And cared not what his daughter did
When she a private volume hid
Beneath her pillow, slumbering nightly.
His lady wife was mad upon
The tales of Samuel Richardson.

Walter Arndt (1963)

But novels, which she early favored,
Replaced for her all other treats;
With rapturous delight she savored
Y   Y
Her honest father, though old-fashioned,
S
About the harm that books might breed;
He, who was never known to read,
Regarded them as empty thrillers
And never thought to bring to light
Which secret volume dreamt at night
Beneath his little da
His wife had had a crush herself
On Richardson   still on her shelf.

Charles Johnston (1977)

From early on she loved romances,
they were her only food... and so
she fell in love with all the fancies
of Richardson and of Rousseau.
Her father, kindly, well-regarded,
but in an earlier age retarded,
could see no harm in books; himself
he never took one from the shelf,
thought them a pointless peccadillo;
and cared not what his daughter kept
by way of secret tome that slept
until the dawn beneath her pillow.
His wife, just like Tatyana, had
on Richardson gone raving mad.

James Falen (1990)

From early on she read romances,
And novels set her heart aglow;
She loved the fictions and the fancies
Of Richardson and of Rousseau.
Her father was a kindly fellow
Lost in a past he found more mellow;

Douglas Hofstadter (1999)

From early on, she read romances;
True life they were for her, not show.
She fell for all the moods and trances
Induced by authors like Rousseau
And Richardson. A friendly fellow,
Her father was old-fashioned, mellow,

But still, in books he saw no harm,
H
Deemed it a minor peccadillo;
Nor did he care what secret tome
His daughter read or kept at home
Asleep till morn beneath her pillow;
His wife herself, we ought to add,
For Richardson was simply mad.

And saw in books no cause for dread;
Instead, because he never read,
He thought of them as dull and boring,
H
What brand of frivolous flim-flam
His daughter clutched all night while snoring.
But on the other hand, his wife
Thought Richardson the spice of life.

Each of these compact fourteen-line verbal packets is a structure that bears to the original packet the relation of analog in the medium of the English language. Each one is clearly the result of myriad tradeoffs involving preservation of imagery, strictness of meter, perfection of rhyme, phonetic patternedness, era exuded by words and phrases, degree of humor, degree of catchiness, degree of familiarity of lexical items, ease of syntactical flow, sequential order of ideas, and much more.

Take, for instance, the word "instead" on line 8 of my translation. Initially, I had line 8 beginning with "indeed," which in some ways is stronger (because "indeed" carries a more emphatic flavor than "instead," and also because, more subtly, the comma-signaled pause that would follow "indeed" strikes me as ever-so-slightly longer and more charged than its counterpart with "instead"), and yet despite these lures, the internal rhyme of "instead" with "dread" preceding it and with "read" following it somehow carried the day in my mind. This is typical of the multidimensional internal conflicts that occur routinely in translation, and each time, one has to weigh all the factors and make a decision.

On a more blatant semantic level, you may note that the imagery in my stanza is that of the daughter sleeping — nay, snoring! — while clinging tightly to a favorite book, whereas the imagery in the original is of the book itself sleeping (or dreaming) beneath the girl's pillow. To what extent is one entitled to manipulate imagery this way and then to claim that the resulting book is "by Alexander Pushkin"?

To what extent did the nonanglophone Russian poet Alexander Pushkin ever say, in describing the girl's father, "and didn't give a tinker's damn"? On the other hand, to what extent did Pushkin ever write the line "nor did he care"? Using the former as line 10 provides a clear whiff of Pushkinesque humor (not to mention having the proper meter, rhyme, and so on), while using the latter as line 10 is bland and flat (and is but four syllables long, where Pushkin's line had, of course, four full iambs on it).

By what right did I feel entitled to insert an alliterative and flippant phrase like "frivolous flim-flam" into the lyrical mouth of Alexander Pushkin? On the other hand, by what right did Vladimir Nabokov feel entitled to insert the graceless and nonidiomatic phrase, "As to his wife, she was herself mad upon Richardson" into the mouth of Russia's greatest poet? For that matter, by what right did James Falen think he could get away with a pseudorhyme like "romances" and "fancies"? As for Walter Arndt, by what right did think he could get away with a nonrhyme like "thrillers" and "pillows"? And by what right did Charles Johnston think he was entitled to portray the girl's novels as "her only food"? By what right did Babette Deutsch rearrange the order of Pushkin's ideas, so as to make the father's old-fashionedness follow (and in parentheses, to boot!) his seeing no harm in books? And how in the world did Oliver Elton feel he was being faithful to Pushkin by using the bizarrely redundant phrase "his lady wife"?

Of course I am feigning outrage here; I have great respect for most of these translators, despite the fact that I see compromises ubiquitously riddling the productions of each of them. The questions just raised were raised purely rhetorically, my intention being to provoke readers into pondering which of these seven rival English-language stanzas might be seen as the most analogous to Pushkin's original stanza (without claiming there is any "correct" answer).

Winding Up: On Associationism and the Cartesian Theater

We have come a long way, starting out by seeing single words as analogs to perceived situations, and ending up by seeing sonnets in different languages as each other's analogs. Somewhere near the midpoint came the crux of this essay, however, which claimed that thinking (at least when isolated from external influences) is a series of leaps involving high-level perception, activation of concepts in long-term memory, transfer to short-term memory,

partial and context-dependent unpacking of chunks, and then further high-level perception, and so forth.

This may sound like no more than the age-old idea of associationism — that we think by jumping associatively from one thing to another. If that's all it came down to, my thesis would certainly be a sterile and vapid noncontribution to cognitive science. But the mechanisms I posit are more specific, and in particular they depend on the transfer of tightly packed mental chunks from the dormant area of long-term memory into the active area of short-term memory, and on their being unpacked on arrival, and then scrutinized. Both transfer and perception are crucial, and in that respect, my thesis departs significantly from associationism.

Some readers, such as the author of *Consciousness explained* (Dennett 1991), might feel they detect in this theory of thinking an insidious residue of the so-called Cartesian theater — a hypothetical theater in which an "inner eye" watches as various images go parading by on a "mental screen," and becomes "aware" or "conscious" of such imagery. Such a notion of thinking leads very easily down the slippery slope of nested homunculi, and thus to an infinite regress concerning the site of consciousness.

. I would gladly plead guilty to the accusation of positing a "screen" upon which are "projected" certain representations dredged up from long-term memory, and I would also plead guilty to the accusation of positing an "inner eye" that scans that screen and upon it posts further representational structures, which trigger a descent via analogy into the dormant depths of long-term memory. I would insist, however, that the label "perception," as applied to what the "inner eye" does, be sharply distinguished from visual or any other kind of sensory perception, since in general it involves no sensory modality in any normal sense of the term (recall the perception of "ethnic cleansing" in a newspaper story). The nature of such abstract or high-level perceptual processing has been sketched out in work done by my students and myself over the years (see Hofstadter and FARG 1995), and I will not attempt to describe it here.

Clearly, since it has been implemented as a computer program (at least to a first approximation), such a model does not succumb to snagging on the fatal hook of infinite regress.

To those who would scoff at the very notion of any "inner screen" involved in cognition, I would point to the large body of work of perceptual psychologist Anne Treisman (e.g., Treisman 1988), which in my view establishes beyond any doubt the existence of temporary perceptual structures created on the fly in working memory (she cal\s them "object files") — a stark contrast to the connectionist-style thesis that all cognition takes place in long-term memory, and that it consists merely of simultaneous conceptual activations (possibly with attached temporal phases, so as to handle the "binding problem") without any type of transfer to, or structure-building in, a distinct working area. Although this more distributed view of the essence of cognition might appeal to opponents of the Cartesian theater, it does not seem to .me that it comes anywhere close to allowing the richness of thought that back-and-forth flow between long-term and short-term memory would allow.

I hope that my speculative portrayal of analogy as the lifeblood, so to speak, of human thinking, despite being highly ambitious and perhaps somewhat overreaching, strikes a resonant chord in those who study cognition. My most optimistic vision would be that the whole field of cognitive science suddenly woke up to the centrality of analogy, that all sides suddenly saw eye to eye on topics that had formerly divided them most bitterly, and naturally — indeed, it goes without saying — that they lived happily ever after. Whatever.

References

Barnstone, W. (1993). *The poetics of translation: History, theory, practice.* New Haven: Yale University Press.

Becker, J. D. (1975). The phrasal lexicon. In R. Schank and B. Nash-Webber, eds., *Theoretical Issues in Natural Language Processing.* Cambridge, MA: Bolt Beranek and Newman.

Blackmore, S. (1999). *The meme machine.* New York: Oxford University Press.

Dawkins, R. (1976). *The selfish gene.* New York: Oxford University Press.

Dennett, D. C. (1991). *Consciousness explained.* Boston: Little, Brown.

Hofstadter, D. (1997). *Le ton beau de Marot.* New York: Basic Books.

Hofstadter, D., and the Fluid Analogies Research Group (1995). *Fluid concepts and creative analogies.* New York: Basic Books.

Hofstadter, D., and Moser, D. (1989). To err is human, to study error-making is cognitive science.*Michigan Quarterly Review.* 28(2):185-215.

Kanerva, P. (1988). *Sparse distributed memory.* Cambridge, MA: MIT Press.

Pushkin, A. S. (1832). *Eugene Onegin: A novel in verse.*

W. Arndt, trans. (1963). Ann Arbor, MI: Ardis Editions.

B. Deutsch, trans. (1998). New York: Dover.

O. Elton, trans. (1995). London: Everyman.

J. Falen, trans. (1995). New York: Oxford University Press.

D. Hofstadter, trans. (1999). New York: Basic Books.

C. Johnston, trans. (1977). New York: Penguin Books.

V. Nabokov, trans. (1964). Princeton, NJ: Princeton University Press.

F. Sobotka (1991). London: Bristol Classical Press.

Schank, R. (1982). *Dynamic memory.* New York: Cambridge University Press.

Treisman, A. (1988). Features and objects: The fourteenth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology* 40A:201-237.

Weaver, W. (1955). Translation. In W. N. Lock and A. D. Booth, eds., *Machine translation of languages.* Cambridge, MA: MIT Press.

Previously published in:
The Analogical Mind: Perspectives from Cognitive Science,
Dedre Gentner, Keith J. Holyoak, and Boicho N. Kokinov (eds.).
Cambridge MA: The MIT Press/Bradford Book, 2001, pp. 499-538.

Illustrations by Douglas Hofstadter; used by permission.
(Previously unpublished.)