

Speechstuff and Thoughtstuff:

Musings on the Resonances Created by Words and Phrases via the Subliminal Perception of their Buried Parts

Douglas R. Hofstadter
Istituto per la Ricerca Scientifica e Tecnologica
loc. Pantè di Povo, I 38050 Trento, Italy

A Mysterious Piece of Fan Mail is Buried and Resurfaces

Several years ago I received an envelope from a reader of my writings, in which I found a few pages photocopied from a science-fiction magazine and a short note saying "You might enjoy this", or words to that effect, but no further explanation. I had heard of the piece's author, Poul Anderson, and knew he was a highly respected writer of science fiction, but had never read anything by him. I momentarily glanced at the strangely-titled piece but didn't even try to read it. It went into a pile of correspondence and remained thus buried for months. Finally one day it resurfaced, and as I was feeling a little less pressed, I decided to give it a chance. Here, then, are the first three paragraphs of Poul Anderson's "Uncleftish Beholding" (it is found in its entirety in Appendix I):

For most of its being, mankind did not know what things are made of, but could only guess. With the growth of worldken, we began to learn, and today we have a beholding of stuff and work that watching bears out, both in the workstead and in daily life.

*The underlying kinds of stuff are the **firststuffs**, which link together in sundry ways to give rise to the rest. Formerly we knew of ninety-two firststuffs, from waterstuff, the lightest and barest, to ymirstuff, the heaviest. Now we have made more, such as aegirstuff and helstuff.*

*The firststuffs have their being as motes called **unclefs**. These are mighty small: one seedweight of waterstuff holds a tale of them like unto two followed by twenty-two naughts. Most unclefs link together to make what are called **bulkbits**. Thus, the waterstuff bulkbit bestands of two waterstuff unclefs, the sourstuff bulkbit of two sourstuff unclefs, and so on. (Some kinds, such as sunstuff, keep alone; others, such as iron, cling together in chills when in the fast standing; and there are yet more yokeways.) When unlike unclefs link in a bulkbit, they make **bindings**. Thus, water is a binding of two waterstuff unclefs with one sourstuff unclef, while a bulkbit of one of the foreststuffs making up flesh may have a thousand or more unclefs of these two firststuffs together with coalstuff and chokestuff.*

At first I couldn't quite tell what was going on. With peculiar, alien words like "worldken" and "workstead", the first paragraph exuded a decidedly ancient aroma, but it was unclear exactly what it was talking about. No doubt it had something to do with "what things are made of", but the rest of it was really very blurry. What in the world could it mean to say "today we have a beholding of stuff and work"? I knew that science-fiction writers occasionally play with language quite creatively in order to give a story a particular flavor, and I was a bit reminded of Tolkien's mythical style of writing. Then in the second paragraph I came across "firststuff" and "waterstuff", which I recognized as heavy-handed anglicizations of the German terms *Grundstoff* and *Wasserstoff*, meaning "chemical element" and "hydrogen", respectively.

Splitting the Etym

Suddenly it crossed my mind that the whole piece might be some kind of systematic conversion, on a strict morpheme-by-morpheme basis, of some long-forgotten essay about chemistry (or possibly alchemy) from German into English. This seemed like a very interesting idea. In any case, my first stride forward made it fairly easy to figure out some of the nearby sentences.

Unfortunately, as I proceeded, my hypothesis started to get undermined. First of all, the second paragraph mentioned 92 elements as well as some new ones recently made. Since I knew that the transuranic elements — those beyond the “classic 92” — were discovered only this century, this was surely not a translation of some long-forgotten essay. Secondly, I knew that in German the only elements whose names end in *-stoff* are the most common and most familiar ones, such as oxygen and nitrogen and a few others, certainly not uranium and its heavier siblings — yet the paragraph used the terms “ymirstuff”, “aegirstuff”, and “helstuff” for the heaviest, newest elements. This made it seem distinctly less likely that Anderson’s linguistic game involved translating from German.

It then briefly flitted through my mind that in the language that might be called “Hitlerese” — that is, the Latin-free, Greek-free, pure Aryan German that Hitler promulgated during his Reich — every last element-name *would* have ended in *-stoff*, but this seemed an extremely unlikely basis for Anderson’s piece. And further weakening the German hypothesis was the fact that to render the idea of “element”, Anderson would almost certainly have written “groundstuff”, not “firststuff”, if he had really been trying to imitate the flavor of the German term *Grundstoff*. So scratch that idea.

My next guess was that Anderson was translating in some very literal manner from ancient Greek. After all, I knew that our word “hydrogen” comes from ancient Greek and that, when taken apart into its etymological atoms “hydro” and “gen”, it is isomorphic to “waterstuff”. It also seemed plausible to me that the Greeks had built their word for “element” out of two pieces meaning “first” and “stuff”. What further strengthened the Greek hypothesis for me was my decipherment of the eerie, almost biblical, opening sentence of the third paragraph: “The firststuffs have their being as motes called *unclefts*.” Not having grown up three hundred years ago, I had no idea what “motes” meant, but “uncleft” made me think “not split”. It didn’t take me too long then to recognize the analogy between this and our word “atom”, which fissions into the pieces “a” and “tom”, meaning respectively “without” and “part” in ancient Greek. So I inferred that the sentence must mean “Elements are composed of particles called atoms”, which gave me the meaning of “mote” in the context of this essay.

My hypothesis was now that Anderson’s essay was a morpheme-by-morpheme translation of an article about chemistry or physics from ancient Greek into English. Obviously, no such article was ever really written in ancient Greek, so it had to be a translation of an article as it *might* have been written in ancient Greek, if the ancient Greeks had known of atoms, molecules (“bulkbits”, in the essay), and other related physical and chemical phenomena. A bit strained, but certainly an interesting premise.

However, as I read on and thought further, I started seeing that the truly distinctive feature of Anderson’s text was its complete avoidance of Latin and Greek roots in any way, shape, or morph. For instance, even as standard and simple a Latin-based word of contemporary English as “number” was studiously avoided (replaced by “tale”, a very old word for “number” in English, cognate to *Zahl* in present-day German and *tal* in present-day Swedish). Similarly, the Latin-based words “rare” and “table” were avoided in favor of “seldom” and “board”, and the Greek-based word “energy” became “work”. As I considered these and many similar facts about the piece, I gradually came to see that it was in all likelihood not a translation from *any* other language, ancient or modern, but simply the result of a strict self-imposed constraint — a decision to write in the language that one might call “pure Anglo-Saxon”, meaning the subset of English that comes solely from Germanic and Celtic roots, with Latin, Greek, and all other roots being totally barred. In sum, this essay was a result not of translating *from* a foreign language into English, but rather from English *into* a foreign language!

Savoring the Strange New Language

Now that I understood what lay behind Anderson's article, I could see it was a very elegant and clever *tour de force*, and I wanted to savor it fully. I decided that the only way to do so would be to convert it lock, stock, and barrel into normal-sounding contemporary English; otherwise I would surely miss some of its subtleties. Being already well on my way, I simply continued to the end. It took a long time, and it turned out that I was lucky to have gotten my doctorate in physics, in the sense that some of the sentences would have been very hard to decode without a solid knowledge of modern physics. Curiously, in fact, there were a couple of tiny conceptual errors in the original text that my knowledge of physics allowed me to recognize, and I patched up those two spots in the original with very slight rewordings. My translation is given as Appendix II.

There were some delightful moments of clarity, such as when I recognized that "forward bernstonish lading" meant "positive electric charge", or "ownship" meant "property". In these cases, it certainly helped to know German words like *Ladung* and *Eigenschaft*. It also helped a lot to know that "electron" comes from the Greek for "amber" and that in German "amber" is *Bernstein*. Similarly, converting Anderson's "offhanging on the makeup of the kernel" into "depending on the composition of the nucleus" was facilitated by my prior studies of both Germanic and Romance languages (for example, the French and German words for "depend" both essentially mean "hang from", and *Kern* is German for "nucleus"). I thought that "for a showdeal" was a rather charming Anglo-Saxon way of saying "for example", although on the other hand "lump beholding" was a strikingly ugly yet well-chosen rendering of "quantum theory".

For me, one of the high points of the whole essay was this surrealistic yet perfectly accurate rendition of Albert Onestone's celebrated equation: "Work is like unto weight manifolded by the fourside of the haste of light." I found the last several words a hilarious way of saying " c^2 " — extremely concrete and literal, yet perfectly sensible when analyzed. It certainly didn't come from German, and I doubted it came from Greek. It was just Anderson's cute way of totally avoiding the Latinate traces present in the word "square".

Perhaps my most satisfying triumph came at the very end, when I managed, after considerable struggle, to figure out that "The Warehouse of Dreamishness and Worldken Sagas" was none other than the magazine I'd enjoyed reading as a kid — namely, "The Magazine of Fantasy and Science Fiction". After the fact, this made perfect sense, but before it, what opacity! I could imagine what delicious fun Anderson himself must have had when composing this article.

Ancient Words for Modern Discoveries of Ancient Entities

From this description of how I came to understand the origin of Anderson's brilliant little "truthwork" (as he might call it), I would now like to pass to a discussion of its *effects* on a reader. Certainly it is nothing if not disorienting and strange-sounding. Obviously, this is due at least in part to the collision between the ancient flavor of the language being used and the modernity of the ideas under discussion — in other words, a seeming contradiction between form and content. But it is much more complex than just this. To explain what I mean, I have to describe a number of subtle mental and linguistic phenomena.

Consider the contemporary English way of talking about submicroscopic phenomena. Physicists use such terms as "molecule", "proton", "neutron", "photon", "gamma ray", "neutrino", "momentum", "angular momentum", and many others. I was exposed to all these and a host of related technical terms at a very early age because my father was an experimental physicist whose research involved probing the structure of atomic nuclei with a linear accelerator of

electrons. I immediately fell under the spell of these terms, especially the names of particles; to me they exuded infinite mystery and romance. As many of my friends know, during recess time in second or third grade, I used to pretend that I was a neutrino (or perhaps an antineutrino), and I got my playmates to pretend that they were photons or protons or other elementary particles, and we all went zipping around the playground — massless ones like myself at the speed of light, the others not quite that fast — and “elastically scattered” off of each other.

I vividly remember that for me, a large part of the romance of these terms came from their alien sounds and structures. Thus, suffixes such as “-cule”, “-ton”, and “-tron”, although they sound very pedestrian to me now, on first exposure sounded very exotic, and made me dream of infinitesimal, invisible, indescribable, forever inaccessible entities. (Of course, I had been told of, and had read about, these things; I wasn’t just guessing!) Small orthographic details like the Greek “ph” rendering of the “f” sound in “photon” seemed eerie; even little vowel pairs like the “eu” in “neutron” seemed alien and curious. The charm or power of these terms, in short, came in good measure from their *opacity* to my child’s mind.

Very similar feelings were evoked in me at almost exactly the same time but by an entirely disjoint set of words — the names of dinosaurs and other creatures, as well as the names of the geological epochs in which they lived. Thus words like “brontosaurus”, “triceratops”, “plesiosaurus”, “eohippus”, “tyrannosaurus”, “pterodactyl”, “stegosaurus”, “trilobite”, “paleozoic”, “pleistocene”, and even the innocent-seeming “epoch” — all these were both magically evocative and exceedingly frightening to me at that time (and of course this was some forty years before the dinosaurs had their second heyday, thanks to Crichton, Spielberg, & Co.). The many strange new sequences of sounds and combinations of letters — the “pt” of “pterodactyl”, the “eo” inside “eohippus” and “paleozoic”, even the hard “ch” of “epoch” — seemed to be in and of themselves microscopic versions of the mysteries that these words denoted. In my mind, the exotic names were *of a piece* with the mysteries for which they stood.

Now of course this is not particularly surprising. People always tend to think that their own native language’s name for some object or idea is completely natural and “right”, and that those in other languages are not nearly as suitable. What is peculiar, however, is that I was reveling not in the transparency of these names, but in their opacity, their strangeness, their nondecomposability. The last term perhaps hits it most on the head. One had to swallow these long words, whether in physics or in paleontology, as wholes — one could not hear clues inside them suggesting what they were all about. At least a child certainly couldn’t.

Strothmann’s Thesis about the Semantic Transparency of German

Some years later, as a young adolescent, I began to study foreign languages, and became intoxicated by them. First it was French, which became my strongest language thanks to a year spent in Geneva, and then in my mid-teens I took up German, Italian, and Spanish. Given my familiarity with French, I found the other romance languages very beautiful but not extremely alien. German, on the other hand, seemed very peculiar. I remember our professor, a German native and a famous authority on his language, proudly repeating over and over again to our class that no German ever needed a dictionary — *ein Wörterbuch* — because every single compound word could be understood at once by breaking it into its components (e.g., *Wörter* and *Buch*). German was so *logical!* (It occurred to me to wonder why, if no German dictionaries existed, the term *Wörterbuch* even existed, but of course the answer was obvious — other languages, being less logical than German, *do* need dictionaries, whence a term is needed for those books.) Many

years later, when I lived in Germany for several months, I again heard this claim about the transparency of the German language, which I will here call “Strothmann’s thesis” in my professor’s honor.

Of course, Strothmann’s thesis tacitly assumes that one knows all possible components first, a list of which would certainly number in the several thousands, and would include many exceedingly rare items. But even given that rather dubious assumption, Strothmann’s thesis turns out to be completely untenable, as I had suspected from the very outset. How could anyone ever intuit, for instance, that whereas *Satz* means “sentence” or “set”, various compounds built from it, such as *Aufsatz* (“on-set”), *Absatz* (“off-set”), *Einsatz* (“in-set”), and *Aussatz* (“out-set”) mean “homework”, “heel”, “bet”, and “leprosy”, respectively? What would ever suggest to anyone not already in the know that *aufmachen* (“on-make”) means “to open”, *auffallen* (“on-fall”) means “to be noticeable”, and *aufhören* (“on-hear”) means “to stop”? Or that *Beispiel* (“by-game”) would mean “example”, that *Unterhaltung* (“under-holding”) would mean “conversation” as well as “entertainment”, that *zunächst* (“to-next”) would mean “first of all”, that *sowieso* (“so-as-so”) would mean “in any case”, that *durcheinanderbringen* (“through-one-another-to-bring”) would mean “to mess up”, that *soweit* (“so-far”) would mean “ready”, that *damit* (“there-with”) would mean “so that”, that *hinrichten* (“thither-judge”) would mean “to execute [someone]”, that *beliebig* (“belove-ish”) would mean “random/arbitrary”, and on and on and on?

To be sure, one could make very similar points about the opacity of English. It’s quite hard to understand, for example, why a compound word made out of “under” and “stand” should mean “understand”. What sense does that make? Or consider the verb “to make out”. Why in the world should it have such a motley collection of meanings as “to discern”, “to fare”, “to fill out”, and “to pet heavily”? No one in their right mind could ever be expected to make out these various meanings from “make” and “out” alone. And then there is “to make up”, which can mean, among other things, “to invent”, “to become reconciled”, and “to apply facial cream to”. No one in their right mind would have made up such a diverse set of meanings for a single term. And so on and so forth. I would be the first to admit that just as many examples of the opacity of compounds could be adduced for English as for German. The difference, however, is that we English speakers don’t make grand claims about the supposed transparency and logicality of our language, whereas Germans — at least some — certainly do.

When I was in Germany, running again into Strothmann’s thesis motivated me to do something I had wanted to do years earlier — I compiled a longish list of to-me-completely-opaque German compounds, such as those shown above, and I showed it to various people who made the claim. To some extent, seeing it convinced them that German is not nearly so *benutzerfreundlich* as it is often made out to be.

Ironically, however, through my efforts to refute Strothmann’s thesis, I slowly came to recognize that there was actually some truth to the thesis, in the sense that German words very often *are* indeed built up in a way that tends to be more transparent to a native speaker than English words or French words are. Thus, that *Durchschnittsgeschwindigkeit* means “average speed” is pretty obvious as long as you know that *Durchschnitt* means “average” and that *Geschwindigkeit* means “speed”. Another example would be *Wortbedeutungslehre* (“word-meaning-theory”), which quite straightforwardly means “semantics”. In all likelihood, my old professor’s claim was based in large part on longish words like those, or on the even longer ones for which German is notorious, such as the term

Agrarstrukturverbesserungsmaßnahmen

(“measures taken to improve the agricultural structure”), which I once ran across

in a newspaper. To an adult German speaker, this blockbuster of a word falls apart very simply and effortlessly into *Agrar/Struktur/Verbesserungs/Maßnahmen*, and its meaning is obvious — no dictionary needed.

The Vagueness of the Notion of “Word”

Of course, Germans are by no means the only ones to make such giant words — monsterwordbuilding is a vice they share with many other languages, including Dutch, Swedish, and others. In fact, I cannot resist inserting here the longest word I have ever run across in an everyday context (as opposed to the often-forced examples in the Guinness Book of Records), one that I saw on a sign hanging on the wall above a wastepaper basket in a university building in Lund, Sweden:

pappersåtervinningsstunnetömningsschema

which breaks up as *pappers/återvinnings/tunne/tömnings/schema*. Taken piece by piece, this means “paper-recycling-barrel emptying schedule”, or perhaps more gently, “emptying schedule for the paper-recycling barrel”. (Incidentally, *återvinnig* itself breaks up further into *åter* and *vinnig* — “backgaining”, more or less.) When I gawked at this “word”, people seemed surprised, since to them it was so obvious what it meant. No one seemed to find it funny, as I did.

Despite appearances, the transparency of such monstrous constructions has very little bearing on Strothmann’s thesis, because such words are really just long phrases run together into a long unit — phrases that would make perfect sense in *any* language. That certain languages encourage blankspaceless wordfusion and that others discourage it is pretty much an arbitrary convention developed over centuries, as opposed to a radical difference between languages.

Take, for instance, Italian. Few people think of Italian as a language filled with long compound words, and yet the single-word utterance *Diamogliene!* requires several words to be rendered in English: “Let us give some of it to her”, or, more tersely, “Let’s give her some”. We anglophones *could* have run them all together as “letsgivehersome”, and the Italians *could* have separated the components as *Da iamo le ne* (“give/we/her/some”), but it just happens that over the course of centuries, the Italians decided to run certain verbs and certain pronouns in certain contexts together into single words, a little bit like our “wouldn’t’ve”, which some people might count as a single word (and which many adult Americans shamefully misspell as “wouldn’t of”, revealing that they don’t even understand how it is put together on a conceptual level). On the other hand, when the Italians wish to say, “We are giving some of it to her”, they simply turn the parts around a little bit, and insert one blank space: *Gliene diamo*. And to say, “We are giving some of it to you”, they insert *two* blank spaces: *Te ne diamo*. And to say, “We gave some of it to you”, they use four words: *Te ne abbiamo dato*. As I said, it is very arbitrary — almost capricious. Just what *is* a word, anyway? We think we know, but do we?

The Chinese¹ are far vaguer than we in the West are about what words are. The reason for this vagueness is that there is simply no counterpart to the blank space in written Chinese; characters are simply juxtaposed, left-to-right, right-to-left, or top-to-bottom, without any kind of conceptual separator. People make divisions in their minds wherever they imagine them to make the most sense. Curiously, there are two words often used in Mandarin for “word” — *zì* and *cí* —

¹ My renditions of Chinese in this article utilize the now-prevalent *pinyin* romanization, but due to limitations of the word-processing program I have, the symbols denoting two of the four tones in Mandarin are somewhat nonstandard. In particular, to denote the first tone, I use an umlaut instead of a short horizontal bar, and to denote the third tone, I use a right-side-up circumflex instead of the usual upside-down one. Second and fourth tones are as usual.

and many Chinese don't even know the difference in meaning between them. The former, strictly speaking, means "single character", whereas the latter corresponds more to what we Westerners mean by "word" — a conceptual unit made out of however many characters are necessary to render it, which most often means two or three rather than one. For instance, both *zìdiān* and *cídiān* — literally "character canon" and "word canon" — are used in Mandarin to render the English word "dictionary". Each of these terms is itself an example of the notion of a *cí* rather than a *zì* — in fact, each term is a two-*zì* *cí*. However, it gets blurrier when one asks a Chinese how many *cí* are in the three-character string *cídiānlì*, which means "in (a/the) dictionary". Some people will say "two", meaning that *cídiān* is one *cí* and *lì* is another, whereas others will say that *cídiānlì* is one word, somewhat like the locative case of a noun in a language like Russian.

And it gets far blurrier when one looks at more complicated cases. Take the phrase *méiyōuyìside shū*, meaning "uninteresting book". I have separated it into two parts, one corresponding to each English word. But a Mandarin speaker might well perceive *méiyōuyìside* not as one word but as two words (first *méiyōuyì*, roughly meaning "not-have-interest", and then *de*, an immensely frequent modifier-yielding suffix meaning something like "-ish"), or perhaps as three words (*méiyōu yìsi de* — "not-have/interest/-ish") — or even as four words (*méi yōu yìsi de* — "not/have/interest/-ish"). There is just no "right answer" to the question "How many words are involved?", because "word" is a western concept, a notion artificially imported into discussions of Chinese.

Despite the imperfect match between the structure of Chinese and the notion of "word", linguists studying Chinese have struggled hard to try to find a precise Chinese counterpart to the concept of "word", but unfortunately there is no clear consensus, despite many theoreticians' adamant assertions one way or the other.

Is Strothmann's Thesis a Universal Fact about All Languages?

Germans (and the Dutch and the Swedes and so on) are willing to run shorter words together to form longer ones in contexts where we English speakers would most often use spaces or hyphens as separators (or as joiners, if you prefer looking at it that way). What fundamental difference is there, for example, between writing "second hand book store", "second-hand book-store", "secondhand bookstore", and "secondhandbookstore"? There isn't really any. It's just a rather arbitrary convention that in English we usually opt for the third version.

Even if "secondhandbookstore" were a single English word, it would hardly be a marvelous and special fact about English that this long word was instantly comprehensible to a native speaker. Well-constructed blankspaceless wordchains should *always* be clear, whatever language they are in. Thus Spanish has the three-piece chain *limpiaparabrisas* (*limpia-para-brisas*) meaning, on a roughly literal level, "washerblockerbreezes", which, to a native Spanish speaker, is very clear — just as limpid as the term "breeze-blocker washer" would be to a native English speaker — in other words, "windshield wiper". Similarly, Italian has the chain *reggiasciugamani* (*reggi-asciuga-mani*), which literally is "holder-dryer-hands" or, turned around, "hands-dryer holder" — or in more common parlance, "towelrack".

If Strothmann's thesis is merely the claim that *extremely long* compound words in German make sense without needing to be explained, then it's hardly a surprise. The analogous claim holds just as much for romance languages. And why shouldn't it? Romance-language speakers, no less than Germans and Swedes, string together long chains of words for the purpose of making sense; so do all human beings. The fact that such chains can be quite effortlessly decoded by native speakers is hardly a revelation.

Some Very Long Compounds

In a sense, we in English are just as guilty as the Germans, the Dutch, and others are of putting together long and fairly hard-to-digest chains of words; the only difference is that ours are a little more disguised because we keep the blanks whereas they swallow them. Consider these examples, all of which are genuine in the sense that I ran across each of them in an official printed document of one sort or another:

Palm Beach invitational piano competition winners concert series
General Motors on-site supplier assessment team summary
aluminum soup pot cover clearance sale
product liability insurance premium problem
power supply cost recovery reconciliation refund
space station payload contamination compatibility workshop

Despite their blank spaces, some of these make even the Swedish and German behemoths cited above look quite tame, by comparison. (And note that in the last one, both "payload" and "workshop" are compound words on their own.)

In theory, though regrettably seldom in practice, compound modifiers in English are supposed to be hyphenated in order to clarify what is modifying what. Thus one is supposed to write "high-energy physics" rather than "high energy physics", or "soup-pot cover" rather than "soup pot cover". This useful convention can sometimes be crucial in resolving ambiguities, such as that of the phrase "the best known cure for cancer", which in this unhyphenated form would tend to suggest "the best of all currently known cures for cancer" whereas with a hyphen between "best" and "known", it would definitely mean "the most famous cure for cancer". Another example is the sentence "The single armed guard was quickly overpowered by the angry mob", which, as is, refers to a solitary soldier carrying a weapon, whereas with a hyphen ("single-armed") it would conjure up an image of a soldier with a missing upper-torso limb.

Despite the clarity that hyphenated compound modifiers lend, hyphens are most often completely neglected, a fact that can lead to very opaque writing. For instance, consider this description of someone's physics research, quoted exactly as I found it in a magazine:

He has done fundamental work on one loop flavor changing processes.

This is very hard just to parse, let alone to understand, until you put in a couple of hyphens:

He has done fundamental work on one-loop flavor-changing processes.

Of course I'm not claiming that thanks to the two hyphens, any lay person could now easily understand the sentence, but at least a radio announcer, say, could parse it and thus read it aloud sensibly.

Hyphens, Dashes, Levels, and Hierarchies

The careful use of hyphens in long compound modifiers can help considerably. Thus some of the above examples, rewritten with hyphens, become much easier to parse, even if not to fully understand:

General Motors on-site-supplier assessment-team summary
aluminum soup-pot-cover clearance-sale
power-supply cost-recovery reconciliation refund
space-station payload-contamination compatibility-workshop

Actually, these hyphens, though helpful, don't really do justice to the complexity of the structures. After all, "space-station payload-contamination" is itself a compound modifier for what follows, and thus should also be hyphenated in some manner. We could imagine a compromise between German and English word-building conventions that did it this way:

spacestation-payloadcontamination compatibilityworkshop

First, the three most tightly-bound compounds are run together to form three "pure" (*i.e.*, spaceless) words. Next, the weaker bond between the first two "pure" words is rendered by a hyphen, thus creating a compound modifier for the third one. Finally, the weakest bond of all is represented by a blank space. The hierarchical structure of the phrase is thus captured explicitly by these conventions. (I have to admit, I don't really know if I've parsed the phrase right; I have no idea what a "compatibility workshop" would be; on the other hand, I can't see any more plausible way to assimilate the word "compatibility".)

Unfortunately, even a powerful hierarchy-capturing symbolism such as the one suggested above would be very hard put to represent the full meaning of the phrase about the piano concerts. Below I have tried as well as I could to capture the strength of the various bonds by exploiting not just word fusion, hyphens, and blanks, but also by allowing different hyphen-like symbols having different strengths. Specifically, I have used the longer en-dash or minus-sign “–” to represent a weaker link than the shorter hyphen “-”, and the even longer em-dash “—” to represent a yet weaker link, although it is still stronger than a blank space. (When reading out loud, you can think of the length of the line as representing the amount of time to pause, with the blank space counting as the longest line, hence the longest pause.)

Palmbeach–invitational-pianocompetition—winners concertseries

An alternative way of representing the exact same hierarchical structure is through multiple levels of parenthesization:

((Palm Beach) [invitational (piano competition)]) winners) (concert series)

Here, the more deeply-nested parentheses and brackets represent the more tightly-bonded substructures, and the outermost (and largest-size) parentheses represent the most weakly-bonded ones.

But no matter how you notate it, the problem is that some words should be conceptually bonded to neighbor *words* ("piano" and "competition", for instance, which in the first display were run together to form a single word), whereas other words should be conceptually bonded to neighbor *structures* (*i.e.*, to small or large compounds considered as units). To be more concrete, to what other unit would you say the word "winners" is truly bonded, on a conceptual level? To the simple word "competition" immediately to its left? Yes, clearly — the winners are obviously winners of a competition. But that's not the whole story. How about to the run-together word "pianocompetition"? Yes, "winners" is also conceptually bonded to that, to a good degree. But this does not undermine or deny the existence of the previous bond. Well, then, what about to "invitational-pianocompetition"? Yes, indeed, "winners" is bonded to *that* structure as well! So now we have three distinct bonds, all coexisting, linking "winners" to various structures at different levels. And how about to "Palmbeach–invitational-pianocompetition"? Of course!

But this means that "winners" is semantically or logically bonded to structures at four different hierarchical levels (at least). Thus we see that what is modifying

“winners” is extremely vague and blurry. It’s not that it is in reality bonded to just one and not to the others; “winners” is genuinely linked, in the mind of any reader, to *each* of those phrases, although I emphasize that this effect is certainly unconscious and nonverbalizable. It would be impossible to capture in a simple and clear notation the “superposition” of adjectival effects just described; in fact, it is such a blurry blend of effects that it would make no sense to try to formalize it or make it precise. In each different reader’s mind, the effect must be subtly different.

Moreover, it is not clear exactly what is modifying the term “concertseries”. In some sense it is merely a *winners*’ concert series, which would suggest something like this way of looking at the entire phrase:

Palmbeach-invitational-pianocompetition winners-concertseries

But if it is structured this way, then there is no direct bond at all between “competition” and “winners”, which seems quite illogical. Somehow, it would seem, “winners” has to have bonds that go both leftwards *and* rightwards. There is no obvious way to accommodate this within a parenthesis-based notation, because such a notation would force one to choose one or the other direction. That is, either the parenthesized structure shown above, or this:

{(Palm Beach) [invitational (piano competition)]} [winners (concert series)]

You can’t get “winners” to group with both its left and right neighbors simultaneously without doing violence to the whole notion of parenthesization.

Long-distance Bonds inside Word Chains

There is still more to the story of this phrase. In addition to the existence of various types of bonds between *adjacent* words or substructures in such a compound structure, there is another effect entirely, in which arbitrarily *distant* components resonate with each other. In this particular phrase, for instance, one of the very strongest of all the inter-word bonds would surely have to be that between “piano” and “concert”, a long-distance bond that saucily leaps right over two intervening words. On this coarsest of levels, a reader is simply noticing the fact of *piano concerts*. This effect could be roughly captured by putting the conceptually linked words in larger, heavier print:

Palmbeach-invitational-*piano*competition—winners *concert*series

And for those people who like classical music but who are not that concerned over details, there is the quite different long-distance bond between “Palmbeach” and “concertseries”. Such a person, on arriving home, would be likely to say to their spouse, “Hey honey, today I saw a poster for a concert series next month in Palm Beach — want to go?”

Palmbeach*-invitational-pianocompetition—winners *concertseries

In either of these cases, two far-apart units are tightly bonded despite their spatial separation, a bit reminiscent of the way that distant chemical components inside a long chain (*e.g.*, a protein or a strand of RNA, consisting of many subunits strung together sequentially by uniform types of bonds) can bond together if the chain twists around in a complex three-dimensional manner (as proteins and RNA indeed do) in such a way as to bring the linearly-far-apart components into physical proximity. More concretely, you can imagine a long noodle twisted around so that regions near its two ends are stuck to each other.

Frame Blends — Unconscious Mental Superpositions

The truth of the matter is that the mental effect of this phrase on a reader is a very big blur — and yet the phrase is one that the human mind takes easily in stride and feels quite comfortable with, never pausing to consider the enormous complexity of what is going on. Attempting to use a highly structured formalistic notation with many levels of hyphens and dashes (or of parentheses and brackets), as well as various levels of boldface, and so on, to “clarify” the structure of the phrase actually does nothing but becloud it.

The superposition or blur of meanings that I have just described provides an example of what I call a *frame blend*. The term “frame” in this context is borrowed from a number of sources, but owes probably most to the American sociologist Erving Goffman, who used it — at least so I perceive it — to mean essentially *a situation as perceived by someone*. In other words, a frame is a personal mental construct that attempts to represent a situation, whether simple or complex, whether specific or generic. And frame *blending* — a notion that I don’t think Goffman himself explicitly considered, though it is very much in his spirit — is thus the phenomenon wherein two or more frames are constructed and merged in someone’s mind, usually without the person’s being in the slightest aware of any such blurring.

One type of frame blend occurs when there is a *single* ambiguous situation and two or more different interpretations of it contribute simultaneously to someone’s ruminations about it, not necessarily in equal amounts. Another type of frame blend occurs when there are two different but *similar* situations, and fragments of interpretations of each are spliced together to form a kind of hybrid representation, neither fish nor fowl, but partly each. In the discussion above, we were dealing with a blend of the first kind, where a single long phrase has different ways of being structured in the mind and these various possibilities all simultaneously contribute, to varying degrees, to what one naïvely takes to be its univocal, clear meaning. We will encounter other examples of the frame-blend phenomenon later on.

Semantic Transparency, Translucency, and Opacity: A Spectrum

Let us now get back to Strothmann’s thesis, which we left in the lurch quite a while ago. The real question that Strothmann’s thesis seems to me to raise is not that of monsterwordchaintransparency; rather, it concerns the transparency of *standard* compound words — words of sufficient frequency and generality that one would be shocked not to find them in a “wordbook”. Typical examples in German would be these: *Umwelt* (“surround-world”, hence “environment”), *Handschuh* (“hand-shoe”, hence “glove”), *Brustwarze* (“breast-wart”, hence “nipple”), *Säugetier* (“suckle-beast”, hence “mammal”), and *Wiedergutmachung* (“again-good-making”, hence “[war] reparations”), most of which are at least reasonably close to transparent, given the meanings of their constituents.

On the other hand, German also has many standard compound words that exemplify what might be called “semantic semitransparency” or perhaps better, “semantic translucency”. Two examples mentioned earlier are the two-part compounds *Wörterbuch* (which might easily have meant “thesaurus” or “book with text but no pictures”, but which instead happens to mean “dictionary”) and *Durchschnitt* (“through-cut”, which might well have meant “cross-section” or “amputation”, but in fact means “average”). Others include *Kindergarten* (“children-garden”, which might have meant “playground” or “children’s vegetable garden”, but instead means “kindergarten”), *Lautsprecher* (“loud-speaker”, which might have meant “megaphone” or “person with a booming voice”, but instead means “loudspeaker”), *Fingerspitzengefühl* (“finger-tip-feeling”,

which might have meant “the sensation of having itchy fingers” or “the feeling of being tickled”, but instead means “sensitivity”), and *Seitensprung* (“side-jump”, which might have meant “digression” or “the act of dodging a thrown object”, but instead means “extramarital affair”). I could go on, but this will suffice to get the idea across. Basically, semantic translucency, as illustrated by the preceding list, is the phenomenon whereby a term seems to make eminent sense after one knows it, but is by no means trivial to guess before the fact.

Semantic Transparency: A Highly Subjective Affair

The question that arises in my mind about the alleged transparency of such compounds is this: *transparency to whom?* To little children, to teen-agers, to adults, to professional etymologists, or to foreigner learners? How transparent is a term like *Sprechstundenhilfe* (literally “speech-hours-help”, but meaning in fact “[medical] receptionist”) to a German child, even one who knows all three components (or two, depending on whether *Sprechstunden* is itself seen as a unit)? How transparent is a term like *Meistbegünstigungsklausel* (literally “most-promotion-clause”, but meaning in fact “most-favored-nation clause”) to a bright contemporary German teen-ager, or even to a random German adult? Suppose a highly intelligent German adult from the 1800’s were transported into our century; how transparent would that person find terms like *Wechselsprechanlage* (“exchange-speech-installation” — in short, “intercom”), *Rundfunkansager* (“round-spark-on-sayer”, meaning “radio announcer”), *Fernsehbildröhre* (“far-see-picture-pipe”, meaning “TV tube”), *Vervielfältigungsapparat* (“manifolding-apparatus”, meaning “duplicating machine”), *Wünschelrutengänger* (“desire-rod-walker”, thus “dowser”), and so on?

Obviously to make sense of such terms, one needs to be familiar with the culture that uses them; they are not self-explanatory in the way that, say, *Wettervorhersage* (“weather foretelling”, literally) or *Wellenreiterbrett* (“waverider-board”, thus either “surfboard” or “water ski”) quite clearly are.

To make the discussion as simple as possible, let us take the word *Rundfunk*, whose two components mean “round” and “spark”, respectively, and which as a whole means “radio”. Since it denotes an extremely common item in this day and age, children learn it at a very early age and probably never even hear it as a compound word at all.

If this claim sounds implausible to you — if it strikes you that no one could fail to hear such very obvious parts — then consider the following English words that most children know by the age of ten or twelve: “maybe”, “therefore”, “careful”, “become”, “because”, “cupboard”, “doughnut”, “wardrobe”, “briefcase”, “handsome”, “offhand”, “without”, “withstand”, “understand”, “awful”, “cockpit”, “cocktail”, “however”, “always”, “nobody”, “anyway”, “band-aid”, “kickstand”, “sidekick”, “haywire”, “bulldozer”, “ballpoint”, “touchdown”, “upset”, “pocketbook”, “broadcast”, “skyscraper”, “dovetail”, “today”, “highway”, “highjack”, “passport”, “upshot”, “free-for-all”, “welcome”, “breakfast”, “toadstool”, “mushroom”... I could go on and on. Once your attention is explicitly drawn to the fact that these usually-unanalyzed words are in fact made out of pieces, then in most cases you can see a *bit* of logic behind them, and in some cases the logic even comes to seem inevitable, *a posteriori*. But to children, these words are just swallowed whole; it never crosses their mind for a split second that “band-aid” has something to do with aiding anything or that the “al” in “always” is the same as the “all” in “free-for-all”.

Even to adults, it may come as a shock that such a word is understandable in terms of its parts. A couple of years ago, my wife, a sophisticated user of English, exclaimed with delight when she learned that the Italians say *grattacielo* for

"skyscraper" — it seemed to her so graphic and concrete. When I pointed out that our word and their word are exactly isomorphic, she was floored. It had just never occurred to her.

It may sound backwards, but sometimes the parts are learned considerably after the compound is learned. I would guess that *Rundfunk* is that way for many German youngsters. A typical example in English is "airport", which in this day and age is learned by many toddlers of just one or two years of age, long before they have the foggiest idea of what "air" or "port" means. My daughter, who is almost three now, still says "hairport" — perhaps she thinks airports have something to do with hair, but more likely is that she doesn't have any theories at all about her word "hairport". She just accepts it for what it is, just as she accepts the word "pacifier", pronouncing it perfectly without trying to analyze it.

The Special Yet Typical Case of the Word "Pacifier"

This word "pacifier" is in fact one of the most charming and striking examples of the phenomenon of children being oblivious to parts that are clearly there to sophisticates. For those unfamiliar with it, "pacifier" is the term used all over America to denote the soft plastic "pseudo-nipples" that parents stuff in the mouths of agitated babies to calm them down — in other words, to *pacify* them. (In Britain, pacifiers are standardly called "dummies".) "Pacifier" is one of the words learned earliest by American babies, and to many people it sounds perfectly ordinary to hear a two-year-old crying "Want pacifier! Want pacifier!" To my ear, by contrast, such utterances have always sounded extremely weird, because the child is unwittingly referring to this object of desire by specifying its function of pacification. It's almost as if the child were saying, "Give me the object that calms me down when I'm agitated!" — which, in a certain sense, the child actually *is* saying — but of course to the child, "pacifier" is just a semantic atom admitting of no dissection into semantic parts. As far as the baby is concerned, "pacifier" has no more inner structure and logic than does its far less frequent synonym "binkie", or their Italian counterpart, *ciuccio*.

In Dutch, incidentally, the term for "pacifier" is *fopspeen*, whose first component, *fop*, means "hoax" or "fool", and whose second component, *speen*, means "nipple" or "teat". Thus a Dutch baby who cries for a *fopspeen* — a "hoaxteat", shall we say? — is in a sense asking for something even weirder than its American counterpart, because on one level, it is literally asking to be *tricked!* But of course a Dutch baby has no idea of this irony.

There is a further irony to the word "pacifier" — namely, by the time one is old enough to understand its semantic pieces and how they are put together and the ironic flavor of the overall term, one has long since ceased to have any understanding of how wonderful it *feels* to have such an item in one's mouth — rather, one is repelled by the ideal! In other words, an adult understands very well the *intellectual fact* that such items pacify babies, but only a baby *knows how* such an item really works. Thus babies, who haven't the foggiest idea about the word's innards, have a very deep first-person appreciation of the *pacification* itself, whereas grown-ups who fully understand the *word* have merely a third-person understanding of what is going on. (Similarly, it is only small children who *really* know how thrilling the word "Christmas" is; to us adults who know all about Jesus' birth and religious holidays and compound words and can understand slick slogans like "Put Christ back into Christmas", the delirious joy identified with the word "Christmas" in a child's mind is but a dim memory, contaminated by the overlay of many seasons of weary, dreary Christmas shopping.)

The Widespread Illusion of the Semantic Transparency of German

But back to *Rundfunk*. If a group of intelligent foreign learners of German were shown this term and asked to guess what it means, they would have a very hard time of it. What could “round-spark” mean? An ingenious guess would be “ball lightning”; another cute conjecture might be “electron” (don’t you visualize an electron as a tiny sparky sphere?). Even if the hint is given that the flavor of *rund* here is “scattered around” rather than “circular”, it is very unlikely that the idea of radio or radio waves will come to mind. Of course, after the fact, when one thinks of radio waves emanating outwards from an antenna tip, much as concentric ripples emanate from a stone’s splash in a pond, the term absorbs those connotations and it seems, retroactively, to make sense.

Now there are thousands upon thousands of similarly opaque compounds that German-speaking children absorb without any problem in their first several years, many of them before even learning to read or write. I am talking about words like *vielleicht* (literally “much-easy”, but meaning “maybe”), *einfach* (literally “one-fold”, but meaning “simple”), *nachholen* (“after-fetch”, but meaning “make up for”), *Zimmermann* (“room-man”, but meaning “carpenter”), *Kronleuchter* (“crown-lighter”, but meaning “chandelier”), *Auseinandersetzung* (“out-of-one-another-putting”, but meaning “argument”), *Aufzug* (“up-pull”, but meaning “elevator”), *Mannschaft* (“man-ness”, but meaning “team”), *Bahnhof* (“road-court”, but meaning “railway station”), *preisgeben* (“price-give”, but meaning “surrender”), *vollkommen* (“full-come”, but meaning “perfect”), *ausgezeichnet* (“out-drawn”, but meaning “excellent”), *vorschlagen* (“fore-hit”, but meaning “suggest”), *Bergbau* (“mountain-construction”, but meaning “mining”), *Morgenland* (“morning-land”, but meaning “Orient”), *aussehen* (“out-see”, but meaning “appear”), *Schildbürger* (“sign-citizen” or “shield-citizen”, but meaning “dunce”), *Bürgersteig* (“citizen-path”, but meaning “sidewalk”), *Rauschgold* (“drunk-gold”, but meaning “tinsel”), *Eiweiß* (“egg-white”, but meaning “protein”), *Scharfrichter* (“sharp-judge”, but meaning “executioner”), *wahrscheinlich* (“true-seemly”, but meaning “probable”), *Kreisbahn* (“circle-way”, but meaning “orbit”), *umkommen* (“around-come”, but meaning “perish”), *sich ausdrücken* (“out-press oneself”, but meaning “express oneself”), *Wechselwirkung* (“change-effect”, but meaning “interaction”), and on and on. I don’t want to bore you, but I want to convince you that a ten-year-old’s vocabulary includes literally *thousands* of such semi-opaque to fully-opaque compounds.

Keep in mind that I am not claiming that English is any different or any better. Who could have predicted that “airport” would be the word for what might just as easily have been dubbed a “plane station” or a “plane field”, or that “bookworm” would mean “intellectual”, or that “warhead” would mean a bomb aboard a rocket, or that “crackpot” would denote a totally deluded thinker, or that “copycat” would designate a tiny kid who admiringly apes a slightly older kid, or that “underdog” would mean, well, “underdog”? In fact, English-speaking children probably have to absorb many *more* terms that are opaque than do German-speaking children, since English is so pervaded by romance roots, which most native speakers do not understand in isolation.

The strange thing is that many adult German speakers, even after seeing examples like the ones above, will still insist, if asked if German words are by and large transparent, that they *are* indeed so. This illusion — and I claim it truly is an illusion — is caused by their decades of exposure to terms like *Rundfunk*, *Bergbau*, *preisgeben*, and *vorschlagen*, which they retroactively hear as perfectly sensible compounds, even though when they first learned them, they probably did not think of them as compounds at all. Thus a German is likely to claim that *umkommen* is just as sensible and as lucid, even when one first encounters it, as is

Staubsauger (“dust-sucker”, thus “vacuum cleaner”). This, however, is simply untenable. Any outsider who knows the meanings of its two components can trivially guess what *Staubsauger* means, whereas *umkommen*, unless it is embedded in a context that gives it away, is virtually foreigner-opaque — no less opaque than the American idioms “kick the bucket” and “buy the farm”, also meaning “to die”, or, for that matter, the phrase “to come around”, which means “to finally get convinced”.

“Words in My Native Language Mean What They Sound Like!”

The word *umkommen* is an informative example. To adult German speakers, the prefix *um-* long ago acquired an unconscious halo of connotations of dying, thanks precisely to this word, and so, by a kind of backwards logic, many German-speaking adults come to think that the full word makes sense because of the meanings of its pieces.

Many speakers of English are similarly convinced that English words often “sound just like” the phenomenon they represent — perhaps “chicken”, perhaps “breast”, perhaps “slippery” or “dungeon”, and so forth. (What could *possibly* sound danker, colder, more medieval and forbidding than the word “dungeon”?) Of course, it is with particular vehemence that many English speakers will insist that Anglo-Saxon vulgarities such as “shit” and “fuck” have intrinsically repugnant sounds hinting strongly at their meanings. (Why, then, are “shuck” and “fit” perfectly innocuous English words? Do they, too, sound just like what they mean?) It is almost impossible to convince someone who hears a word this way that the connection is essentially arbitrary, and that the phonetic-conceptual similarity they claim to hear is illusory.

In any case, German speakers, when first asked their opinion about Strothmann’s thesis, almost always take for granted a huge base of words that foreign learners do not have. They do not remember how or when they learned all those words, and they claim not to have consulted dictionaries (except for checking on spelling, which, ironically, is almost perfectly regular).

I might add that I myself have little memory of using a dictionary to learn English; most of the time I must have guessed at meanings from context, and sometimes I of course asked help from people who knew more than I did. But the fact that I seldom consulted a dictionary didn’t in any way instill in me the illusion that all English words have self-evident meanings.

Wordstews and Semi-opaque Proper Nouns

One German word I find of considerable interest is *Volkswagen*. This Hitler-era term is of course transparent to native speakers the moment they hear it: it falls instantly apart into *Volks* and *Wagen*, thus clearly meaning “people’s car” or, if you prefer, “folks’ wagon”. On one level, this is how Germans hear the term. However, on another level, it certainly is not heard that way — at least not if my own personal experience with proper nouns that are compound nouns has any generality, and I would guess it does.

Two of my favorite English-language examples of this phenomenon have to do with buildings in New York City. One is the term “Empire State Building”, which I have known virtually my entire life. To understand this example, one has to know that when native speakers pronounce this term, they put a slight stress on the word “State”. This holds as well when the shorter name for the building, namely “Empire State”, is used. As a child, I heard this term over and over again, and of course said it over and over again. I have to admit, the Empire State was without any doubt my favorite building in the world. I knew both constituents in

its name but it never occurred to me to wonder why they were combined in this phrase, or what it meant. Rather, I took the phrase as an unanalyzed chunk.

Quite independently, as a child fascinated by geography, I had learned that the state of New York, just like every other state, has a nickname, and that that nickname is none other than "Empire State". (Here I insert a caveat for non-native speakers: the pronunciation of this term is not at all the same as the name of the building. Here, the accent falls distinctly on the word "Empire" — in fact, on its first syllable — while "State" is said rather softly. When people mean the building, they always say "*Empire State*", and when they mean the state, they always say "*Em pire State*".) Although I knew this nickname, I never had connected it in my mind with the building. And then one day, *bang!* — the two facts collided in my mind and all of a sudden I had an epiphany: the name of the tall building in New York City actually came from the nickname for the state it was located in! I had never put two and two together before — my favorite building's name had seemed completely chunk-like, completely indivisible to me. Or rather, its parts had been visible to me but I had never tried to figure out how they added up to the whole.

A similar term is "Madison Square Garden", the name of the place where countless famous boxing matches have taken place. The accent in this expression falls without any doubt on the first syllable of "Garden", with a slightly weaker accent falling on the first syllable of "Madison": "*Madison Square Garden*". I heard this said a thousand times when I was growing up. And being a great lover of New York City, I also heard the term "Madison Square" (accented on the second word) quite often, and knew it denoted a place in Manhattan. But it honestly never occurred to me that *Madison Square Garden* was located in *Madison Square*. Rather, I thought *Madison Square Garden* was some kind of garden that was square in shape! Why not? After all, boxing rings are squarish in shape. (On the other hand, it didn't enter my mind to ask why a boxing ring would be called a "garden". Children accept so many mysteries!)

It came to me one day as a complete revelation that the building was named for the place it was located in — it was *Madisonsquare Garden*, not *Madison Squaregarden*. The irony was that, by the time I had that flash of understanding, the original *Madison Square Garden* had been knocked down and a new one, still bearing exactly the same name, had been built in a different location, some ten blocks away — definitely not in *Madison Square*. (And, by the way, I still have no idea why this huge arena was and still is called a "garden". Adults accept so many mysteries!)

I hope these two examples suggest how the term *Volkswagen* might be heard by a native German speaker: namely, as an obviously compound word yet a *semi-opaque* compound word. Perhaps the first few times it is heard, its pieces stand out and the resemblance to an indisputably straightforward noun phrase like "water pistol" or "restaurant supplier" is strong; but after it is repeated many times, the parts start to mingle and merge into each other, something like the ingredients of a stew as they cook longer and longer together, and the end result of this identity-blurring process is a very different kind of term. Incidentally, many speakers of English actually pronounce *Volkswagen* as "folks' wagon"; in their minds, the "stew" effect must take place in a very similar manner.

To make this "stew effect" crystal-clear, let me ask you: How often do you think to yourself that the name "New York" really means "new-world namesake of the English town York"? Has this ever occurred to you, even one time in your life? Probably for most Americans, especially New Yorkers, it never has — yet there is no doubt whatsoever that they "hear" and "see" both pieces inside it, in some curious sense of the terms "hear" and "see". Analogously, do you really hear the words "motors" and "mills" with their ordinary associations when you

say “General Motors” or “General Mills”? This is what I mean by my suggestion that when Germans hear *Volkswagen*, they don’t really hear it in the way that one might naively suppose they do. It is more like the way English speakers hear a common term like “doughnut” — just a phonetic unit made of familiar parts that has a familiar meaning only vaguely related to those parts.

French Compounds with Invisible Components

There are, to be sure, counterparts to this phenomenon in every language. In French, there are some stupendous examples of compounds seldom noticed by native speakers, such as *beaucoup* (“many/much”, but in literal terms, “beautiful blow”), *maintenant* (“now”, but in literal terms, “hand-holding”), *lieutenant* (“lieutenant”, but in literal terms, “place-holding”), *quand même* (“anyway, even so”, but in literal terms, “when even”), *cependant* (“however”, but in literal terms, “this-hanging”), *parmi* (“among”, but in literal terms, “by midst”), and *parce que* (“because”, but in literal terms, “by that which”).

The words *fainéant* and *vaurien* both mean essentially “lazy lout” (*fait-néant* is literally “does nothing”, and *vaut rien* is literally “is worth nothing”), but because each of them is missing a (completely silent) letter, it is not clear how many people notice the parts inside them. *Néanmoins* (“nonetheless”) is just one “t” short of being *néantmoins* (“nothing-less”). Similarly, *sourire*, meaning “smile”, lacks just the silent “s” that would make it be *sous-rire*, which means “sub-laugh” or “under-laugh”, and *toujours*, meaning “always”, lacks the silent “s” that would make it be *tous-jours*, meaning “all days”. I remember being struck by these words when I first learned them, but then I was a teen-ager, and a foreign one, to boot; native speakers, by contrast, learn most of them at age two or so, which makes it dubious that they are aware of their inner structures, either then or later.

Two curious French compounds are *vinaigre* (“vinegar”) and *gendarme* (“police officer”). Each of them is very obvious when you look at it in the right way: *vin aigre* simply means “sour wine”, and *gens d’armes* simply means “people of arms”. However, the spelling of *gendarme* is “off” just barely enough that few people realize where it comes from. And the pronunciation of *vinaigre* suggests that it should be divided, if at all, as “vi-naigre” than as “vin-aigre”, so that its meaning is obscured by its pronunciation (analogous to the case of “wardrobe”, which sounds more like “war-drobe” than “ward-robe” — or, for that matter, to “vineyard”, which sounds like “vin-yurd” instead of “vine-yard”).

A curious midway case is that of *monsieur* — literally, “my sire” or “my lord”, but, as virtually everyone knows, it actually means “sir”, “mister”, “gentleman”, and often just plain “man”. The word is pronounced quite irregularly — as if it were written “mecieu” — so that its two constituents are quite disguised. From its pronunciation in the singular, no one would ever guess what it meant. And yet in the plural, it becomes *messieurs*, sounding like “mécieu”, whose unexpected first syllable virtually forces one to go back and reparse the singular as *mon* followed by *sieur*, and thus to see the parts inside the whole after all. Certainly every French speaker comes, sooner or later, to feeling the compoundness of *monsieur* at some level, but how much this hazy recognition remains conscious is very unclear. In that sense, it is probably very much like “New York” or *Volkswagen*.

The word for “today” — *aujourd’hui* — means, when dissected, “on the day of *hui*” (*hui* being an old way of saying “today”) but today no one hears that structure inside it. Nor, I think, do people hear the three clear-as-a-bell pieces inside *après-midi* (“afternoon”) — literally, “after-half-day”, which also is how the Germans say it: *Nachmittag*. I myself remember noticing with interest, when I was a kid, that the afternoon is the time of day that comes *after noon*. What a surprise!

The French have two words for “day”: *jour* and *journée*. Likewise, “morning”

is both *matin* and *matinée*, “evening” is both *soir* and *soirée*, and “year” both *an* and *année*. Italian is analogous, with *-ata* playing the role of *-ée* (thus *giorno* and *giornata*, *notte* and *nottata*, and so on). The strange thing is that both *-ée* and *-ata* look suspiciously like feminine forms of past participles of verbs (nonexistent verbs such as *soirer* and *annare*). (Spanish, with its *jornada*, meaning “day”, extends this pattern.) Usually, a noun formed as the feminine past participle of some verb designates the product of doing that verb, as in *soufflée* or *sonata*. The implication would seem to be that *nottata* and *serata* and *année* and *journée* all came from verbs. But what would such hypothetical verbs — and they are hypothetical because none of them exists — mean?

In English one could ask a similar puzzle about words designating times of the day: why do “evening” and “morning” look like *present* participles of verbs? What would “to even” and “to morn” mean? But nobody hears the “-ing” that way, so the question is never asked.

The Curious Catchiness of Semantically Anomalous Phrases...

Italian has a pair of contrasting expressions that I find very interesting. The way one says “sweet dreams” in Italian is *sogni d'oro* — literally, “dreams of gold”. This is as one would expect. What one would not expect at all is *sogno d'ori* — “dream of golds”. That would sound ridiculous, and any Italian would laugh at it. And yet, how does one say “tomatoes” in Italian? The answer is: *pomodori* — and the meaning of this term is “apple of golds”. What is fascinating is that only a generation or two ago, people often said *pomidoro* — “apples of gold” — but this is now very rare. I have also been told that in certain places, one occasionally hears, or at least used to hear, *pomidori* — “apples of golds”. Clearly, over the past few decades, the parts of this particular compound word have been growing gradually more and more submerged inside the whole, in a most striking progression.

Another example of this phenomenon is the term *lungarno* (“along-the-Arno”), which is a name applicable to any street that runs along the river Arno. (There are analogous terms in Italian for streets along other rivers: *lungadige*, *lungotevere*, *lungopò*, and so on.) I once saw in a guide to Florence a reference to all the different *lungarni*, which, in my mind, conjured up a vision of a “street along the Arnos” — that is, a single street that goes first down *one* Arno, then down *another* Arno, and so on. And yet it is a standard term in Florence, and nobody bats an eyelash at it.

The word for “hello” in Italian is *buongiorno*, and in French, *bonjour*. In many languages, the term for “hello” is formed in a similar way — thus, in Spanish, *buenos días*. But I was caught quite by surprise when I heard radio announcers saying, in a very effusive way, *¡Muy buenos días!* and *¡Muy buenas tardes!* These phrases rang very odd to my non-native ears. After all, no one would ever say *Très bonjour* or *Molto buongiorno*; they both sound absurd — no less absurd than “Very good-bye” sounds to my anglophonic ears. One simply does not hear *bonjour* and *buongiorno* as nouns modified in a normal way by adjectives — the components are too buried. And yet — the components of these two terms cannot be all *that* deeply buried, because native speakers effortlessly adjust both components as a function of the time of day and the gender of the noun involved: *bonsoir*, *bonne nuit*, *buona sera*, *buona notte*. This would make it seem quite plausible that the adjective for “good” could comfortably be intensified by the addition of “very” — but no, it’s just not the case. However, in Spanish, doing so is fine.

Another Spanish example is the very friendly farewell, *¡Adiocitos!* This is a diminutive, in some sense, of *adios*, which is literally “to God”. Thus *adiocitos* means, roughly, “to godlets”. But of course no one who says it intends in any way to allude to polytheistic or paganistic notions — it is heard simply as a very

colloquial form of farewell, perhaps like “See ya later!” or *Ciao ciao!* On the one hand, then, this construction does bring the component “dios” to the foreground, yet on the other hand, all that is heard by a typical listener, as the phrase passes by in the bat of an eye, is a humorous “spin” on the usual way of saying “good-bye”.

...Otherwise Known as “Greasy Spoonerisms”...

Anomalous constructions like *adiocitos*, *lungarni*, and *pomodori* exemplify a surprisingly widespread phenomenon that David Moser (whose lovely phrase “splitting the etym” I borrowed above) and I have dubbed *greasy spoonerisms*. Let me swiftly explain the strange term, since it amusingly illustrates what it names. A common pejorative metaphor in American English to describe a cheap and low-quality restaurant is “greasy spoon”. Thus one might say either of the following:

Jack’s Diner *really* is a greasy spoon.

Jack’s Diner is a *really* greasy spoon.

Notice, however, that the strategies behind these two nearly-identical phrases are entirely different. The first one intensifies the metaphor by modifying the phrase “is a greasy spoon”, whereas the second one descends inside the metaphor itself and intensifies only the adjective “greasy”. The first strategy is perfectly normal and acceptable; the second one is anomalous, and consequently Moser and I, with a certain degree of humor, refer to it as a “greasy” usage.

To bring out more vividly this contrast between “normal” and “greasy” modification styles, I offer the following sentence, which mixes the two in a subtle and slippery way:

We often ate in a small and rather greasy spoon.

The noun phrase here obviously cannot be read as “small-and-rather greasy-spoon”, since “small” and “rather”, not being syntactically parallel parts of speech, cannot sensibly be yoked together this way; on the other hand, neither can it be read as “small-and-rather-greasy spoon”, since the intention is obviously to talk about a *greasy spoon*, not about a mere *spoon*. Thus when you analyze it carefully, something is clearly very wrong. And yet to most native speakers of American English, this sentence just glides right by; people may detect a vaguely humorous flavor as they read it, but little more.

The next two sentences play in a more blatant manner with the same metaphor:

Jack’s Diner is a much greasier spoon than Jill’s Eatery.

Jack’s Diner is the greasiest spoon in town.

Now in these, native listeners will probably sense, though most likely without consciously realizing what is going on, the internal tampering with what ought to be an inviolate phrase. The upshot is a mildly-felt suggestion that these eating establishments are literally *spoons* that have different degrees of *greasiness*. This semi-conscious imagery tickles the linguistic palate.

How could one get across the meanings of these two sentences while keeping the basic metaphor yet without tampering greasily with it? Well, it’s not all that easy, but it can certainly be done. Thus, instead of the first sentence, one could say, “Jack’s is *much more* of a greasy spoon than Jill’s” and instead of the second one, “Jack’s is more of a greasy spoon than any other place in town” or even “Jack’s is the *biggest* greasy spoon in town”, where “biggest” here is a code word for “most extreme example of the category”, and doesn’t at all refer to the physical size of the establishment. However, these come to mind less quickly, and are less colorful and less compact than the earlier sentences.

The color and vividness of the phrase “greasier spoon” suggested to Moser and me that the term “greasier-spoon—ism” might be an amusing generic way of referring to any example of this general linguistic phenomenon. However, once we’d concocted this phrase, we couldn’t resist the temptation to playfully readjust its own bonding strengths, thus turning it into “greasier spoonism”, which, by virtue of this twist, now amusingly exemplified its own meaning. This was a cute term, and we could have — perhaps should have — stopped there.

However, for better or for worse, we got further carried away by the winds of playfulness; this was engendered by the strong resemblance of our new term “spoonism” to the standard word “spoonerism”, which is the name of another famous class of errors. (Spoonerisms are swaps, intentional or accidental, of the initial consonants of successive words, such as saying “toin coss” for “coin toss”, or “chell specker” for “spell checker”. The notion can also be generalized to include internal sound-swaps such as “stark mocket” for “stock market”. Needless to say, spoonerisms exist in all languages; an example of a spoonerism I accidentally made in Italian was to say *dare sper contato* when intending to say *dare per scontato*. As for the name “spoonerism”, it has nothing to do with spoons; it comes from one Reverend Spooner, who was notorious for making this type of speech error.) Unable to resist temptation, we simply picked up “er” from the end of “greasier” and plunked it down in the middle of “spoonism”, thus giving rise — via a spoonerism-like sound-swapping operation, as a matter of fact — to the term “greasy spoonerism”. It’s a rather tortuous etymology, and yet not that atypical of how words and phrases evolve naturally over the centuries.

To recapitulate, then, greasy spoonerisms are not a variety of spoonerisms, and they are greasy only in a completely new sense of the term — roughly meaning “syntactically or phonetically appealing but semantically incoherent”. For example, “The spoonerism you are now reading is just about as greasy as they come” would be a very confusing sentence to anyone who didn’t know what a greasy spoonerism is.

...Or Even as “Iced Teaspoonerisms”

A self-service restaurant that I have gone to many times has for years had a sign on one of the silverware bins near the entrance, saying “ICED TEASPOONS”. I probably read right through this sign many times without thinking anything at all, but one day I noticed its hilarious inaccuracy: what the bin contains is not a bunch of frozen spoonlets, but a bunch of room-temperature spoons used for stirring iced tea! All of a sudden, the two words struck me as being completely absurd — yet I guessed that I was probably the only person who had ever noticed this at all.

The fact that the word “teaspoon” is a tightly-bonded chunk, together with the fact that the noun “iced tea” is weakly bonded, constitutes a pressure much greater than the pressure exerted by mere *logic* to make the binding between “iced” and “tea” stronger than that between “tea” and “spoon”. If, by some fluke, the latter pressure had won out, then the sign might have said “ICED-TEA SPOONS” — far better, in *my* eyes, than “ICED TEASPOONS”, but probably far worse, in most people’s eyes. Such a sign would tend to push people towards pronouncing it as “*iced-tea spoons*”, when in fact they intuitively want to say “*iced tea-spoons*”, preserving the stronger chunk and sacrificing logic. Therefore, the actual sign in the cafeteria is one of my all-time favorite greasy spoonerisms; indeed, because it is such a superb example, it occurred to me that “iced teaspoonerisms” might be a better name for the phenomenon, in which case “iced” would be synonymous with the new meaning of “greasy”. And then one might say, “The teaspoonerism you are now reading is just about as iced as they come” (and if this sounds like utter nonsense to you, don’t worry — it probably is!).

An example nearly isomorphic to the cafeteria sign just described is the extremely common phrase “used bookstore”. Have you ever noticed anything in the slightest odd about it? Probably not. And yet, it’s not the *store* that is old — it’s the *books*. To get this across crisply and sharply, one should thus bind “used” to “book”, not to the compound word “bookstore”. But that would necessitate breaking the very tight preexistent chunk “bookstore” and making a different bond supersede it, as follows: “used-book store”. Moreover, the pronunciation associated with this different way of writing it would be this: “*used-book store*”, which sounds very strange to the ears of people who have a lifetime of experience saying and hearing “used *book-store*”, illogical though the latter may be. As a result, nobody would ever advocate the changes that, in the name of logic or underlying semantics, I am suggesting. Cheap syntax defeats the semantic pressure hands down — and the bookstorism you are now reading is just about as used as they come!

The Surprising Frequency of Spooneristic Greasiness

I keep on running into examples of this phenomenon, each one outdoing earlier ones in absurdity. Thus, to my amazement, I have come across books entitled *Incompressible Aerodynamics* and *Single Electronics*. The former is about the dynamics of incompressible air; the latter is about how single electrons behave in conductors and semiconductors. Given that these are both solemn, no-nonsense scientific treatises, I doubt that the authors or publishers found anything funny in either title. Many names of scientific professions are also unintentional greasy spoonerisms, and are heard by most people with no amusement, two notable examples being “high-energy physicists” (who need not be particularly energetic people) and “abnormal psychologists” (who probably tend to be quite normal).

On the other hand, many language-sensitive people are aware, on a certain level, of greasiness and exploit it deliberately to good effect, even if they have no name for it and could not exactly articulate what is going on. Thus, a book about the scientists and engineers developing Reagan’s hugely misguided “Star Wars” project several years ago was called *Star Warriors*, in which the standard agent-denoting “-er” suffix was appended, in a mildly distorted form, to the second word in the term, thereby turning “Star” into a greasy modifier of “Warriors”. It’s quite clear that this title was a deliberate attempt at being catchy and clever, although it’s not a particularly humorous phrase. There are many similar examples of playful greasiness used in advertising and publicity.

I cherish two greasy spoonerisms uttered by my wife Carol about our son Danny, both involving greasy usages of “little”. The first — “Let him enjoy his little babyhood!” — was uttered when he was about a year old. Obviously, it wasn’t his babyhood that was little; rather, Carol was intending “-hood” to apply to the entire phrase “little baby”. A couple of years later, when Danny was no longer always so angelic, Carol commented with exasperation, “He can be the littlest jerk you ever saw!” I was flabbergasted by the greasiness of this phrase, and laughed out loud. Carol, however, didn’t see anything funny whatsoever about what she had said. Only after I’d nongreasily paraphrased it as “He can be the most extreme example of the category ‘little jerk’ you ever saw!” did she clearly see its illogicality. A friend later pointed out that her meaning could be caught more succinctly by the paraphrase “He can be the *biggest* little-jerk you ever saw!”

Other Languages are No Less Greasy

Regrouping and combining parts of expressions in a “greasy” manner is a surprisingly efficient, or perhaps I should say *compact*, technique for intensification and other types of semantic tweaking, and for that reason it is

widespread in all the languages I have encountered. For example, consider these two French and Italian phrases: *la presque totalité* and *la quasi totalità*. Literally, these become “the almost totality”, meaning “almost the entire set”, as in *la presque totalité des spectateurs* or *la quasi totalità degli spettatori* (“almost the entire set of spectators”). To an English ear they sound initially very strange, but then we anglophones are not so innocent; we, too, greasily abuse the word “almost”, at least in some technical contexts, such as “almost periodicity”, a well-known term in mathematics. Here, “almost” gives the surface appearance of being an adjective modifying the noun “periodicity”, whereas in fact it is an adverb modifying the adjective “periodic”. The only semantically logical way of building up this hierarchical term would be “almost-periodic-ness”, but no one but *no one* would write it this way. The mere fact that “periodicity” is an already-existing and familiar word just wipes out all rival pressures due to logic. Similar analyses hold for the French and Italian terms just mentioned.

Then there is “almost everywhere convergence”. Even if you leave off “almost”, it still is disturbing: “everywhere convergence” means “everywhere-converging-ness”, but once again, who would buy that, let alone “almost-everywhere-converging-ness”? The irony is that use of too much logic makes the phrases *less* comprehensible rather than *more* so. We humans are apparently more comfortable accepting, in a relatively mindless manner, whatever our prior lexicon and syntax jointly press upon us, than dealing with the precision of exquisitely-punctuated hierarchical notations.

Two further Italian examples are the colloquial phrases *d'accordissimo* (“sure thing”, “right on”) and *in frettissima* (“with extreme haste”). These intensified phrases are both made from simple prepositional phrases — the adjectival phrase *d'accordo* meaning “in agreement” and the adverbial phrase *in fretta* meaning “with haste” — by the standard device of adding *-issimo* to an ordinary adjective or adverb. (For example, the adjective *buono* [“good”] becomes *buonissimo* [“extremely good”] and the adverb *bene* [“well”] becomes *benissimo* [“extremely well”].) This same device is then applied even though *d'accordo* and *in fretta* are phrases, not simple words. Thus what actually happens is that *-issimo* gets underhandedly added to nouns, an illegal act that normally would hurt italophonic ears sorely. (An analogous illegal act in English would be to say “with veryhaste” or even “with hastest”.) Even the underlying gender of the nouns shows through, since *fretta*, which is feminine, gets the feminine “a” at the end, whereas *accordo*, which is masculine, keeps the masculine ending “o”. Although this grammatical agreement of the suffix with the noun brings out the blatant illogicality of the whole construction, it doesn’t matter; people hear these phrases as making perfect sense, *e bastissima!* It would be far less weird if, say, *d'accordo* had collapsed, as *pomodoro* did, into a single apostrophe-free word, so that women said *Sono d'accorda* while men said *Sono d'accordo*, but that isn’t the case in Italian — at least not so far.

Cannibalistic Iced Teaspoonerisms

A most interesting subclass of the iced-teaspoonerism/greasy-spoonerism phenomenon involves cases where the semantically logical construction would demand two copies of one word to occur in a row — a syntactic awkwardness that causes the unconscious mind to balk and drop one of them. This phenomenon, wherein one word or short phrase “eats” an identical one, has been termed, in the literature on human error-making, a *cannibalism*. (Cannibalism is a special case of the more general phenomenon of *haploglossy*, which is the dropping of any element.) Sometimes this type of word-dropping gives rise to blatant errors that the speaker would immediately recognize as such, whereas in other cases the effect is

so subtle that speakers will insist till the cows come home that what they said is perfectly acceptable.

Thus I once heard someone refer to “a very highrise”, meaning, obviously, “a very *high* highrise”. This is an indisputable example of an error, and consequently no one would dispute it. Similarly, I myself once told someone that I was “a very newcomer to Ann Arbor”, and I also described Escher and Bach as having been “very latecomers” to my book *Gödel, Escher, Bach*, in the sense that I had written a great deal of the book before thinking of letting them play significant roles in it. These two utterances were certainly errors on my part. Non-native speakers of English might think that mere insertion of a blank space in front of “comer” would fix everything up in both cases (“a very new comer”, “very late comers”); however, we anglophones simply don’t use “comer” in that way (in fact, you won’t even find “comer” in a dictionary). The proper fixes would have to be “a very new newcomer” and “very late latecomers”.

I have also heard someone refer to “a very hot dog”, meaning “a very *hot hot* dog”, and I myself once said, “He’s a really hotshot”, meaning “a really *hot* hotshot”. Analogously, a colleague once remarked, “The Japan Prize will probably be given to some very bigwig”, obviously meaning “to some very *big* bigwig”. These three examples are also all pretty clear-cut cases of errors.

On the other hand, one of my favorite examples of a borderline case is this remark: “Paul has a very sweet tooth” (“to have a sweet tooth” means “to like sweets very much”). The person who said this, and several people with whom I have discussed this example, weren’t convinced by the analogy to “Jack’s Diner is a very greasy spoon”; rather, they felt that “sweet” here is somehow functioning as a perfectly *ordinary* modifier of “tooth”, so that intensification via “very” should work in the standard way. I remember strongly insisting at the time that this was a cannibalistic greasy spoonerism, and that the proper way of phrasing it should be “Paul has a very *sweet* sweet tooth”. Now, however, I think that that is a flawed analysis, and I find the only logical fixes to be sadly flat: “Paul has quite a sweet tooth” or “Paul has very much of a sweet tooth”. That flatness is perhaps one reason why the greasy version sounds so appealing.

A similar sentence I recently caught myself in the act of penning is this: “As parents, we were very soft touches.” To one part of me, this sounds incredibly natural and right, yet another part of me protests, insisting that this feeling is an illusion, that the phrase is syntactically incoherent, and that the only proper way to say it is to decannibalize it into “we were very *soft* soft touches” — because, after all, “soft touch” is a *chunk* whose parts can’t be tampered with. But then the first part of me fights back by asking, “If the adjective ‘soft’ and the noun ‘touch’ can fuse into a tight unit, why can’t a similar fusion happen with the adjectival phrase ‘very soft’ and the noun ‘touch’?” As these inner voices battle back and forth, the sum total of me starts to lose all bearings of correctness and coherency and I no longer know what’s right and what’s wrong.

My wife once remarked, “It only happens once in a blue, blue moon”, meaning, of course, “It happens incredibly rarely”. This on-the-fly trick of hers for intensifying “once in a blue moon” was very subtle. It bypassed the far more obvious route “once in a *very* blue moon” (although I would guess that unconsciously, she most likely did sniff this route out a little ways before abandoning it). Both alternatives are greasy, to be sure, but to my ear, “very blue moon” sounds ever-so-slightly greasier, because “very blue” forms such a strong unit that it leaves “moon” sitting there a little isolated and naked, thus violating the “blue moon” chunk, whereas saying “blue, blue moon” almost magically has the effect of leaving “blue moon” virtually intact while at the same time intensifying “blue”, and thus intensifying the phrase as a whole. Altogether, this was a tricky maneuver by which Carol somehow managed to leap through two hoops at once.

I am guilty of having once said, "I felt like a somewhat weirdo." Where along the spectrum does this utterance fall? One is tempted to redo the bondings, and to write "a somewhat-weird-o", as if "-o" were a productive suffix by which nouns were commonly made out of adjectives. However, although this happens at times, "-o" is far from being a standard suffix, and in any case, "somewhat weird" is not a simple adjective to which one can attach a nominalizing suffix. Another alternative would be to see this as a case of cannibalism, and to propose the fix "I felt like a somewhat *weird* weirdo." However, this sounds very wrong. All weirdos being by definition weird, this says essentially nothing. A far better, although sadly bland, fix would be "I felt somewhat like a weirdo." It just lacks all the pizzazz of the original utterance.

My mother once remarked about a small but rich California town, "Danville is probably a pretty hotbed of Republicans." Does the standard cannibalism fix work here? That is, did she mean "a pretty *hot* hotbed"? I don't think so, although it is tempting. Once again, I am disappointed to say that the only *logical* fixes that I can come up with are along the disappointing lines of "quite a hotbed", "pretty much of a hotbed", and so forth. Examples like this make it fairly obvious why there is considerable pressure to follow the greasier and often cannibalistic routes, silly though they sound under scrutiny.

In the innocent-seeming prose of this very paper, the phrase "an obviously compound word" occurred a few pages ago. You probably read right through it. Certainly I did — in fact I did so many times, as I proofread that section over and over. But then one time, I realized that what I *really* meant was either "an obvious compound word" or else "an obviously *compound* compound word". Or was it?

Greasy Spoonerisms and Frame Blends

One of the subtlest examples I have ever come across of the greasy-spoonerism phenomenon is this sentence, here quoted verbatim as it was uttered to me by a scientist I had just met:

I'm a mathematical physicist — a *very* mathematical physicist.

The first half is simply a standard greasy title for a technical profession, analogous to "fuzzy logician" and so many others. Nothing too surprising here. It's the leap taken in the second half that really throws things off. All of a sudden, the words "mathematical" and "physicist" are supposed to be taken as a *standard* adjective-noun pair. The problem is that, since they come on the heels of the preceding phrase, such a univalent interpretation is impossible; instead, a blur of both meanings is heard. That is, one can't help hearing the word "very" both as a *greasy* modifier of the tight chunk "mathematical physicist" and as a *normal* modifier of the detached adjective "mathematical". This is a typical cannibalism, whose fix should be "a *very mathematical* mathematical physicist", yet that sounds preposterously pedantic.

It may seem terribly nitpicky of me to be dissecting this little utterance in such gruesome detail, but I feel that it is precisely here that we once again come very close to the core issues of this paper: How are combinations of words felt in the unconscious? What do they do to one? How much are parts heard inside wholes? What kinds of resonances are awakened, deep down? And my feeling is that the answer is far from black-and-white: that in fact one very often experiences a superposition of two or more different hearings, forming an intangible frame blend of a sort that is anathema to traditional linguistics as well as to traditional computational models of language and thought, wherein *disambiguation* is the central goal and to which the suggestion of the existence of ineradicable, essential ambiguity is intolerable.

The issues raised by this example are perhaps easier to see in the following strongly analogous statement:

Sean used to be a high jumper — a *very* high jumper.

Once again, the first half is a run-of-the-mill greasy term (remember that it's not the *jumper* that is high, but the *jump* or the *jumping*); we need not dwell on it. It's the second half that again seems to be asking the listener to suddenly hear the whole as its parts — but here it seems clear that a cannibalism has been committed as well. Thus the meaning of the second half might be rendered as "Sean was a *very high* high-jumper." The only problem is, just as before, that it's not the high-jumper who was very high, but his high-jumping. How, then, to fix it? The only solution I can see is the excruciatingly strained-sounding phrase

Sean used to be a high-jumper — a *very high-jumping* high-jumper.

Not too beautiful, is it?

Incidentally, I feel compelled to point out that the bond-strengths in what I've just displayed are *still* all wrong, but if I were to correct them as well, throwing in en-dashes and em-dashes and whatnot, the whole thing would be so disgusting-looking that you might feel a need to throw up before reading on. I thus hope you are glad I chose to spare you — but if you feel like fixing this still-flawed sentence up yourself, you are welcome to do so.

On High Jumpers and Heavy Sleepers

You may feel uncomfortable with my rejecting "a *very high* jumper" as a well-constructed phrase — it sounds so appealing, so natural, so smooth. I agree, it does. However, this does not mean that there isn't something very strange about it. There is even something strange about "a high jumper". At this point, you are surely thinking, "Now he's gone too far!"

Perhaps you will understand my discomfort with "a high jumper" if I simply turn it around: "a jumper who is high". This sounds like a description of someone who either is on top of a big rock (physically high) or is tipsy (mentally high) — but in it there is not the slightest implication that this person is an athlete who does the high jump. Part of the reason is that in the reversal, the chunk "high jump" has been destroyed — but that's not the full reason. After all, even if I retain the chunk — "a high-jumper who is very high" — it *still* sounds like I'm talking about a drunkard or a rock-climber, as opposed to a talented high-jumper. By similarly turning around such innocent-seeming stock phrases as "a big eater", "a fast driver", "a beautiful diver", and "an elegant dresser", you will find that they are all far less straightforward than they appear.

A native speaker of French described herself quite un-self-consciously to me as *une très petite dormeuse*. At first, I took her quite literally to be saying that she was a *tiny* sleeper, an image that threw me for a loop, but then a moment later I realized that she simply meant "someone who sleeps very little" — close in meaning to our term "a light sleeper." I had never reflected upon that phrase, or its opposite, "heavy sleeper", before hearing the French phrase. All of a sudden, these terms seemed very strange.

We can make use of hyphens, en-dashes, and em-dashes, as was done above in the discussion of very long compound noun phrases, in an attempt to "logically purify" the phrase "heavy sleeper": thus, "a heavily-sleep-er", or even "person who heavily-sleep-s" (it has to be "heavily" rather than "heavy", because "sleep" here is a *verb*, not a noun). These notations expose the underlying meaning of the phrase, but the phrase itself, taken literally, conjures up a totally different and rather droll image. Yet it is seldom taken literally or even heard literally. I

wonder how likely it would be that an obese person who said, "I'm a very heavy sleeper" would be aware of the irony. Quite unlikely, I'd guess. Nor do I think that most listeners would hear the humor in it. Likewise, I'd bet that the *très petite dormeuse* had no notion of the potential humor in the phrase she was using.

Greasy Number-words

Earlier, I described the pathway by which "greasy spoonerism" evolved, with the novel term "spoonism" being replaced by the more familiar "spoonerism". This phenomenon, whereby an attractive standard word replaces an ugly artificial one produced by some mechanical process, may seem idiosyncratic, but in fact it is a relatively common occurrence in the evolution of compound words. The innocent-seeming ordinal number "twenty-first" is a familiar example that you may never have noticed in English. For the sake of perspective, let us consider it in conjunction with its French counterpart *vingt-et-unième*. In either language, the basic mechanical strategy is to add the standard ordinalizing suffix ("th" in English, *-ième* in French) to the cardinal number — thus, "twenty-one—th" or *vingt-et-un—ième*. But once this logical step is taken, there arises intense phonetic pressure to *merge* the suffix with the immediately preceding component — that is, to *submerge* it therein. Thus the word "one" furtively slinks away from its logical mate "twenty" and greasily sidles up to "th", giving "twenty—oneth". This greasy regrouping yields the ugly nonword "oneth". Now comes the step I was alluding to. The strange-sounding "oneth" is such a close semantic cousin to the standard ordinal "first" that there ensues a slide down the path of least resistance, with the upshot being our standard term "twenty—first". When you look at it this way, "twenty-first" is in essence an *incoherent arithmetical addition* of a cardinal number ("twenty") with an ordinal number ("first") — very, very greasy indeed!

The French, in a valiant attempt at being less greasy (*i.e.*, more coherent), stop one step earlier than we do; instead of surrendering to the temptation of converting the gauche-looking nonword *unième* into the familiar and alluring *premier*, thus making *vingt-et-premier*, they boldly resist and stick with *vingt-et-unième*. The French ordinals thereby seem less greasy than ours. To be *completely* ungreasy, however, this ordinal would really have to be written as follows: *vingt-et-un—ième*, and it would have to be pronounced correspondingly — namely, with the nasal *un* followed by *ième*, and no elision between them. But the French aren't willing to sound *that* ugly — not even in the name of semantic coherence!

Number-words as Quintessential Wordmelts

Words for numbers — especially numbers between 1 and 100 — are in general a most interesting category of compounds, because ideally, in any given language, they should form a perfect, flawless pattern, yet they almost never do. Instead, each individual number-word is the result of a unique constellation of phonetic and semantic pressures interacting over hundreds or thousands of years.

Just look at our own number-words to see some of the weird effects. How can one explain the irregularity of the sequence "eleven", "twelve", "thirteen"? Or the pronunciation of "thirteen", "fourteen", "eighteen", and "nineteen" (with their doubled "t" sound)? Why is the suffix "-teen" rather than "-ten"? Why does "forty" not contain a "u"? And so on.

There are even greater mysteries in the French number-words. Why, as they move up the teens, do they suddenly switch from rather opaque terms like *quinze* and *seize* to perfectly transparent compounds like *dix-sept* and *dix-huit*? Why do they say *vingt-et-un* but not *vingt-et-deux*? What's the logic behind *vingt*, *trente*, *quarante*, *cinquante*, *soixante*? Why do they say *vingt-et-un* but *quatre-vingt-un*? Why

do they say *soixante-et-onze* but *quatre-vingt-onze*? Why an “s” in *quatre-vingts* but no “s” in *quatre-vingt-un*? Why, for that matter, do the French not say *septante*, *octante* (or *huitante*), *nonante* — as the Swiss and Belgians do? Hardly a perfect crystal.

But the irregularities of number-words in English and French seem like child’s play in comparison with those in Hindi. Like ours, the Hindi number-words reflect the base-ten system, and are thus trying to form a regular ten-by-ten matrix, but for whatever historical reasons, the components in virtually every one of these compound words wound up melting together in very unpredictable ways, like wax candles left sitting against each other in the blazing sun for several days. In every Hindi number between 11 and 99, one still sees clear traces of its origin, but to an outsider like me, there is scarcely a single one of them that seems completely predictable. I would have a very hard time learning to count to 100 in Hindi. Without further ado, here is the subtle pattern I am alluding to:

<u>1–10</u>	<u>11–20</u>	<u>21–30</u>	<u>31–40</u>	<u>41–50</u>	<u>51–60</u>	<u>61–70</u>	<u>71–80</u>	<u>81–90</u>	<u>91–100</u>
<i>ek</i>	<i>gyarah</i>	<i>ikkis</i>	<i>ikattis</i>	<i>iktalis</i>	<i>ikyavan</i>	<i>iksath</i>	<i>ik’hattar</i>	<i>ikyasi</i>	<i>ikyanve</i>
<i>do</i>	<i>barah</i>	<i>bais</i>	<i>battis</i>	<i>bayalis</i>	<i>bavan</i>	<i>basath</i>	<i>bahattar</i>	<i>bayasi</i>	<i>banve</i>
<i>tin</i>	<i>terah</i>	<i>teis</i>	<i>taim tis</i>	<i>taim talis</i>	<i>tirpan</i>	<i>tirsath</i>	<i>tihattar</i>	<i>tirasi</i>	<i>tiranve</i>
<i>car</i>	<i>caudah</i>	<i>caubis</i>	<i>caum tis</i>	<i>cavalis</i>	<i>cauvan</i>	<i>caum sath</i>	<i>cauhattar</i>	<i>caurasi</i>	<i>cauranve</i>
<i>pa mc</i>	<i>pandrah</i>	<i>paccis</i>	<i>paim tis</i>	<i>paim talis</i>	<i>pacpan</i>	<i>painsath</i>	<i>pac’hattar</i>	<i>pacasi</i>	<i>pacanve</i>
<i>chah</i>	<i>solah</i>	<i>chabbis</i>	<i>chattis</i>	<i>chiyalis</i>	<i>chappan</i>	<i>chiyasath</i>	<i>chihattar</i>	<i>chiyasi</i>	<i>chiyanve</i>
<i>sat</i>	<i>satrah</i>	<i>sattais</i>	<i>saim tis</i>	<i>saim talis</i>	<i>sattavan</i>	<i>sarsath</i>	<i>sat’hattar</i>	<i>sattasi</i>	<i>sattanve</i>
<i>ath</i>	<i>atharah</i>	<i>atthais</i>	<i>artis</i>	<i>artalis</i>	<i>athavan</i>	<i>arsath</i>	<i>athhattar</i>	<i>atthasi</i>	<i>athanve</i>
<i>nau</i>	<i>unnis</i>	<i>untis</i>	<i>untalis</i>	<i>uncas</i>	<i>unsath</i>	<i>unhattar</i>	<i>unyasi</i>	<i>navasi</i>	<i>ninyanve</i>
<i>das</i>	<i>bis</i>	<i>tis</i>	<i>calis</i>	<i>pacas</i>	<i>sath</i>	<i>sattar</i>	<i>assi</i>	<i>navve</i>	<i>sau</i>

For your information, here the letter “c” is always pronounced roughly as “ch”, and the boldface “r” and “t” represent retroflex consonants. Moreover, the vowel “a” represents the bland schwa sound (as in “woman”), whereas bold vowels are long — thus “a” and “i” are roughly as in “father” and “machine”. Finally, boldface “m” represents a nasalization of the preceding vowel.

Of course, Hindi is written not in roman letters, but in the same script as Sanskrit was written — Devanagari. This system of writing is based on symbols that represent entire syllables consisting of consonant-clusters followed by simple or complex vowels, such as “ka”, “tu”, “bho”, “trai”, “strau”, and so forth. Thus the consonant and the vowel following it are fused into a single symbol, in which one can usually but not always see both parts as individuals. Moreover, consonant clusters such as “bl” and “str” are also captured by fused symbols — the technical term is *ligatures* — made up of pieces that suggest the individual sounds. But just like the number-words shown above, such consonantal ligatures are very much like pieces of wax that have melted together over millennia, sometimes leaving almost no trace of the original pieces.

At the opposite end of the spectrum from the extremely irregular number-words in Hindi are the number-words in Chinese. If you know how to count to ten in Chinese, then you can immediately count to 99 with no problem.

<u>1–10</u>	<u>11–20</u>	<u>21–30</u>	<u>31–40</u>	<u>41–50</u>	<u>51–60</u>	<u>61–70</u>	<u>71–80</u>	<u>81–90</u>	<u>91–100</u>
<i>yǐ</i>	<i>shíyǐ</i>	<i>èrshíyǐ</i>	<i>sānshíyǐ</i>	<i>sǐshíyǐ</i>	<i>wūshíyǐ</i>	<i>liùshíyǐ</i>	<i>qǐshíyǐ</i>	<i>bǎshíyǐ</i>	<i>jiūshíyǐ</i>
<i>èr</i>	<i>shíér</i>	<i>èrshíér</i>	<i>sānshíér</i>	<i>sǐshíér</i>	<i>wūshíér</i>	<i>liùshíér</i>	<i>qǐshíér</i>	<i>bǎshíér</i>	<i>jiūshíér</i>
<i>sān</i>	<i>shísān</i>	<i>èrshísān</i>	<i>sānshísān</i>	<i>sǐshísān</i>	<i>wūshísān</i>	<i>liùshísān</i>	<i>qǐshísān</i>	<i>bǎshísān</i>	<i>jiūshísān</i>
<i>sì</i>	<i>shísì</i>	<i>èrshísì</i>	<i>sānshísì</i>	<i>sǐshísì</i>	<i>wūshísì</i>	<i>liùshísì</i>	<i>qǐshísì</i>	<i>bǎshísì</i>	<i>jiūshísì</i>
<i>wǔ</i>	<i>shíwǔ</i>	<i>èrshíwǔ</i>	<i>sānshíwǔ</i>	<i>sǐshíwǔ</i>	<i>wūshíwǔ</i>	<i>liùshíwǔ</i>	<i>qǐshíwǔ</i>	<i>bǎshíwǔ</i>	<i>jiūshíwǔ</i>
<i>liù</i>	<i>shíliù</i>	<i>èrshíliù</i>	<i>sānshíliù</i>	<i>sǐshíliù</i>	<i>wūshíliù</i>	<i>liùshíliù</i>	<i>qǐshíliù</i>	<i>bǎshíliù</i>	<i>jiūshíliù</i>
<i>qī</i>	<i>shíqī</i>	<i>èrshíqī</i>	<i>sānshíqī</i>	<i>sǐshíqī</i>	<i>wūshíqī</i>	<i>liùshíqī</i>	<i>qǐshíqī</i>	<i>bǎshíqī</i>	<i>jiūshíqī</i>
<i>bā</i>	<i>shibā</i>	<i>èrshibā</i>	<i>sānshibā</i>	<i>sǐshibā</i>	<i>wūshibā</i>	<i>liùshibā</i>	<i>qǐshibā</i>	<i>bǎshibā</i>	<i>jiūshibā</i>
<i>jiǔ</i>	<i>shíjiǔ</i>	<i>èrshíjiǔ</i>	<i>sānshíjiǔ</i>	<i>sǐshíjiǔ</i>	<i>wūshíjiǔ</i>	<i>liùshíjiǔ</i>	<i>qǐshíjiǔ</i>	<i>bǎshíjiǔ</i>	<i>jiūshíjiǔ</i>
<i>shí</i>	<i>ershí</i>	<i>sānshí</i>	<i>sǐshí</i>	<i>wūshí</i>	<i>liùshí</i>	<i>qǐshí</i>	<i>bǎshí</i>	<i>jiūshí</i>	<i>yǐbāi</i>

If you examine this array, you will see that it is as purely regular, as purely crystalline, as would be a table of numerals written out in decimal notation. This might suggest that *all* Chinese compound words form a perfectly logical, rational system. Nothing could be further from the truth: compounds in Chinese are just as crazy and jumbly as in any language on earth.

The Nature of Compounds in Chinese

Chinese is a language that is built on compound words as deeply as is German. As I remarked earlier, single characters — each of which is pronounced as exactly one syllable — are of course the building blocks of the language. But single characters are by no means the natural units that Chinese speakers draw upon to make larger structures — no more than we English speakers construct sentences by stringing together long series of monosyllables. We anglophones string together *words*, and the Chinese do analogously — they mentally assemble long strings of *cí*, not *zì*.

There is a strong pressure in modern Mandarin Chinese to make words of exactly two characters. A simple example is *dàxiàng*, the standard way of saying “elephant”. The curious thing is that this word consists of *dà*, which means “big”, followed by *xiàng* which also means “elephant”. But people almost never say *xiàng* alone. On the other hand, when it is used inside compounds, “elephant” is usually rendered by *xiàng* alone — thus *xiàngyá*, literally meaning “elephant tooth”, is the word for “ivory”. It would sound like a joke if someone said *dàxiàngyá*. The second component, *yá*, is a one-syllable term for “tooth” used just in compounds; to say “tooth” by itself, one uses the two-character word *yáchǐ*, each of whose components on its own means “tooth” or “cog”. Curiously, *yá* by itself can also mean “ivory”, as in *yákuài*, which means “ivory chopsticks”. And of course, the stand-alone word for “chopsticks” is not just *kuài*, but the two-syllable word *kuàizi*, meaning roughly “chopstick-things”. On and on it goes.

A second strong pressure in modern Mandarin is, whenever possible, to juxtapose pairs of two-syllable words, thus giving elegant four-syllable structures (*e.g.*, *feicháng yōuqián* — “extremely wealthy”, but literally “un-usual have-money”). However, this does not mean that there are no three-syllable words — there are plenty of them (*e.g.*, *yōuyìsi* — “interesting”, but literally “have-meaning”), just as there are one-syllable words. There is thus a real-time metrical pressure in spoken Mandarin — a pressure to assemble structures having a “good rhythm” — that has no real analogue in any of the western languages I have studied.

In any case, contrary to a popular myth, the vast majority of Chinese words are polysyllabic, not monosyllabic, which implies that the vast majority of Chinese words are compounds. However, as in German, these words range all over the map in terms of their degree of opacity. Some are very clear *a priori*, some are comprehensible (or somewhat so) only after the fact, and some are just totally dumbfounding both before and after you know their meanings. Just as Strothmann’s thesis falls apart, so would its Chinese counterpart. Let us look at some examples.

Perhaps my favorite example of all Mandarin compounds is the word for “thing”: *döngxi*, which literally means “east–west”. Yet no native speaker *ever* hears it that way. It is just as seamless-seeming to Chinese as are *beaucoup* and *sourire* to French speakers, or “crackpot” and “understand” to English speakers. I suppose one can imagine that its etymology has something to do with the idea that from east to west, from north to south, all that one ever finds is *things*. But virtually no one ponders its etymology: for 999 speakers out of 1,000, *döngxi* is just an unanalyzed chunk.

A handful of other examples, escalating in opacity, are these: *xiāoshüo* (“small-tell”, which, once you know that it means “a novel”, you can sort of see why), *mǎshàng* (“horse-upon”, meaning “immediately”, which seems not too illogical), *míngbái* (“bright-white”, but it means “to understand”, and this is definitely not obvious), *fángshì* (“house-stuff” or perhaps “domestic-matters”, but it actually means “sexual intercourse” [though only in the case of a married couple]), *guàchī* (“hang-tooth”, but it means “to mention” — a compound just about as opaque as “crackpot” or *dōngxi*).

It may come as a disappointment to readers with a romantic impression of China that Chinese placenames, when translated character-by-character into English, sound as bland as the names that property developers in America have learned to give to their housing tracts. Thus *Xīhú* means “Westlake”, *Shànghái* means “Sea-upon”, *Héběi* means “River North”, *Guāngzhōu* means “Broad Province”, *Xiāngshān* means “Sweet Hill”, and so forth. The places start to seem far less exotic and more mundane when you hear their names this way. On the other hand, is this how the Chinese *really* hear them? As I asked earlier, how much do we hear the two pieces in “New York” or in “New England”? How much of the “holly” and the “wood” do we hear in “Hollywood”? How much do we hear “washing” and “town” inside “Washington”, or “spring” and “field” inside “Springfield”? For that matter, how much do we hear “shake” and “spear” inside “Shakespeare”?

Dinosaur Names: Better Transparent or Opaque?

This gets us back to some of the questions asked in the very early parts of this essay, about the flavors evoked subliminally by the names of such exotic things as subatomic particles and dinosaurs. In this regard, the contrast between Chinese and English, being very extreme, is highly pertinent.

Here are the Chinese names for some of the best-known dinosaurs, together with their piece-by-piece translations as well as their standard English names:

<i>lēilóng</i>	thunder dragon	brontosaurus
<i>bàwánglóng</i>	tyrant-king dragon	tyrannosaurus
<i>yǐshǒulóng</i>	wing-hand dragon	pterodactyl
<i>jiànóng</i>	sword dragon	stegosaurus
<i>sānjiāolóng</i>	three-horn dragon	triceratops

All the rest of the dinosaurs, of course, have similarly transparent names in Chinese. This would certainly suggest that the image — or, better yet, the aura of connotations — that swirls around these beasts in the mind of a Chinese speaker is quite different from ours. After all, because their names all end in *lóng*, dinosaurs are being suggested as varieties of *dragons*, and dragons of course play a central role in Chinese mythology. The Chinese culture is so permeated by the notion of *lóng* that it’s almost impossible to imagine that large parts of the feelings associated with dragons would not seep into children’s absorbent and uncritical minds.

Furthermore, a term like “wing-hand” is so much more concrete and mundane than the exotic, strangely-spelled word “pterodactyl” that one can’t help but assume that the feelings evoked thereby have to be radically different. Of course, if one knows ancient Greek, one realizes that “pterodactyl” means nothing other than “wing finger”, which is almost the same as the Chinese term — but most of us don’t know ancient Greek. On the other hand, we do sense *something* behind a strange structure like “pterodactyl” — we sense an ancient and alien way of describing things, we sense the existence of parts inside the word, even if they are in a completely opaque language. And when we repeatedly hear the

ending “-saurus”, we of course build up a sense of its implications even if we are never told explicitly that in ancient Greek it means “lizard”.

Transparent Words for Elementary Particles Suggest Classical Behavior

From dinosaurs, let us move to particles. Here is a table of a few particle names, similar to the one shown earlier for dinosaurs:

<i>zhīzî</i>	essence seed (“ur-bit”?)	proton
<i>diànzî</i>	electric seed	electron
<i>zhōngzî</i>	middle seed	neutron
<i>zhōngwēizî</i>	middle seedlet	neutrino
<i>guāngzî</i>	light seed	photon

It is a bit misleading to choose “seed” as the sole translation of *zî*, since that very frequent and densely-laden character has a complex range of meanings, including “person”, “son”, “egg”, and “kernel” — one might even say “bit”, to use Poul Anderson’s term.

Once again, the clear impression comes through that the Chinese names are far more concrete than ours are. In fact, compare the anglicized Chinese terms “middle seed” and “middle seedlet” with Poul Anderson’s terms “neitherbit” and “weeneitherbit” — there is a very strong feeling of similarity. To me, such concreteness carries with it strong although subliminal suggestions that the objects being described are very ordinary objects, except much smaller. To use a more precise technical term from physics, this kind of terminology suggests that these tiny things are like *classical* particles, rather than something radically different, something that cannot be visualized in any ordinary way, something whose behavior is altogether unfamiliar and radically counterintuitive. I need not dwell on it, but readers will probably realize that I am in essence talking about the famous wave–particle duality of quantum mechanics, and such unexpected consequences of it as Heisenberg’s uncertainty principle and the nondeterminacy of the laws of physics, which Albert Onestone so abhorred.

It is this subliminal appeal to standard, classical intuitions about the behavior of things that I find so misleading about both the Chinese words and Poul Anderson’s words. Subatomic particles are nothing at all like tiny “bits” or “seeds”. They are entities having no precedents or analogues in our macroscopic lives. To me, they *demand* terms that suggest their totally alien nature.

And yet, there is a bit of a paradox here. When I was a young boy, I found opaque terms like “photon” and “neutrino” endlessly suggestive and mystery-laden — but over the years, I gradually got used to them and they became quite routine items in my vocabulary. Then, when I read “Uncleftish Beholding”, my mind was once again jolted and disoriented for a few brief days by Anderson’s eerie terms, concrete though they were, and to some extent I was again able to savor that feeling of strangeness and awe that I had first known when I encountered the more opaque terms. What a surprising turnaround!

Emotional Issues Stirred up by Compound Words

One fundamental question I am grappling with in this exploratory essay is how much one is aware of the parts making up the wholes that one so effortlessly perceives and produces in using language. A second basic question is the effect on native speakers of a language in which the presence of very obvious parts is pervasive: how differently do they perceive the world from speakers of a language in which parts are rarely visible and generally harder to see?

I have to admit that I have a certain emotional preference for languages in

which the parts tend to be somewhat cryptic rather than always on the surface. For example, once when I was living in Germany and reading a German newspaper, I came across an article that spoke about war reparations being paid by Germany to Israel. The article used the standard German word for “reparation”, namely *Wiedergutmachung*. Maybe some people could have just read on, but I could not: I was absolutely floored by this compound word. Taken apart, it means “again-good-making”. I simply couldn’t believe what I felt was being suggested — that Germany was going to be able to “make everything good again” by paying Israel money? That six million dead Jews could be “fixed up” with some cash? I found this tacit implication unbelievably crude and repugnant. Someone could argue that one should simply take *Wiedergutmachung* as a chunk and not dissect it into its parts. Or one could argue that “reparations” really means almost the same thing, in the sense that it is a cognate to “repair”, which means “fix”. But in English, one doesn’t hear that connection — or if we do, it’s very tenuous at best.

Another example of the crudity of visible parts that I ran across in German was the word for “nipple”: *Brustwarze*, which means “breast-wart”. The female breast is certainly a shape of great beauty to many people, and the nipple is seldom considered to be a blotch or blight on it. Yet what can one possibly think upon encountering this term? Is a nipple not being called a “wart” in the plainest possible of terms? And is that not slandering the beauty of breasts? A German woman to whom I posed this question merely shrugged her shoulders and said, “But isn’t that exactly what a nipple is?” She wasn’t in the slightest put off by the term; it just seemed accurate to her. By contrast, the native speakers of English — male and female — that I asked about this term were either bothered or amused by it, never indifferent.

The native German speaker’s reaction suggested an interesting hypothesis to me. Maybe translating *Warze* as “wart” is just a tiny bit inaccurate — and for no other reason than the fact that *Warze* is associated with nipples! In other words, where “wart” has essentially zero positive associations in an anglophonic mind, the word *Warze* for Germans has at least a slightly positive aura — aureola? — of associations, because of its role in the word *Brustwarze*. A German speaker thus doesn’t feel so appalled that a nipple is seen as a kind of a wart, because after all, *certain kinds of warts are nice!* It sounds a bit circular, but in the end I think it makes some sense. The upshot would be that the extremely natural-seeming equivalence *Warze* = “wart” is simplistic; that the two strongly cognate words, although they designate the same entities (warts, of course) when used in isolation, do not behave identically when embedded in compounds (at least in *one* compound!), and hence the equality is imperfect.

Shortly after writing these remarks about warts and breastwarts, I was surprised to hear myself call a friend a “worry wart”. Though I had meant to use the gentlest of epithets, the word “wart” suddenly jumped out at me, with its cloud of repulsive connotations. Yet I knew that no native speaker would hear any repulsive flavor in this stock phrase of American English. This suggested to me that in English as well as German, even the ugly word “wart” can go close to unheard. This gave me pause for thought about the repulsiveness of “breastwarts”.

Similar issues come up with the word *Reisschleim*, which denotes a kind of rice gruel. Its second component, *Schleim*, is a cognate to our “slime”; and like “slime” for us anglophones, *Schleim* is used by Germans to describe the moistness on the slippery body of a reptile, the trail left on the ground by a slithering slug, or in general, any kind of repulsive-seeming ooze (most often yellow and green). Thus how could something to eat conceivably be called “slime”? Well, the answer, according to this hypothesis, is that it is *not* called “slime”; it is called *Schleim*, and that word, though cognate to and almost synonymous with “slime”, has an aura of connotations that is ever so slightly, ever so subtly different.

Appealing though this hypothesis of non-equivalence is, I am not sure that I really buy it. In its place, I can offer a slightly different hypothesis: that German speakers are every bit as repulsed by the words *Warze* and *Schleim* when they occur in isolation as we anglophones are by "wart" and "slime", but that when those words are buried inside certain compounds, they just don't hear them any more. Or rather, they hear them in a certain sense, but that sense involves unconsciously tampering with the associations to fit the context. Thus, they suppress some of the negativity surrounding warts and slime, and retain other, more neutral properties, such as the idea that a wart is a protuberance, that it is darker than the surrounding area, that it is slightly irregular in shape, and so on; or that slime is kind of gooey, perhaps a bit warmish, hard to stir, and so on.

Another possibility is, of course, that German speakers are completely repulsed both by nipples and by gruel, and don't mind letting the world know it.

I don't really know the ultimate answer. All I know is that as an English speaker, I cannot help but feel appalled at what to me is the undeniable crudity or bluntness of certain German compound words, such as the ones just mentioned, and others such as *Elternteil* ("parents-part", which is the only way of saying "parent"), *Säugetier* ("suckle-beast", which means "mammal"), *häßlich* ("hateful", but it means "ugly"), *Fingerhandschuh* ("finger-hand-shoe", one way of saying "glove"), *Denksportaufgabe* ("think-sport-exercise", but it means "brainteaser"), and so on. I know that we in English have strange words too, but it seems to me that comparable linguistic or conceptual crudity, if that is what it is, is far rarer. Or perhaps we have just as much crudity if you take the trouble to dissect some of our Latin-based and Greek-based words etymologically, but most speakers are completely unable to do so, and hence the parts are truly absent for them.

Some More Animal Names

My wife, when I told her that the German word for "tortoise" — *Schildkröte* — translates literally as "shield-toad", practically howled with laughter. She found this unimaginably silly, and instantly commented, "Shield-toad'?! That's like calling an eagle a 'feather-cow'!" And I had to agree that "shield-toad" is a rather comical and implausible characterization of a tortoise. Would anyone, on first hearing it, be able to guess what it meant? Incidentally, the Dutch and the Swedes are with the Germans, on this one: *schildpad* and *sköldpadda* are compounds having precisely the same origins.

We say "hippopotamus" and the very sound of the word suggests roly-poly-ness and makes us smile; however, we have no idea of the origin or structure of this funny-sounding term. The Chinese say *hémâ*, and what they hear is also what the Swedes hear in their term *floodhäst*: "river horse". There's no mystery there, although calling a hippopotamus a variety of horse seems pretty far-fetched to this anglophone. The Germans say either *Flußpferd* ("river horse") or *Nilpferd* ("Nile horse"). The latter is admittedly kind of cute, rather than crude.

We say "rhinoceros" and in the very sound we feel something awkward and queer-looking. (Or do I fool myself? Could such a word have been the name for something extremely graceful?) Of course, we don't know where it comes from. The Germans, on the other hand, say *Nashorn*: "nosehorn". Pretty direct.

How would we English speakers have felt about anteaters if they, like hippopotamuses and rhinoceroses, had inherited their name from ancient Greek, and thus been yclept "myrmecophaguses"? Would it have made any difference in our feelings at all? That question is largely what this paper is about.

The Germans pull no punches with their name for a skunk: *Stinktier*, or "stinkbeast". Once again, it seems to me that this is crude and unfair, because it focuses unnecessarily on just one aspect of the poor beast. Just because someone

has occasional body odor does not mean that they deserve forever being called “Stink-arms” and nothing else. (And don’t tell me that *Stink* has some pleasant connotations for Germans precisely because of its embeddedness in their word for this adorable little animal — I won’t believe it!)

And finally, we anglophones don’t hear a trace of “lion” inside “leopard”, yet italophones do, since *leopardo* is obviously a compound made of *leo* and *pardo*, and to an Italian ear, the Latin *leo* roars “lion” loud and clear, even if the true Italian word for “lion” has two extra letters (*leone*). This second component *pardo* is very curious. Italians encounter it again in *ghepardo* (“cheetah”) and in *gattopardo*, which clearly is some kind of cat (*gatto*); however, few Italians have even a hunch what *pardo* means. Most dictionaries don’t have any entry for it at all, and when as a second choice, you look up the compound *gattopardo*, you find that dictionaries waffle hugely on what one of these creatures is, ranging from a leopard to a cheetah to an ocelot to a serval. If you look long and hard enough, you will finally discover that *pardo* is just an antiquated word for “leopard”, but for our purposes, that doesn’t much matter; the interest for us is the variable degree to which these animal names using *pardo* seem decomposable, ranging from the near-transparency of *gattopardo* (it was once even acceptable to write *gatto pardo*) to the translucency of *leopardo* and then the quasi-opacity of *ghepardo*, whose *ghe* is as semantically vacuous as the “chee” of “cheetah”.

Upwrapping

Do we, or do we not, hear parts? The answer, as this essay has attempted to demonstrate, is incredibly complex and shaded, amazingly context-dependent. I conclude my explorations with an anecdote that harks back to Anderson’s piece.

My *Doktorvater* Gregory Wannier — my Ph.D. advisor — was a physicist who grew up speaking Swiss-German but moved to the United States in his twenties and lived there for most of his life. Since he was linguistically talented, his English eventually became as good as his German, and he could easily pass for a native American. But one day, as we were discussing the hydrogen atom, the term “waterstuff” popped out of his mouth before he knew what was happening. He caught himself instantly, replaced it by “hydrogen”, and simply went on. I never heard him make this or any similar error again. It was an amazing little German glitch in the midst of perfectly American-sounding English.

Did this slip mean that Wannier actually *conceived* of hydrogen as some kind of stuff associated with water? Or did it simply mean that that day, for whatever reason, the German language was activated more than usual in his brain, and the word *Wasserstoff* was somehow triggered and then a mechanical piece-by-piece translation was done unconsciously? Or did Wannier *always* have to fight to suppress saying “waterstuff” when he spoke English, no matter what the circumstances were, and this time his hidden censors were simply too slow to the draw? I will of course never know.

What I do know is that for me, having the unheard makeup of words brought out in a sharp and brainteasing way, whether in a tiny tongueslip by my old teacherfather or in Poul Anderson’s greatly clever and longlasting gameword, is an enlightening and deeply ghostshaking lifehappening, and can sometimes even churn up a huge and hazy cloud of closeknit mindgropings, which when written down might, with luck, reach across to other thinkbeasts and in them trigger wonderings of a likeful kind. Such would at least be a happy outcome.

* * *

* * *

*

Appendix I:

Complete text of Poul Anderson's "Uncleftish Beholding"

For most of its being, mankind did not know what things are made of, but could only guess. With the growth of worldken, we began to learn, and today we have a beholding of stuff and work that watching bears out, both in the workstead and in daily life.

The underlying kinds of stuff are the *firststuffs*, which link together in sundry ways to give rise to the rest. Formerly we knew of ninety-two firststuffs, from waterstuff, the lightest and barest, to ymirstuff, the heaviest. Now we have made more, such as aegirstuff and helstuff.

The firststuffs have their being as motes called *unclefts*. These are mighty small: one seedweight of waterstuff holds a tale of them like unto two followed by twenty-two naughts. Most unclefts link together to make what are called *bulkbits*. Thus, the waterstuff bulkbit bestands of two waterstuff unclefts, the sourstuff bulkbit of two sourstuff unclefts, and so on. (Some kinds, such as sunstuff, keep alone; others, such as iron, cling together in chills when in the fast standing; and there are yet more yokeways.) When unlike unclefts link in a bulkbit, they make *bindings*. Thus, water is a binding of two waterstuff unclefts with one sourstuff uncleft, while a bulkbit of one of the forestuffs making up flesh may have a thousand or more unclefts of these two firststuffs together with coalstuff and chokestuff.

At first it was thought that the uncleft was a hard thing that could be split no further; hence the name. Now we know it is made up of lesser motes. There is a heavy *kernel* with a forward bernstonish lading, and around it one or more light motes with backward ladings. The least uncleft is that of everyday waterstuff. Its kernel is a lone forwardladen mote called a *firstbit*. Outside it is a backwardladen mote called a *bernstonebit*. The firstbit has a heaviness about 1840-fold that of the bernstonebit. Early worldken folk thought bernstonebits swing around the kernel like the Earth around the Sun, but now we understand they are more like waves or clouds.

In all other unclefts are found other motes as well, about as heavy as the firstbit but with no lading, known as *neitherbits*. We know a kind of waterstuff with one neitherbit in the kernel along with the firstbit; another kind has two neitherbits. Both kinds are seldom.

The next greater firststuff is sunstuff, which has two firstbits and two bernstonebits. The everyday sort also has two neitherbits in the kernel. If there are more or less, the uncleft will soon break asunder. More about this later.

The third firststuff is stonestuff, with three firstbits, three bernstonebits, and its own share of neitherbits. And so it goes, on through such everyday stuffs as coalstuff (six firstbits) or iron (26), to ones more lately found. Ymirstuff (92) was the last until men began to make some higher still.

It is the bernstonebits that link, and so their tale fastsets how a firststuff behaves and what kinds of bulkbits it can help make. The worldken of this behaving, in all its manifold ways, is called *minglingken*. Minglingers have found that as the uncleftish tale of the firststuffs (that is, the tale of firstbits in their kernels) waxes, after a while they begin to show ownships not unlike those of others that went before them. So, for a showdeal, stonestuff (3), headachestuff (11), potashstuff (19), redstuff (37), and bluegraystuff (55) can each link with only one uncleft of waterstuff, while coalstuff (6), sandstuff (14), germanstuff (32), tin (50), and lead (82) can each link with four. This is readily seen when all are set forth in what is called the *roundaround board of the firststuffs*.

When an uncleft or bulkbit wins one or more bernstonebits above its own, it takes on a backward lading. When it loses one or more, it takes on a forward lading. Such a mote is called a *farer*, for that the drag between unlike ladings flits it. When free bernstonebits flit by themselves, it may be as a bolt of lightning, a spark off some faststanding chunk, or the everyday flow of bernstoneness through wires.

Coming back to the uncleft itself, the heavier it is, the more neitherbits as well as firstbits in its kernel. Indeed, soon the tale of neitherbits is the greater. Unclefts with the same tale of firstbits but unlike tales of neitherbits are called *samesteads*. Thus, everyday sourstuff has eight neitherbits along with its eight firstbits, but there are also kinds with five, six, seven, nine, ten, and eleven neitherbits. A samestead is known by the tale of both kernel motes, so that we have sourstuff-13, sourstuff-14, and so on, with sourstuff-16 being by far the mostfound. Having the same tale of bernstonebits, the samesteads of a firststuff behave almost alike minglingly. They do show some unlikenesses, outstandingly among the heavier ones, and these can be worked to sunder samesteads from each other.

Most samesteads of every firststuff are unabiding. Their kernels break up, each at its own speed. This speed is written as the *half-life*, which is how long it takes any deal of the

samestead to shift itself into half as much. The doing is known as *lighttrotting*. It may happen fast or slowly, and in any of sundry ways, offhanging on the makeup of the kernel. A kernel may spit out two firstbits with two neitherbits — that is, a sunstuff kernel — thus leaping two steads back in the roundarounds board and four weights back in heaviness. It may give off a bernstonebit from a neitherbit, which thereby becomes a firstbit, and thrusts the uncleft one steady up in the board while keeping the same weight. Often, too, a mote is given off with neither lading nor heaviness, called the *weeneitherbit*. In much lighttrotting, a mote of light with most short wavelength comes out as well.

For although light oftenest behaves as a wave, it can be looked on as a mote — the *lightbit*. We have already said by the way that a mote of stuff can behave not only as a chunk, but also as a wave. Down among the unclefts, things do not happen in steady flowings, but in leaps over midway bestandings that are forbidden. The knowledge-hunt of this is called *lump beholding*.

Nor are stuff and work unakin. Rather, they are groundwise the same, and one can be shifted into the other. The kinship between them is that work is like unto weight manifolded by the fourside of the haste of light.

By shooting motes into kernels, worldken folk have shifted samesteads of one firststuff into samesteads of another. Thus did they make ymirstuff into aegirstuff and helstuff, and they have afterward gone beyond these. These heavier firststuffs are all highly lighttrottish and therefore are not found in the greenworld.

Some of the higher samesteads are *splitly*. That is, when a neitherbit strikes the kernel of one — as, for a showdeal, ymirstuff-235 — it bursts it into lesser kernels and free neitherbits; the latter can then split more ymirstuff-235. When this happens, weight shifts into work. It is not much of the whole, but nevertheless it is awesome.

With enough strength, lightweight unclefts can be made to togethermelt. In the Sun, through a row of strikings and lighttrottings, four unclefts of waterstuff in this wise become one of sunstuff. Again, some weight is lost as work, and again this is greatly big when set beside the work gotten from a minglingish doing such as fire.

Today we wield both kinds of uncleftish doings in weapons, and kernelish splitting gives us heat and bernstoneness. We hope to do likewise with togethermelting, which would yield an unhemmed wellspring of work for mankindish goodgain.

Soothly we live in mighty years!

* * *

Besides his newbooks and truthbooks, the writer has forthshown in *Likething Worldken Sagas/Worldken Truth*, *The Warehouse of Dreamishness* and *Worldken Sagas*, and other roundaroundses.

Appendix II:

Complete text of my translation of “Uncleftish Beholding”

For most of its existence, humanity did not know what things are made of, but could only guess. With the advent of science, we began to learn, and today we have a theory of matter and energy verified by observation, both in the laboratory and in daily life.

The basic types of matter are the *elements*, which combine in various ways to give rise to the remaining types. Until recently, we knew of ninety-two elements, from hydrogen, the lightest and simplest, to uranium, the heaviest. Now we have made more, such as neptunium and plutonium.

Elements are composed of particles called *atoms*. These are extremely small: the number of atoms in one ounce of hydrogen is equal to 2 times 10 to the 22nd power. Most atoms can combine to make what are called *molecules*. Thus, the hydrogen molecule consists of two hydrogen atoms, the oxygen molecule of two oxygen atoms, and so on. (Some types, such as helium, do not combine; others, such as iron, bond to form crystals when in the solid state; and there are yet other states of matter.) When different atoms combine in a molecule, they make *compounds*. Thus, water is a compound of two hydrogen atoms with one oxygen atom, while a molecule in a biological cell may contain a thousand or more atoms of these two elements, along with carbon and nitrogen atoms.

At first it was believed that the atom (Greek for “partless”) was a solid object that could not be further divided; hence the name. Now, however, we know that it is composed of smaller particles. There is a massive *nucleus* with a positive electric charge, and around it one or more light particles with negative charges. The smallest atom is that of ordinary hydrogen. Its nucleus is a single positively charged particle called a *proton*. Outside it is a negatively charged particle called an *electron*. The proton has a mass about 1840 times that of the electron. Early scientists thought electrons orbited the nucleus like the Earth around the Sun, but now we know they are more like waves or clouds.

In all other atoms there are also other particles, roughly as massive as the proton but with no charge, known as *neutrons*. We know a type of hydrogen with one neutron in its nucleus along with the proton; another type has two neutrons. Both types are rare.

The next larger element is helium, which has two protons and two electrons. Ordinary helium also has two neutrons in the nucleus. If there are more or fewer, the atom will soon split. More about this later.

The third element is lithium, with three protons, three electrons, and its own share of neutrons. And so it goes, on through such everyday materials as carbon (six protons) or iron (26), to ones discovered more recently. Uranium (92) was the last, until people began to produce elements even beyond it.

It is the electrons that bind, and so their number determines how an element behaves and what kinds of molecules it can help make. The science of this phenomenon, in all its diversity, is called *chemistry*. Chemists have discovered that as the atomic number of the elements (that is, the number of protons in their nuclei) increases, after a while the elements begin to manifest properties similar to earlier ones. Thus, for example, lithium (3), sodium (11), potassium (19), rubidium (37), and cesium (55) can each bond with only a single atom of hydrogen, while carbon (6), silicon (14), germanium (32), tin (50), and lead (82) can each bond with four. This is readily seen when all are displayed in what is called the *periodic table of the elements*.

When an atom or molecule gains one or more extra electrons, it acquires a negative charge. When it loses one or more, it acquires a positive charge. Such a particle is called an *ion* (Greek for “traveler”), since the attraction between opposite charges moves it around. When free electrons move, it may give rise to a bolt of lightning, a spark from some solid substance, or the ordinary flow of electricity through wires.

Returning to the atom itself, the more massive it is, the more neutrons as well as protons in its nucleus. Indeed, soon the number of neutrons is larger. Atoms having the same number of protons but different numbers of neutrons are called *isotopes*. Thus, ordinary oxygen has eight neutrons along with its eight protons, but there are also varieties with five, six, seven, nine, ten, and eleven neutrons. An isotope is identified by the total number of both nuclear particles, so that we have oxygen-13, oxygen-14, and so on, with oxygen-16 being by far the most common. Having the same number of electrons, the isotopes of an element behave almost the same chemically. They do reveal some different properties, especially among the more massive ones, and these can be used to separate isotopes from each other.

Most isotopes of every element are unstable. Their nuclei split, each at its own rate. This rate is measured by the *half-life*, which is the length of time it takes any quantity of the isotope

to convert itself into half as much. This process is known as *radioactive decay*. It may happen quickly or slowly, and in any one of several different ways, depending on the composition of the nucleus. A nucleus may eject two protons with two neutrons — that is, a helium nucleus — thus jumping back two places in the periodic table and four units in mass. It may emit an electron from one of its neutrons, which thereby becomes a proton, thus shifting the atom one step up in the periodic table while keeping its mass constant. Often, a chargeless, massless particle, called the *neutrino*, is emitted. In much radioactive decay, a particle of light of very short wavelength emerges as well.

Although light most often behaves like a wave, it can be regarded as a particle — the *photon*. We have already said in passing that a material particle can behave not only like an object, but also like a wave. In the atomic realm, things do not happen continuously, but in jumps, skipping over intermediate states that are forbidden. Research on this subject is called *quantum theory*.

Nor are matter and energy unrelated. Rather, they are fundamentally the same, and one can be converted into the other. Their relationship is that energy is equal to mass times the speed of light squared.

By shooting particles into nuclei, scientists have converted isotopes of one element into isotopes of another. Thus they have made neptunium and plutonium from uranium, and they have even gone beyond these. These more massive elements are all highly radioactive and therefore are not found in nature.

Some of the higher isotopes are *fissile*. That is, when a neutron hits the nucleus of one — uranium-235, for example — it breaks it into smaller nuclei and free neutrons; the latter can then split more uranium-235. When this happens, mass is converted into energy. It is not a large percentage of the mass, but nevertheless it is impressive.

With enough energy, lightweight atoms can be made to fuse. In the Sun, through a series of collisions and radioactive decays, four atoms of hydrogen become a single atom of helium. Again, some mass is lost as energy, and this energy is enormous when compared to the energy obtained from a chemical process such as combustion.

Today we use both kinds of atomic processes in weapons, and nuclear fission provides heat and electricity. We hope to do likewise with fusion, which would yield an unlimited source of energy for human profit.

Truly we live in great times!

* * *

Besides his novels and essays, the author has published in *Analog Science Fiction/Science Fact*, *The Magazine of Fantasy and Science Fiction*, and other periodicals.

Presented at the Nobel Symposium 92 in Stockholm, Sweden, August 1994.

In Sture Allén (ed.), *Of Thoughts and Words — Proceedings of the Nobel Symposium 92: The Relation Between Language and Mind*, pp. 217–266, 1995. Imperial College Press and the Nobel Foundation.

Distributed by World Scientific Publishing Co. Pte. Ltd, River Edge, New Jersey, and London, UK.