TIKKI TIKKI TEMBO

The Chemistry of Protolanguage

Yuri Tarnopolsky
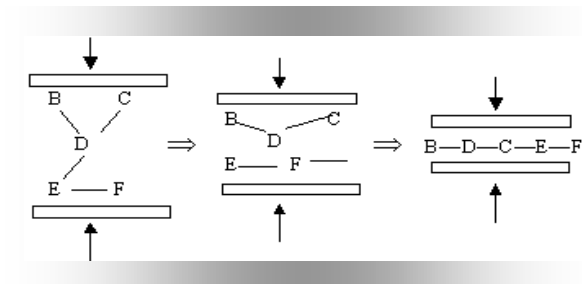
2004

ABSTRACT

Protolanguage (Derek Bickerton) in linguistics corresponds to an evolutionary stage preceding the grammaticalized language as we know it. It could be possible to reconstruct the principles of protolanguage by turning to most general principles of evolution in a larger picture, of which chemistry is a relevant part. Both linguistics and chemistry are discrete combinatorial systems. Considering the chemical origin of life, chemical analogies might offer some insight into the origin of mind, language, and society, all of which developed on the platform of life. The conceptual basis for discrete combinatorial systems, including chemistry and language, can be found in Pattern Theory (Ulf Grenander) where ideas, utterances, and molecules are *configurations*. To draw the parallel further, chemistry uses its own *language* of chemical nomenclature to represent non-linear molecular structures as linear strings of symbols. Chemistry pays particular attention to the intimate mechanisms of structural transformations. A tentative concept of the mechanism of protolanguage generation is suggested as kinetically controlled linearization of a typically non-linear observable configuration through a non-observable thought. Generation of linear expressions in protolanguage is viewed as a process of generalized chemistry, going from a typically non-linear initial state through a transition state toward the linear output, under the constraint of a maximized preservation of configuration topology.

CONTENTS

The words and verses differ, each from each,

Compounded out of different elements...

Lucretius (*De rerum natura*, II)

The order and connection of ideas is the

same as the order and connection of things.

Spinoza (*Ethica*, I, VII)

1. INTRODUCTION

Mark S. Baker in his book *The Atoms of Language* (Baker, 2001) drew a consistent analogy of linguistics with chemistry. Within the *Principles and Parameters* framework, a language is similar to a chemical element in the sense that it is a combination of certain parameters. Baker acknowledged that most people would associate **words** with atoms of language, but he simply put this view aside as "correct—in one sense." (Baker, 2001, p. 51). He was, of course, right, pointing to the Periodic System as a metaphor for the combinatorial nature of language. Moreover, the book manifested the true chemical spirit: it was built around numerous **observable** linguistic examples, as any typical chemical monograph is built around hundreds of structures and their transformations.

> For a linguist with comparative interests, a large part of the fun of doing linguistic research is searching out all the fascinating, deep, and intricate differences in how languages work. Indeed, the only thing that gives a comparable thrill is discovering the deep and fascinating ways in which they are all the same. (Baker, 1995)

The above quotation reveals to me a kindred soul: I could sign it if *linguist* was substituted for *chemist* and *molecules* for *languages* because chemists deal with concrete

individual and observable objects, while parameters are rather abstract structural invariances.

It would be unbecoming for a linguist to say "It's Greek to me" about chemistry, anyway, but there is some genuine feeling of kinship between both areas. Linguists evoke chemistry as the science epitomizing not only complexity but also its successful conquest.

The kinship was prophesized long before the birth of chemistry and after the birth of the famous Greek Democritus:

> The words and verses differ, each from each,
> Compounded out of different elements—
> Not since few only, as common letters, run
> Through all the words, or no two words are made,
> One and the other, from all like elements,
> But since they all, as general rule, are not
> The same as all. Thus, too, in other things,
> Whilst many germs common to many things
> There are, yet they, combined among themselves,
> Can form new who to others quite unlike.
> Thus fairly one may say that humankind,
> The grains, the gladsome trees, are all made up
> Of different atoms (Lucretius, 1958, Book II).

Chemistry, on its part, has been using extensive linguistic parallels for nucleic acids and proteins since the discovery of their relation. Moreover, much earlier, chemistry developed its own tongue with a lexicon heavily borrowed from Greek and a refined grammar with codified flexions and word order.

Mark Baker's book was the last drop into the bucket of observations that I had accumulated over a significant time. This paper is an attempt of a chemist to view words as atoms of **a chemistry**.

I am a chemist without any linguistic credentials whatsoever, but with a life long interest in languages. I am, to a variable degree, familiar with properties of such languages as Russian (native), English (current), German (studied at school), French, Hungarian, Japanese, Hebrew, and a few others. My very limited hands-on experience with the non-Indo-European languages, as well as a better (but by no means perfect)

knowledge of both Indo-European but diametrically opposite English and Russian, persuaded me that, with all the striking differences in their design, all languages perform the **same function** with the **same means**. Neither the opulence of Bantu languages, with their classifiers and suffixes, nor the intricately woven ribbons of the Na-Dene verbs could shake my conviction. The function is a representation of a non-linear "source," whatever it is, and the means is an optimal linearization of the non-linear representation. When I had looked into linguistic literature, I found plenty of support.

> The vocal-auditory channel has some desirable features as a medium of communication: it has a high bandwidth, its intensity can be modulated to conceal the speaker or to cover large distances, and it does not require light, proximity, a face-to-face orientation, or tying up the hands. However it is essentially a serial interface, lacking the full two-dimensionality needed to convey graph or tree structures and typographical devices such as fonts, subscripts, and brackets. The basic tools of a coding scheme employing it are an inventory of distinguishable symbols and their concatenation. (Pinker and Bloom, 1990).

For over twenty years I have been watching the development of Pattern Theory (Grenander, 1976-2003), sometimes from a close distance, regarding it as a general approach to complex systems consisting of atom-like elements and connecting bonds. It became clear to me that this mathematical theory of everything nicely covered not only molecules and languages but also every discrete combinatorial system we could come in touch, and did it with an unprecedented combination of generality and realism.

Furthermore, I have witnessed the entire genesis and evolution of the science of complexity, starting from Prigogine (1984) , who formulated the most fundamental principles of complex natural systems such as life, mind, and society, and further toward Artificial Life (Adami, 1998) where languages and molecules were of the same kin already at the inception (Eigen, 1971-1979). "Natural" here is the opposite of "artificial," such as virtual reality where people can walk on the ceiling and turn into wolves right before your eyes.

I am also familiar with the language of musical notation, a couple of programming languages, and, due to my profession, with the curious language of

chemical nomenclature invented by organic chemistry to verbally communicate the non-linear molecular structure.

Finally, the language of poetry—rarely spoken in everyday life—is my bonus pass to a gym where one can exercise linking distant meanings and close sounds.

The enormous literature comprising computational, formal, traditional, and historical linguistics, Artificial Intelligence, Artificial Life, mathematical structures, physics of open systems, chemistry, and details of Pattern Theory is probed here only highly selectively and superficially. The growing but still manageable bibliography on language evolution and computation has been nicely collected and presented at the website of University of Illinois at Urbana-Champaign (Language Origin, WWW).

My intent is not to formulate a theory—this should be entrusted to professionals—but to offer a new (but organically grown!) spice for the boiling cauldron of linguistic ideas. Whoever likes the aroma can use it for meditation, inspiration, and, who knows, for some fun time after a Ph.D. thesis. I wish to share a widest and most comprehensive—an illusive but honorable goal—view of the intellectual jungle where linguistics and chemistry are of the same blood. In short, if it all boils, then down and up to Pattern Theory.

I believe that my outsider status, as well as the claim for a larger picture, grants me the privilege of choosing my own far-from-academic style—which is just being natural. I cannot walk on the ceiling. But I have another uncommon gift: I see the world with the eyes of a chemist.

I further refer to the following key figures of painting a large picture with language in the landscape: Lucretius, Ulf Grenander, George Zipf, Manfred Eigen, Ilya Prigogine, Walter Ross Ashby, Rene Thom, George Hammond, all of them, except Lucretius and Zipf, natural scientists and mathematicians. I mention other profound linguistic thinkers in the main text. I am sure most of the ideas of this paper can be found in the literature and I apologize if I failed to find them.

I widely use the WWW sources where it is possible. They may die out with time, but a peculiar life-like property of the Web is that the new ones will be cooked, could be searched for, and found, garnished with ads. For better or worse, money will never be out of the larger picture but I hope it will not fill up the entire canvas.

2. PREVIEW OF MAIN IDEAS

My initial thesis is: to compare language and chemistry we have to view them as natural phenomena within a larger picture.

> Language is embedded in human psychology and society, and is ultimately governed by the same principles as galaxies and mesons. (Hurford, 2003, p.38)

We shall look at both from a more general view than either chemistry or linguistics and in both linguistic space and evolutionary time.

From the evolutionary perspective, there must be a fundamental truth in the concept of protolanguage (Bickerton, 1990, 1995, Calvin and Bickerton, 2000), from which the full-blown languages evolved. Bickerton's earlier vigorous and polemic book (Bickerton, 1981) was full of important large-picture ideas, some of which will be echoed here. The recent collective volume (Language Evolution, 2003) with Bickerton's contribution, summarizes the current status of the problem and will be often referred to.

Protolanguage is not just a playground for imagination. The use of protolanguage, within the framework of algorithmic AI, has been discussed and attempted for a simplified communication between a human and a computer or robot, see for example, Varshavskaya (2002) and Billard (2002), but the results were not too encouraging.

If there is something truly universal for all languages, from pre-protolanguage to the modern street slang to the flashy lingo of *The New Yorker* art reviews, and from

English to Mohawk, it must be so not only in space but also in time. Here we will be looking for a **universal property of modern language, applicable also to protolanguage and its subsequent evolutionary record**. By definition, it is something that cannot be found in either extant or extinct formal structures, but only diachronically: along the time axis.

If we deal with a non-grammatical protolanguage we have to abandon all the theories of fully developed language, together with all the contrived and artificial examples and even the entire ontology of the past several millennia of culture. This is a big relief because it is difficult to find a non-trivial linguistic statement which has not been contested during the linguistic wars in which a chemist has no part.

The form of dialogue, which had lost its appeal since Plato, Galileo, and Bishop Berkeley, was revived by Juan Uriagereka (1998) in his popularization of minimalist syntax. It inspired me to design the following introductory exchange, where vague words, like "typical" or "source," are used to avoid obdurate head-on statements and meaningless terms such as meaning (why meaningless? to define meaning, you have first to define meaning).

1.　**Q.**　What is typical for natural sciences?

　　A.　They deal with observable objects and processes.

2.　**Q.**　What is typical for a natural process?

　　A.　A certain parameter, for example, energy, changes in a preferred direction, unless an external influence prevents the natural course. The apple naturally falls down, if not caught by a human hand.

3.　**Q.**　What is typical for molecules and utterances?

　　A.　Both are **structures**. Most generally, structure (not in the sense of "mathematical structure") is a set of elements and a set of pairs, i.e., connections between some of them. Graph, especially colored and labeled, i.e., with values or markers at arcs and nodes, is a fair image of structure. Configuration of Pattern Theory is a better one.

4. **Q.**   Atoms interact. How do words interact?

    A.   They form linear strings: utterances. Some strings stick together, which reveals the affinity of the words to each other, and participate in verbal exchange. Other strings do not hold together and cannot be used.

5. **Q.**   What is the immediate source of the utterance?

    A.   Thought. We do not know what it is except that it is a mental state or process. A thought, unlike an apple, is never shared and never observed—yet. (Very fortunately. But for how long?).

6. **Q.**   What is the source of the thought?

    A.   **Observable** reality: thing, situation, object, process, relation, sensation, information, sign, phrase, text, image, cue, signal, question, remark, utterance— anything that could be shared or witnessed by at least two people.

8. **Q.**   What is the relation between a thought and its expression as an utterance?

    A.   Since we have no means of observing a thought, we have to go to the **source** of the thought outside the individual mind. The **image** of the source preserves topological relations between components of observable reality. The structure of thought is, **hypothetically**, not necessarily linear. Structures in PT (configurations, images, patterns) and chemistry are typically non-linear.

9. **Q.**   What happens when an utterance in protolanguage or language is generated from its source?

    A.   **Linearization** of a typically non-linear configuration. In any case it must happen somewhere between the source and the expression like "I see Og take Ug meat" using Bickerton's example (Calvin and Bickerton, 2002).

10.   **Q.**   What is protolanguage? What is full language?

      A.   **Protolanguage** is a linearization of thought in which the binary connections in the source are represented by pairs of adjacent words, which may not be possible for all pairs in the source.

           Language is same as protolanguage, but the connection can be expressed, in addition to word order, by means of morphemes, regardless of the adjacency.

11.   **Q.**   How does language emerge and evolve from protolanguage?

      A.   By mutation and replication in populations of utterances communicated in a social group. "You say potahto, I say potato." Language **is** a form of life. You say nukelar, I say nuclear.

12.   **Q.**   What is typical for a form of life?

      A.   In addition to replication and mutation, there is the important property of homeostasis, i.e. the ability to restore peace after turbulence and to minimize a deviation from "the middle road," often by taking another evolutionary pathway.

13.   **Q.**   What do language and chemistry have in common, apart from being discrete combinatorial systems?

      A.   First, there is a stage of the fleeting and **non-observable thought** in language generation (it cannot even be remembered if not put into words) and there is a fleeting and typically **non-observable transition state** in a chemical transformation. Second, chemistry has a distinct language of its own, intended to **linearize** nonlinear chemical images.

Next, some points of the dialogue will be expanded in a series of chapters, without going into too much detail, because what truly relates both chemistry and linguistics is the zillion devils in the details.

Finally, the above dialogue will be illustrated by computer-aided examples of linearization at the level of a hypothetical protolanguage called Nean. The computer, however, will act in the dumbest role of a generator of random numbers and could be

substituted just by flipping coins. To attribute any algorithm to a genesis of a natural system equals to designating a creator.

My concluding thesis is that an almost mindless process of linking two atomic names together (Ug big) because the atoms are linked in the source, could be sufficient to launch protolanguage. The random and mindless process of natural selection in the populations of utterances could be sufficient to launch the evolution of full language with its rainforest exuberance of species.

3. CHEMISTRY AND LINGUISTICS: SISTER SCIENCES

It is no wonder that chemistry and linguistics could be jointly found at over 800,000 web pages (in 2003) because any university is a natural place for them to rub shoulders. It is much less common to meet both on the same page of a scientific text. Nevertheless, an acquaintance has been struck about half century ago when chemists compared strings of amino acids and nucleotides with lines of text. Today the Web search for DNA + language delivers almost 1 million sites. "DNA linguistics," as it is called, has become a largely independent direction of research (Searls, 1992).

The time around 1950 was a period of extensive planting of new ideas. The concept of Artificial Intelligence, including automatic translation, was formulated. The application of the idea of **transition state** to molecular structure by George Hammond completely revolutionized chemistry. It was also the time of the articulation of the science of complexity by Ilya Prigogine, the arrival of the "extravagant" and "controversial" ideas of George Zipf, and the advent of the formal linguistics of Noam Chomsky. (I believe that the incredibly fruitful time was the aftershock of the WWII: the soil was fertilized with the ashes).

Marc Baker's book (Baker, 2002) is only one evidence of the interest in chemistry on the part of linguists, although the most significant one. There are other examples of a mutual curiosity.

At the conference *Language as an Analogy in the Natural Sciences* held in Munich in 1997, the chemist Pierre Laszlo presented an essay *Belaboring the Obvious:*

Chemistry as Sister Science to Linguistics. Laszlo (1997) emphasized binary mixing and combining as the way chemical experience was acquired throughout history.

L. Vlasov compared more common in nature metals and less common non-metals with consonants and vowels, expanding the following metaphor:

> Nature speaks to us in the language of chemical compounds. And each of these is a sort of combination of chemical "letters," or elements occurring on Earth. The number of such "words" exceeds three million. But there are only just over a hundred "letters" in the chemical "alphabet." (Vlasov, 1970, Story No.20).

The following was addressed to students of chemistry: "Learning organic chemistry is like learning a new language, a language that is both verbal and pictorial" (Ege, 1989, p.2). So, don't be afraid. Chemistry is fun.

Chomsky (2002) used the comparison of the burgeoning descriptive linguistics with chemistry to express optimism in developing a compact theory after a sufficient body of knowledge has been accumulated.

Chomsky noted that chemistry had achieved understanding of the nature of its invisible objects through the union with physics, which has not shaken the factual knowledge itself. It implies that a further advance in linguistics may depend on including it into a **larger picture** of the world drawn by natural sciences.

Mark Willems entitled his doctoral thesis in graph-theoretical study of semantics *Chemistry of Language* (Willems, 1993). Following his lead, Liu (2002) took up *The Chemistry of Chinese Language* as the title of his thesis. Both dissertations expand the research in "knowledge graphs" initiated by Cornelis Hoede at the University of Twente, the Netherlands. His compact lecture, available on the Web, is an excellent introduction into graph theory, knowledge graphs, and the larger picture (Hoede, 2003).

There are countless ways of representing knowledge, meaning, and logic of an expression. The vast literature on ways and uses of representing a sentence as a graph makes me think that there is simply no way to prefer one to another and more can be invented. It also tells me that we do not know what thought is but are afraid to admit it.

The concept of chemical structure has been ingrained into chemical thinking for almost 150 years. Generically, it means the following (wording could slightly differ):

Properties of a chemical compound depend on the **properties of the constituting atoms and the way they are put together.**

Compare that with:

The meaning of an utterance is some function of **the meanings of parts of that utterance and the way they are put together** (Kirby and Christiansen, 2003).

Quoting the Nobel Lecture of one of the founders of modern chemistry (Pauling, 1954):

In 1861 Butlerov, making use for the first time of the term "chemical structure", stated clearly that the properties of a compound are determined by its molecular structure, and reflect the way in which atoms are bonded to one another in the molecules of the compound.

Probably, much more scattered testimonies of affinity could be found, but the last one is strong enough.

I would add some personal observations. Most linguists and chemists deal with material evidence: speech and text are as completely observable as molecular structure—at least by the modern means of analysis. Theory in both areas, however, deals with structures beyond observation. Finally, the material evidence is enormous in size, but by no means infinite, whatever the linguists may say in their strange and pervasive obsession with infinity. See, for example, Studderd-Kennedy and Goldstein (2003).

4. NOAM CHOMSKY AND JOSEPH GREENBERG

Linguists sometimes speak the language of chemistry without even realizing it. But can we speak the language of linguistics?

In Molier's *Le bourgeois gentilhomme*, a play with linguistic connections, Monsieur Jourdain was surprised to learn that he had been speaking prose all his life. A shock—rather than surprise—may expect a natural scientist entering the world of formal linguistics explaining how we speak, prose or poetry.

What follows are some impressions of a skeptical ignoramus like myself, pumped up for the sake of performance but without any malice. I am just trying to prepare the soil of Pattern Theory for the seeds of both chemistry and linguistics.

The Rhyme and Reason by Juan Uriagereka (1998) was intended as an encyclopedia of formal linguistics for an outsider like myself. It reads as sometimes irritating but irresistible spoof of Dante's *Inferno*. Any delusions about the language one has been mindlessly speaking since childhood are supposed to be cured by elephantine doses of proprietary purgative terminology, strangely duplicating the most common words of our daily usage. Here is an example I found in References (Uriagereka, 1998, p.636):

Uriagereka, J. 1994. Government restrictions on Basque movements. In *Syntactic theory and Basque syntax*, edited by J.I.Hualde and J. Ortiz de Urbina. University of the Basque

When (in 2003) I search the Web for all the words in the title, over 2600 links pop up. When I subtract from the search linguistics and Uriagereka, the remaining 2450 links deal with political restrictions on Basque separatism in Spain.

To me, an ignoramus, the minimalism sounds as a maximalist version of Zen Buddhism with koans like "empty category is not empty." The giant book's layout leaves almost half of it just blank, making one wonder if this is a hint to solving the koan.

For an outsider it is hard to chase away the prophetic vision of Herman Hesse in his *The Glass Bead Game*, which, in my opinion, anticipated the postmodern world obsessed with repertoire and performance and indifferent to substance and spirit. A look at the original works of the formal school shows the state of permanent flux, debate, elimination, and invention, for which the *Magister Ludi* himself is not always responsible. Paraphrasing Goya, the sleep of context multiplies monsters.

The Chomskian linguistics looks to a chemist like a Gothic bestiary of creatures, one more bizarre than another, living in a haunted castle where the zombies of 𝕿𝖗𝖆𝖈𝖊 and 𝕰𝖒𝖕𝖙𝖞 𝕮𝖆𝖙𝖊𝖌𝖔𝖗𝖞 drag their ghostly existence. This world is not for the faint of heart: like in a horror movie, as soon as you, 𝕭𝖔𝖚𝖓𝖉 in 𝕮𝖍𝖆𝖎𝖓𝖘, have be𝕳𝖊𝖆𝖉ed the common 𝕸𝖊𝖗𝖌𝖊, the towering 𝖀𝖀𝖕𝖊𝖗𝖒𝖊𝖗𝖌𝖊 assails you from behind. Furthermore, after you have somehow tackled the midsize 𝕷𝖆𝖇𝖊𝖑, a peewee 𝕾𝖚𝖇𝖑𝖆𝖇𝖊𝖑 crawls into your pants, and behind the disabused 𝕻𝖗𝖔𝖏𝖊𝖈𝖙𝖎𝖔𝖓, the monster of 𝕾𝖚𝖕𝖊𝖗𝖕𝖗𝖔𝖏𝖊𝖈𝖙𝖎𝖔𝖓 𝕽𝖆𝖎𝖘𝖊s on its hind legs. But wait, there is also the bloodcurdling 𝕾𝖚𝖕𝖊𝖗𝖗𝖆𝖎𝖘𝖎𝖓𝖌, even before the end of Part One.

Well, I get some understanding, if not support, from an outstanding linguist:

> We all speak at least one [language] , that one we acquired without a lick of conscious effort, and most non-linguist, in the unlikely event that they opened a copy of Linguistic Inquiry or Natural Language and Linguistic Theory only to find stuff every bit as hard going as genetics or quantum mechanics, would in many cases react by saying 'What's all this nonsense about? Why are they making such a fuss about something that's perfectly simple and straightforward?' Bickerton (2003, p.77).

The growing area of language evolution and computation, however, carries a clear promise to create a new, lighter, funnier, but incomparably more complex and technically intricate post-Gothic world, shielded from a curious trespasser by large and unpublished computer codes where, by definition, nothing is either minimalist or made of atoms.

And yet the minimalist *Inferno* left a definite 𝔱𝔯𝔞𝔠𝔢 in my chemically pre-complexificated mind: there must be some deep (*ergo*, simple) truth in the entire approach.

The "naturalization" of atoms in chemistry and, later, physics, as well as genes in biology, was a result of processing numerical experimental data. It is the absence of numbers that makes the generative linguistics a nightmare for the natural scientist.

Nevertheless, knowing well that mathematics is not only about numbers, and responding as a chemist to the teasing call of my "sister science," I see more than just vague parallels between formal linguistics and chemistry. I have a feeling that Noam Chomsky saw and tried to solve a really monumental scientific problem much ahead of time. The right time, probably, has not yet arrived: the fifty year old artificial intelligence, while beating the grand masters at the chess, still generates ridiculous translations. I feel that Chomsky's appeal to chemical experience is justified.

It is true that the chemists went through a similar stage. The chemical concept of atom, as speculative as genes before molecular biology, had been introduced by John Dalton around 1808, but it took exactly one hundred years before Rutherford could put some physical flesh on the surreal invention of the mind. After that, about fifty years of theoretical maturation followed. For the remaining half century, chemists have quietly immersed into applications and their monetary rewards, now even sending ripples through the stock market with the dotcomoid nanotech.

There is another, more successful than AI, area with a similar childhood: the chemistry-based molecular biology. Formal geneticists were able to map the distribution of genes in chromosomes at the dawn of the twentieth century without a slightest idea of what either gene or chromosome actually was, simply by comparing the observable features of progeny with those of the parents and counting some numbers. The Chomskian linguistics is strikingly similar to the pre-DNA formal genetics (now called transmission genetics), in which an enthusiast can even find parallels between 𝔐𝔬𝔳𝔢 and

𝔐𝔢𝔯𝔤𝔢 of linguistics and the **crossing over** and **linkage** of transmission genetics. The dramatic (and probably lethal) difference is that formal linguistics is based on an idealized standard of language, which we (not having blood royalty in America) can never hear around. The formal genetics, on the contrary, was based on the study of observable **deviations** from the idealized standard biological type. This is how the tiny fruit fly worked for genetics when the frequencies of mutations were counted.

After applying the Ockham's razor—with the barbarism of an ignoramus—to the dense growth of formal linguistics, I (together with some linguists) see the main problem that formal linguistics deals with as the linearization of the non-linear, see Figure 4.1. This is like extruding the globe of dough through the spaghetti-making machine, for which you need to apply some effort.

There is the unobservable "message" (my copout, to avoid terms like meaning, content, and knowledge) and there is its strictly linear verbal expression. The message is not necessarily linear but the expression is. How does the mind do it? For example, how could Shakespeare put on paper his mental images of the story of Hamlet, with its intricate relations, moves, and physical collisions?
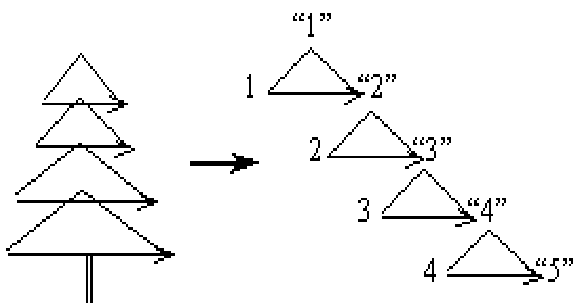
The simplest part of the problem is to portray the development of the story in time. This is done regardless of the language, by arranging consecutively the descriptions of events. The chain consists of phrases. For example, a primitive Hamletesque story could go like this:

Figure 4.1. Linearizing a tree

Enter Ug and Og. Og takes Ug's meat. Ug ponders whether to kill Og. Ug kills Og. Ug dies of spoiled meat.

The drama of the story is in the unexpected consequences. The linguistic drama starts each time when a single link of the diachronic chain must be constructed. There could be several alternatives and which one will be realized creates the suspense in the mind of the cave Shakespeare.

Within the formal framework, it is useless to ask, as a chemist would, what real process stands behind the schema, what its driving force is, how long it would take, how it starts, proceeds, and ends, as well as what the word and its boundaries are, and how we form the initial array of words. Nevertheless, the core of formal linguistics has the power of a well-formulated question which is half the solution.

It is hard to miss the similarity of the problem of linearization to the problem of projection in our perception of visual images: the complex 3D object turns into a flat image on the eye retina. For the humans who, like greyhounds, get most of information from vision, language seems to be an extension of vision in the sense that the same kind of topological many-to-one mapping has to be performed. Language can be called a collective vision. (What is TV? Corporate vision?).

I absolutely do not intend to criticize the formal grammar. Atoms were criticized, too, and, for that matter, what sound idea was not? Just the opposite, I would like to formulate the points of the formal theory that not only make sense but also appeal to a chemist who has survived 𝕾𝖚𝖕𝖊𝖗𝖕𝖗𝖔𝖏𝖊𝖈𝖙𝖎𝖔𝖓 and 𝕾𝖚𝖇𝖑𝖆𝖇𝖊𝖑.

1. **Binarity**. The basic relation between words/atoms is binary in X-bar and chemical bond. Not only there is an attraction between some words, but there is also a bond between groups of words, which tend to stay within the group, as a swarm of midgets (subjacency).

2. **Linearity**. The output of a certain process is linear. This means that a certain typically non-linear structure undergoes a structural transformation, as in a chemical reaction.

3. **Optimality**. The natural direction of the process is toward a minimization of a certain parameter, as in chemical reaction going toward in the state of equilibrium. This point is not exactly a part of the formal framework but it snugly fits the picture and is embraced by Chomsky.

On this positive note, let us now take a peek into the world of Joseph Greenberg who was also attacked and vilified in his time and for the same reason as Noam Chomsky (and may I add Charles Darwin?): his conclusions about evolution of languages were beyond a direct proof. After the sulfurous inferno, however, the world of Joseph Greenberg brings the revitalizing smell of the rhubarb pie to the nostrils of the natural scientist who has just been zenned by Ug, sorry, UG (Universal Grammar).

In the Semitic languages the verb tenses are marked by varying vowels, usually, around three constant letters of the root, as in Hebrew (read Hebrew characters right to left):

אני יכתב   אני כתב   אני כותב

ani iktov   ani katav   ani kotev

I will write   I wrote   I write

The root of the verb is ktv, כתב. The additional ambivalent letters י and ו are sufficient to distinguish between tenses in writing, without explicitly indicating how the words are pronounced, which could be done by diacritic signs. The signs, absent in the examples, are written in Hebrew and Arabic only when absolutely necessary. The last vowel of the verb between ת and ב is not hinted in any way. The **heard but invisible** vowels, therefore, define here the tense, which in writing is marked by adding some **visible but not heard** letters from a limited set.

Figure 4.2 presents the top left hand corner of Table 1 from Greenberg (1990, p.368). The table contains the frequencies of binary occurrences of the first and second consonants of three-letter roots of 3775 Arabic verbs. The work was stimulated by the previously well known fact that Semitic roots do not typically contain identical first and second consonant, while no such restriction exists regarding the second and third ones. Greenberg presents two more tables with correlations of II-III and I-III consonants, which exhausts the matter.

The numbers in the Figure 4.2 are quantitative measures of occurrences of initial pairs of root consonants and they can be converted into probabilities, although the list of all verbs is not a natural system for a natural scientist. It is like a list of all animals in a

forest without their actual numbers. Nevertheless, some conclusions can be drawn from the frequency of a combination of a long neck with long legs or hooves with horns.

What the Table shows is a Shakeaspearean play of affinity and animosity between consonants. The two identical initial consonants categorically refuse to stand side by side, as the arrow along the diagonal zeros indicates, but there are also definite areas of affinity, as around consonante n (No. 12). It could be seen from another Greenberg's table that the last two consonants do not mind being twins.

	II	ʔ	ḥ	ḫ	ʕ	x	ɣ	q
I	1	ʔ	0	3	2	0	3	0	3
	2	h	0	0	0	0	0	0	0
	3	ḥ	0	0	0	0	0	0	8
	4	ʕ	0	3	0	0	0	0	10
	5	x	0	0	0	0	0	0	0
	6	ɣ	0	1	0	0	0	0	0
	7	q	1	3	6	11	0	0	0
	8	k	3	5	4	5	0	0	0
	9	g	3	7	6	5	0	0	0
	10	r	7	9	7	10	4	7	13
	11	l	3	8	12	8	5	6	11
	12	n	6	14	10	12	10	11	23
	13	š	4	7	8	7	3	5	7
		ḍ	3	4	4	3	1	3	5
		s	5	8	12	8	9	1	.
			1	2	3	4	5	6	7

Figure 4.2   A corner of one of Greenberg's tables

Analysis of this kind was one of the first exercises of the newly mastered power of computers (Shannon, 1948). Shannon measured the frequencies of letter and word

pairs and synthesized artificial texts. Greenberg, however, did his work entirely by hand, as, by the way, Zipf did his. The frequency matrices of the type explored by Greenberg and Shannon are linguistic fingerprints. Like DNA analysis, they can be used for recognizing kinship between natural languages as well as styles of individual writers. Among other applications of frequency matrices are the fingerprints of musical styles.

As a non-linguist, I will use the single old paper of Joseph Greenberg, published in 1950, the remarkable time with all the planets in the most auspicious layout for great ideas and the computers in the prenatal fetal position, to illustrate a principle important for both chemistry and linguistics. Scores of modern works are running off the computer mill today.

Shannon's game of generating texts from the probabilities of letter and word pairs still fascinates the public and specialists. The reader who copies the next paragraph and runs it through *Shannonizer* (Senyak, WWW) will easily see why.

Greenberg's tables are something a chemist can fully identify with: the numbers characterise the strength of the bond between morphological atoms, or, metaphorically speaking, of their attraction to each other. In everyday parlance this is what we mean by chemistry—good or bad—of human relations. This is also what we mean by chemistry proper, in which only aggregates of atoms with good attraction are stable and those with repulsion or indifference are ephemeral or non-existent.

The high school chemistry is almost exclusively "good" one. Only when a student goes to a college, he or she discovers that modern chemistry started with the recognition and intense study of structures with bad internal chemistry. They are short-living, rarely observable, capricious, unruly, and cannot be kept in a jar on the shelf, but chemical transformations go through such ghostly chimeras and cannot be explained without them.

This may be an entirely poetic vision, but I associate the **transition states**, as the chimeras are called, with fleeting **thoughts**. Indeed, what is captured and fixed into a stable grammatic form as a generic "thought," like "All people are born equal," is only the stable result of the thought **process**. I suspect that thouhts are unobservable for the same reason most chemical transition states are: they are, as the old chemists used to say, *in statu nascendi*, in the moment of birth.

If thought is a **transition state** and the verbal output is the **final state**, the inavoidable question is what the **initial state** of thought is. Questions of this kind, comprising the entire triad **source—thought—language**, have been generating the Himalayas of literature through millennia. I would prefer to leave the majectic background behind the merciful fog and present the larger picture as a few snapshots taken through a chemical lens. As language is a linearization of an image and image is a projection of an object, my notes are only a projection of both into the mind of a chemist, extruded as a text output, not without the help of a computer.

The main para-linguistic problem that arises in connection with matrices and tables is how to send the entire Greenberg table over the telephone. Same problem exists in chemistry: how to send a structure over the telephone? Without the Internet, of course.

5. CHINESE AND CHEMICALESE

I believe that an organic chemist thinks not in words but in structures. They form the sustaining environment for the chemist, like the animals for the hunter or the plants for the gatherer on the verge of a protolanguage. I definitely see structural formulas or their fragments when I think about familiar compounds. They are the **source** of my **thoughts** about them and they can be shared.

Some of chemical thoughts are visual images of relatively simple structures, like benzene, while others are more complex but seen in all details, especially if the chemist works with them daily. Chemists can mentally operate the structures, similarly to the blindfolded chess player who can remember and operate positions on the board. For chemists, the relation between the object and its representation is exceptionally straightforward, **iconic** (Peirce, 1992).

The development of a separate language in chemistry by the end of the nineteenth century, when the first comprehensive handbook appeared, was driven not so much by the need to communicate with other chemists, which could be better done on paper or blackboard, but, as I believe, from the need to store and search the stressfully mushrooming literature. A similar problem arose millennia ago in China.

The Chinese characters (*Kangxi* in China and *Kanzi* in Japan) consist of standard elements arranged in a certain order. They come from ancient pictograms, later combined into ideograms and supplied by phonetic elements.
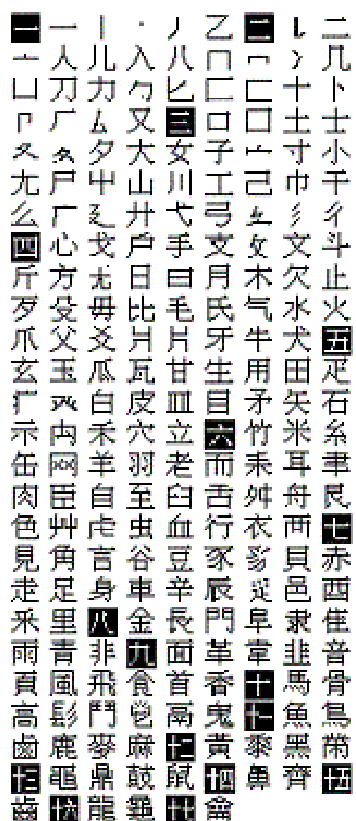
The problem with pictograms and ideograms is that there is no preferred or obvious way to order a very large number of them. In China the stress that such situation inflicted on the business of making dictionaries was resolved around 140 AD by using components called radicals (*bushou*) as "quasi-alphabetic" and partially semantic markers, while arranging them in order of increasing number of brush strokes (a combination of right and down strokes is considered a single stroke). Figure 5.1 (from Zhongwen, WWW) shows the Chinese *bushou* counterpart of alphabet. The white-on-black characters are Chinese numerals from 1 to 17 signifying the number of strokes. The modern system of 214 *bushou* radicals was introduced in the seventeenth century. The Japanese writing system uses the Chinese characters, limited in number by the script reform of 1946 to 1850 characters (now 1942), but is significantly different as a whole because of a separate indigenous phonetic system, not to mention the grammar.

Figure 5.1 Chinese *bushou*

Figure 5.2 shows a smaller part of all ideograms with the radical *ren*, person (top left-hand corner), which is used also as a whole character. As a radical, it takes different forms depending on left or top position. The arrows point to: 1. Feelings (two + person). 2. Partner (person + fire, i.e., at the same fire). 3. Position (person + standing).

The chemical nomenclature, amazingly, uses the same principles of dividing the structures into classes by semantic features, and arranging them within the class by their components and size (IUPAC, 1993). Examples are the radical "person" in Chinese and the class "heterocycle" in chemistry, i.e., cycles built with participation of atoms other

than carbon. If there is more than one cycle, the compounds are ordered in search systems according to the number of cycles, their size, number of non-carbon atoms, and their types, which is all much more complex than the Chinese characters.

Like Kangxi, chemistry uses numerous abbreviations and shortcuts. Now and then, both Chinese characters and nomenclature are ambiguous or just difficult for a novice. For a curious example, see Appendix, 1.
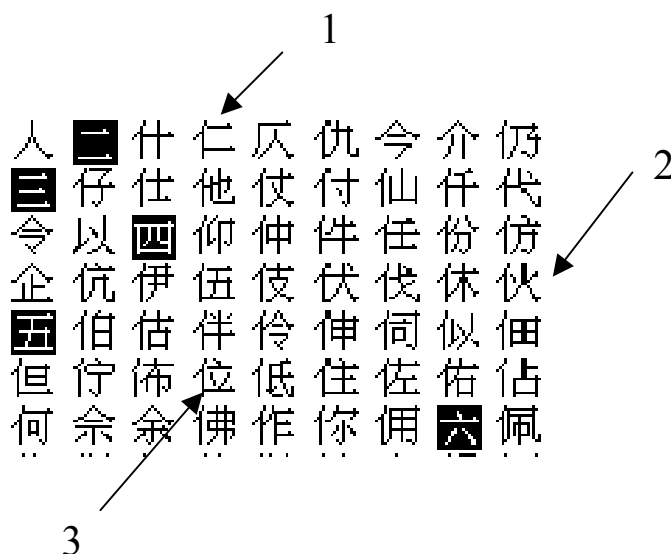

Figure 5.2. Some Chinese characters with the radical **ren**. Inverted characters indicate the number of additional strokes. Explanation of arrows, see text.

The chemical notation allows multiple versions of portraying the same structure, which can be transformed into each other by simple operations, such as permutation, flipping, stretching, rotation, etc. What remains invariant is the topology of the molecule. It can be represented by a molecular matrix, for example, in the case of formaldehyde with indexed hydrogen atoms, Figure 5.3.
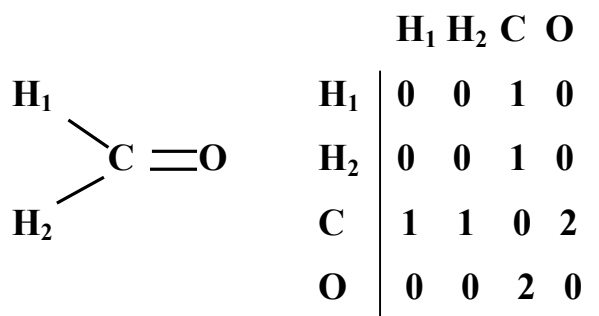
Since the hydrogen atoms in formaldehyde are indistinguishable, the indexes are redundant, which is generally not the case, for example, in methanol with two types of hydrogen atoms, Figure 5.4.

	H_1	H_2	C	O
H_1	0	0	1	0
H_2	0	0	1	0
C	1	1	0	2
O	0	0	2	0

Figure 5.3 Formaldehyde and its molecular matrix

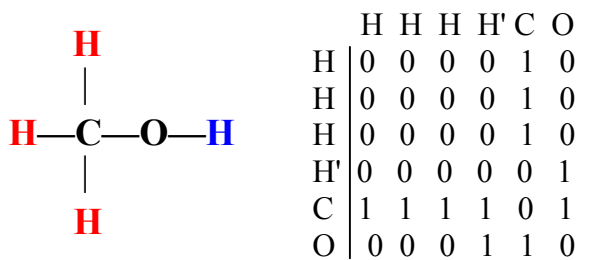|	H	H	H	H'	C	O
H	0	0	0	0	1	0
H	0	0	0	0	1	0
H	0	0	0	0	1	0
H'	0	0	0	0	0	1
C	1	1	1	1	0	1
O	0	0	0	1	1	0

Figure 5.4 Methanol and its matrix

The nonlinear pictographic and ideographic character of Chinese and chemical nomenclature separates the sound from the sign. Both systems of language make the message visually understandable, respectively, for all chemists regardless of language and all literate Chinese speakers regardless of very significant dialect differences. The spoken foreign Chemicalese, however, can be incomprehensible even for a chemist. Thus, in Russian, "benzin" means gasoline, while benzene is called "benzol" (from German).

The musical notation can also be viewed as a quasi-hieroglyphic system where the signs represent the sound and its duration by separate means. But the notes clearly display a property that is much less pronounced in other languages: music, like natural speech in real circumstances, is mostly continuous and it consists of a hierarchy of overlapping segments, including very large ones (unless it is minimalist). For music as language, see Jackendoff (1987, Chapter 11).

The matrix representation of chemical structure is not something that could be found in a textbook: the chemists do not need it, unless they are developing chemical software. The following is yet another representation of the structure of formaldehyde, I believe, not to be found anywhere, although the principle is known in programming as sparse matrix.


C—H$_1$, C—H$_2$, C==O
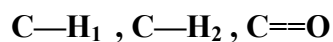
This representation lists the pairs of connected atoms in no particular order and contains the same information as structural and matrix representations. All three can be reconstructed from each other.

The list (sparse matrix) for methanol is: 3 (C—H'), C—O , O—H' .

What neither the matrix nor the list have is ***iconicity*** (Peirce, 1992), a similarity to the molecule itself. Note that similarity is a cardinal notion in Pattern Theory and can be formulated in exact terms. What the list has, however, is linearity and spontaneity. The

latter means a relatively high entropy and **low effort** resulting from a minimum of rules required to assemble the list. The list of binary connections is, paradoxically, a random way to represent order. From a sufficiently long random list with repetitions and reversals of the doublets $(C—H_1 \leftrightarrow H_1—C)$ a complete representation can be easily derived by eliminating redundancy. Thus, the list like this:

$$C—H_1 , H_1—C , C—H_2 , C=O, H_1—C , O=C , C—H_1 , C=O !!!$$
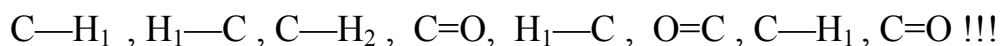
reminds a fresh on the spot account of a witness of a horrible accident to a policemen. The agitated witness is repeating and varying the same phrases while the policeman is jotting some concise notes down. The excitement of the witness creates an internal noise and repetition is a patented way to get the information through a noisy channel.

Our cave-dwelling ancestors are unreasonably portrayed in the movies as unkempt and dirty, which can hardly be seen even among monkeys, but we definitely could not expect from them the eloquence of Demosthenes. The eloquence of a survivor of a plane crash is more probable—the crash of old good animal life.

The most important thing for us is that a highly randomized linear sequence of pairs is sufficient to code a non-linear chemical formula. To preserve the binary relations in whatsoever order is all that is needed. To think in Pavlovian binary relations "ringing"—"food" is something any mammal is good at.
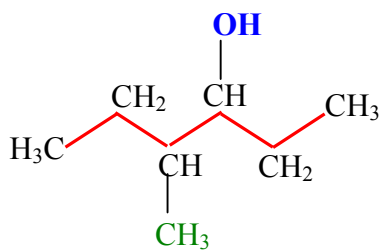
4-methyl-3-hexanol

The language of chemical nomenclature linearizes the formula of the substance on the left by deriving its name from the longest six-carbon chain, here hexan, C_6 . The atoms of the chain are numbered and the groups that substitute for atoms of hydrogen are listed by their "morphemes" *methyl* and *ol* in a prescribed order along with their positions. The components of carbon chains, such as methyl, are called radicals, while –*ol* is an example of a function (hydroxyl). It is more complicated if the substance

contains cycles and non-carbon atoms in long chains. Chemical nomenclature widely uses parentheses and brackets as syntactic means of grouping, nesting, and long-range connections.

Both written Chinese sentence and chemical name are perfectly linear in appearance. There is as little similarity between a structure and its chemical name as between a live person and the sound **ren** that signifies it, but the structure can be reconstructed from the name. **The chemical name is a linearized molecular matrix**. The Chemicalese does it by using the complicated grammar of chemical nomenclature. This is the simplest universal grammar, truly minimalist, and seen in all its nakedness. Still, it does not say anything about protolanguage.

In spite of my declaration that chemists think in structures, the chemical thought is not observable, either. A chemist may associate with the word "formaldehyde" its gross formula CH_2O, structural formula, the image of the bottle, or the smell, depending on context and circumstances. In a discourse, ambiguity, never completely suppressed, is minimal or easily resolvable by a question.

The chemist has all the reasons to believe that the chemical formula is not just a symbolic representation of the molecule but its model or a projection of a model that preserves important properties of the original, namely, its topology and in some cases its metrics. What the chemist keeps in his mind—and the formal linguist brushes off—are the physical properties of atoms and "atoms" and bonds between them.

To summarize, the function of language, so clearly demonstrated by chemical nomenclature, is first and foremost linearization. It turns the **source** (configuration, event, situation) into a string. In the rather rare chemical case, the source, **most probably** (but you never know), differs very little from its invisible mental representation.

The natural for a chemist idea that **the haphazard list of bonds as relates to the rules of chemical nomenclature as protolanguage relates to language** is the central proposition of this paper. There have been currents in linguistics, as well as in the collective painting of a larger picture, flowing in the same direction and branching into a wide delta.

Atoms are real, so are Og, Ug, cave, and a piece of meat, bur what is verb? Can we put a chemical finger on give or take?

6. RENÉ THOM AND IMAGES OF CHANGE

The approach to the language through the main entrance from the Broadway of reality counts solid and eloquent literature, including Charles Sanders Peirce (1839-1914; Peirce, 1992), Lakoff and Johnson (1980), Johnson (1987), Harnad (1990), Cangelosi *et al* in Cangelosi and Parisi (2002). Fodor (1976) developed an essentially structuralist approach to thought without the interface with observability, which is legitimate, but his Mentalese is not spoken here because it is not audible.

The entire topic of the interface between reality and the mind is so heavily burdened by centuries of philosophic discussions, starting from Plato's cave, that it is best not be touched by a chemist. But the topic is alive and exciting, especially, in view of some recent postmodernist attempts—uncharacteristically serious—to unite sciences and humanities. See for example, De Landa (1998).

The opportunity to bridge the "embodiment" linguistics, generative linguistics, and natural sciences was missed when the René Thom published his *Structural Stability and Morphogenesis* in 1972 (Thom, 1975). Her refers (Thom,1975, p. 116) to

generative grammar and draws a parallel between linguistics and biology, both having a hierarchy of morphological levels.
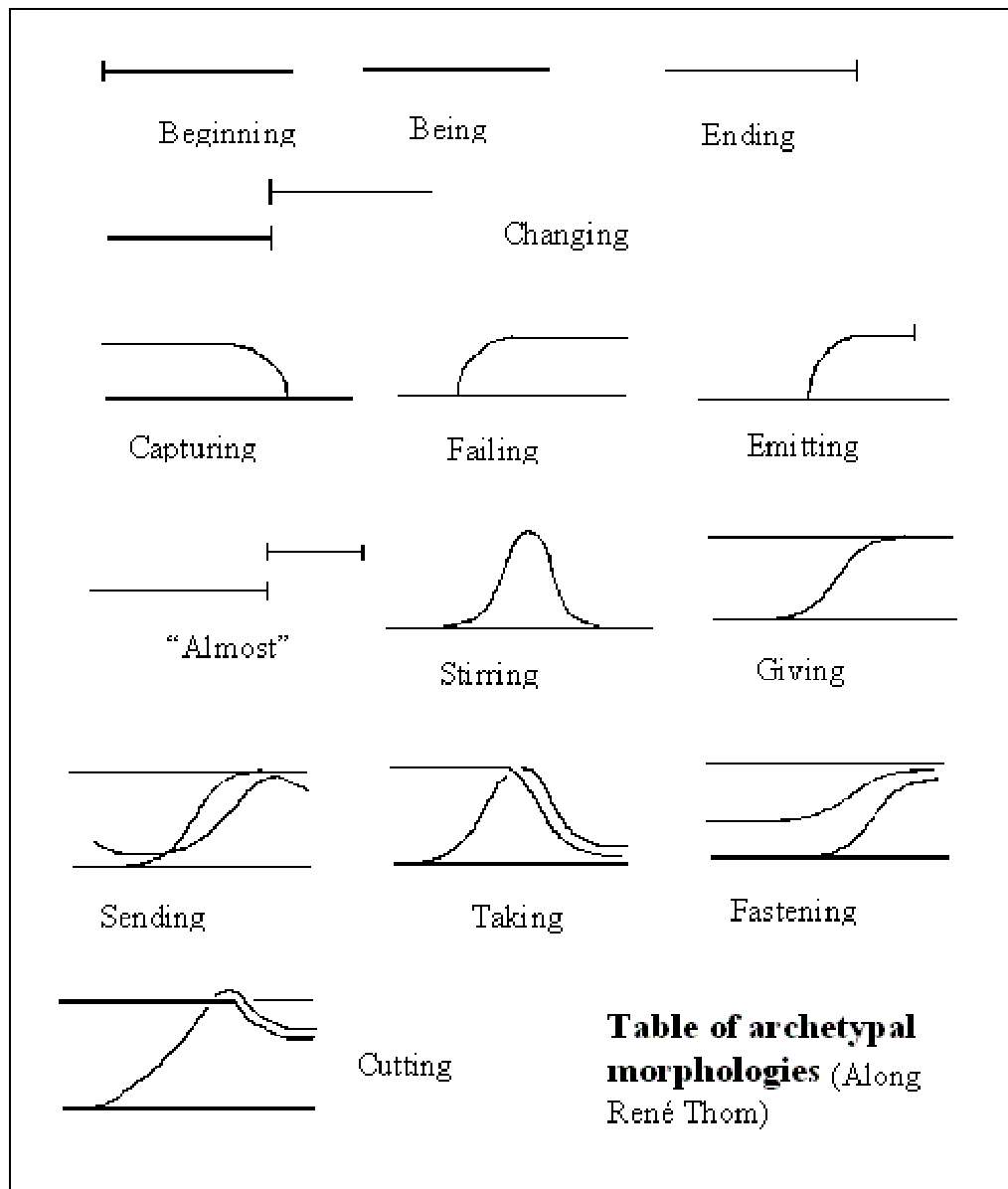

Figure 6.1 Archetypal morphologies of René Thom

The name of Rene Thom (1923-2002) is practically absent from linguistic literature. He was a French mathematician who presented a very general theory and typology of sharp **discontinuous** changes in the evolution of forms. His book was enthusiastically received but rather quickly set aside, probably, as Thom noted, because it did not lead to anything calculable. Besides, it was sometimes easier to see his intent than to follow his thought. Sadly, Thom is fading from scientific memory even faster than Zipf who remains at least anchored there by Zipf law and his "eccentricity."

 David Lightfoot comes closer than anybody to the concept of transition state in linguistics. He mentions Thom in his detailed and hands-on linguistic book (Lightfoot, 1999), which is unique in its thorough attention to the science of complexity, the exactness of the large picture, and the metaphoric power.

Figure 6.1 illustrates Thom's typology of "catastrophic," i.e., abrupt changes. Note, that almost each of them carries a name of a verb, including give. In my chemical interpretation, Thom distinguishes between three stages: stable and prolonged initial and final states, not pinpointed on the time axis, and a short transition between them.

Thom's types may seem nothing but modern ideograms that could be used in an artificial script similar to Chinese. The cave pictures of the hunt, where some animals are chased and other are lying dead or wounded, are, probably, the first pictograms of verbs right at the fork road where language is about to split from art.

The minimalist ideograms of Thom can be compared with some of those suggested by linguists, Figure 6.2.

The *image schemata* of Mark Johnson and George Lakoff (Lakoff and Johnson, 1980, Johnson, 1987) are less abstract than the archetypes of Thom. While the latter clearly distinguish between the tree diachronic stages of any change, the former appear to be just ideograms, like the Chinese characters in Figure 6.3.

Thom's "chreods" capture an important property of transition dynamics and possess some iconicity, while the static characters are purely symbolic and are interchangeable. According to Thom, chreod ("necessary path," the term borrowed from Waddington, 1957) is a stable configuration separated from another configuration by a catastrophic change.

Let us put ourselves in the position of the first humans who have few abstract ideas. They see a piece of meat in the hands of Ug. After a while it goes to the hands of Og. Having witnessed this process, how can we communicate it, for example, as gossip (Dunbar, 1996/98)?

> What is the machinery of language trying to accomplish? The system appears to have been put together to encode prepositional information—who did what to whom, what is true of what, when, where, and why—into a signal that can be conveyed from one person to another (Pinker, 2003, p. 27).
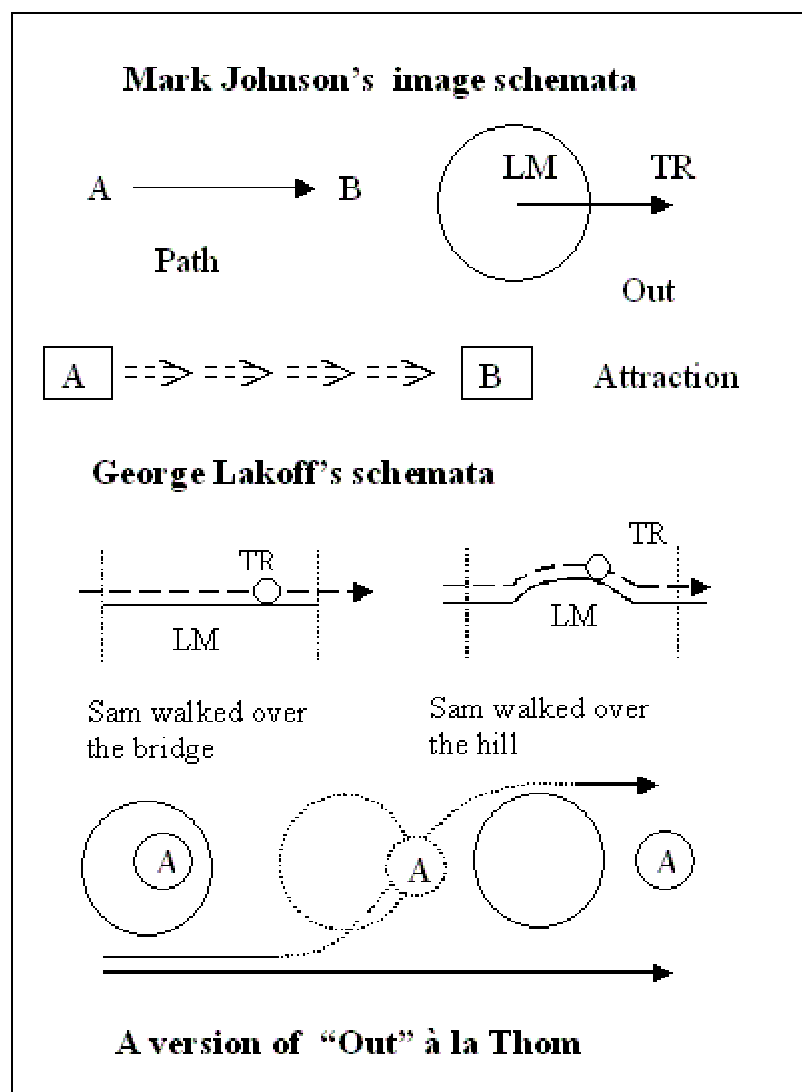

Figure 6.2 Some modern ideograms

In the primitive life who does what to whom and what the alternatives are is of an utmost importance.

行 , xíng, an ideogram for **walk**, "step and stop" or "left step and right step."

出 , chū, a pictogram of sprouting plant used as an ideogram for **out**.

吸 引 力 , xī yǐn lì , **attraction** with components: 吸 mouth, suck, inhale (left side), hand grabbing a person (right side), 引 bow and string, pull , 力 tendon, force.
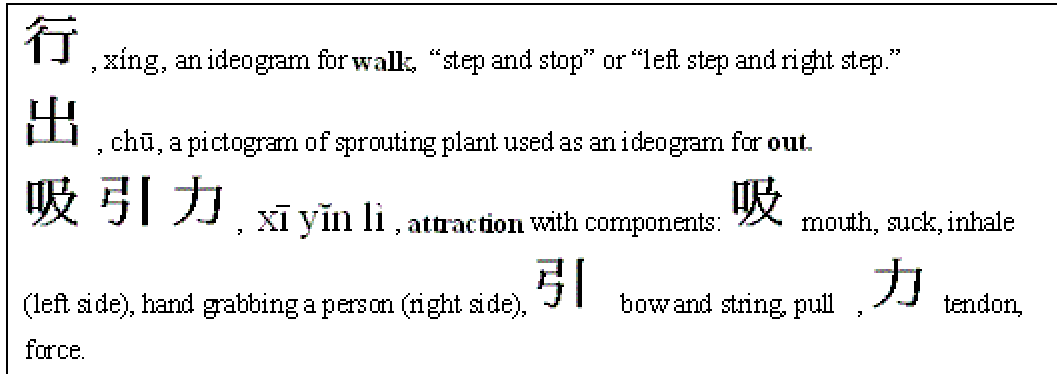
Figure 6.3 Ideographic composition in Chinese writing

Among the enormous variety of attempts to represent graphically what I call external sources of subsequent thoughts, semantic networks (Quillian, 1968) have been the oldest approach in Artificial Intelligence. The arbitrariness of knowledge representation, which reminds to a chemist a relative arbitrariness of writing 2D formulas of 3D molecules, is the reason for the never-ending flow of such graphs. Here is another example.

Cornelis Hoede (Hoede, 2003) and his group build very attractive knowledge graphs as, for example:

$$PLUTO \xrightarrow{\quad EQU \quad} \square \xleftarrow{\quad ALI \quad} DOG$$

The square, called token, is "something" identified (EQU) as PLUTO and similar to DOG: "something like a dog equal to Pluto"

The number of. binary relation types is limited to eight:

EQU : Identity   CAU : Causality
SUB : Inclusional part-ofness   ORD : Ordering
ALI : Alikeness   PAR : Attribution
DIS : Disparateness   SKO : Informational dependency

And yet the simple "Mary and Mike married" requires four tokens and nine relations for representation, Figure 6.4.
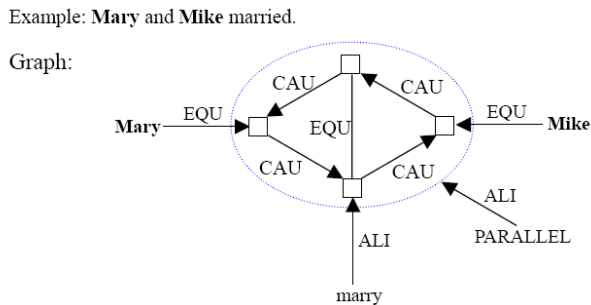

Example: **Mary** and **Mike** married.

Figure 6.4 Example of knowledge representation. From Hoede (2003)

Sowa (2000) provides an excellent guide in this world created by the drive to overcome the suffocating linearity of our language. We see in the postmodern art a similar drive to overcome, in the form of an installation, the Euclidian restrictions on the tangible classical form and its synchronicity. On the contrary, language itself started with overcoming the overwhelming connectedness of the world and passing it through the bottleneck of speech in segments.

The obsession with graphs, shared by chemists, has an underlying agenda in linguistics: to penetrate into the structure of thought. If we take up the vow of abstention from painting the images of thought, the graphs like those in this chapter fall in the category of **images**, which is a term of Pattern Theory. Image is the **observable** configuration generated by the object in the sensory anteroom of an organism, robot, or by an imaging instrument. The actual configuration of the object, which I call here the source, can be to some extent reconstructed from images, as it is done in CAT scans, Mars rovers, or the intelligence reports based on numerous and only partly reliable sources. The chemists do it from instrumental and analytical data.

Speaking half-metaphorically, if it is about medicine, in the **image area** between the source and the thought, the skeleton of the source is cleaned from flesh and prepared to be taken apart and arranged as a lineup of bones for transportation, as an unearthed skeleton of a dinosaur, but this is all we can say about it. Thus, a series of CAT scan

images results in the verbal linear description and diagnosis (intelligence report), with most of the computerized flesh retired to the archive or discarded.

The process of protolanguage generation, from the point of view of a chemist, displays between the external source and the output, passing the stage of **image** (projection of the source), **thought** (invisible), and **speech** (observable):

Source → image → thought → output

The mathematical language of Pattern Theory is the *lingua franca* for all three stages. PT also offers its own approach to configurations in the mind.

The General Mind Model (GOLEM) was suggested by Ulf Grenander (2003) as a natural application of Pattern Theory (Grenander 1976-1995) to systems of high complexity

The internal output of GOLEM, which is a state of GOLEM's mind, is a configuration called *idea*. It consists of atomic generators connected in a certain order, is limited in size, and is preceded and followed in time by other ideas. The idea can be spontaneous or induced by an external stimulus. It is selected along a probability distribution for generators and their connections.
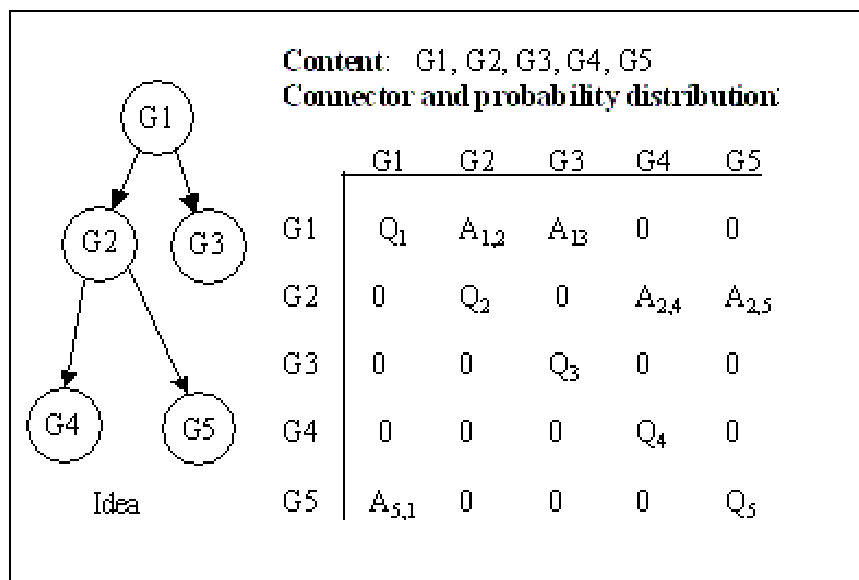
Content: G1, G2, G3, G4, G5
Connector and probability distribution:

	G1	G2	G3	G4	G5
G1	Q_1	$A_{1,2}$	$A_{1,3}$	0	0
G2	0	Q_2	0	$A_{2,4}$	$A_{2,5}$
G3	0	0	Q_3	0	0
G4	0	0	0	Q_4	0
G5	$A_{5,1}$	0	0	0	Q_5

Figure 6.5 . Configuration of an idea, along Grenander (2003), modified

Figure 6.5, modified from Grenander (2003) shows a typical idea of GOLEM in a matrix form. It is a list of all generators in the ***content*** of the idea, together with the ***connector*** graph. We can expect, that in a further evolution of GOLEM, this representation will persist.

The nature of generators here is irrelevant. They can be physical objects, ideas, emotions, memories, instincts, and all atomic components of a *state* of the system, in accordance with the *compositional* principle of PT. In linguistics, the generators are words or their roots, no other question asked.

The probability distributions Q and A, from which the idea is stochastically selected, are calculated from the previous Q and A , as well as from the external input, and are subjected to dissipation in the form of forgetting. GOLEM's mind, therefore, is a thermodynamically open system. While Q gives the probabilities of generators for content selection, A gives the acceptor functions (affinities of generators toward each other) from which the probabilities of bond couples in the connector are calculated.

GOLEM's **idea**, therefore, is a quartet <CONTENT, CONNECTOR, A, Q>. In the case of mental activity, probability distribution Q , in very general terms, can be compared with the priorities of an agenda: some items are more urgent than others. To use another metaphor, the state of mind is a pandemonium of generators, most of which are silent, and some of those few that are heard are louder than others. Accordingly, A reflects the strength of the bonds between generators and, therefore, the relevance and consistency of the content.

In the current Version 3.1 of GOLEM the output is a graph. Its internal precursor is **connector** in the form of a sparse matrix of the connector graph, which is essentially the list of bond couples between the generators selected into the **content**. It reflects the content and topology of the idea but its interpretation can be confusing for generators of high arity, as well as for fragmentary ideas. This confusion, however, is an intrinsic property of thought before it has been verbalized. While all the other known to me graphic approaches to thought aim at expurgating any ambiguity, Grenander's GOLEM proudly carries the poster: TO ERR IS HUMAN.

7. CONFIURATIONS, PATTERNS, AND NEAN

The key to the evolution of protolanguage, language, society, and life itself is lost in an unobservable past. All we can say is that evolution starts with something extremely simple and accumulates complexity by simple steps.

In this chapter, pursuing the large picture, I am going to pass over the insurmountable stack of literature on evolution. I will firmly hold onto my chemical intuition and experience, as well as the major ideas of Pattern Theory, which for me has been a meta-chemistry. Along the way, the simplest and a simple, but not the simplest, mini-grammars for protolanguage will be described in a few lines.

Pattern Theory (PT) is a branch of mathematics that embodies a very **simple** and known since Democritus and his proponent Lucretius (1958) principle of generating **complexity**: complex structures are produced by **combining** simple elements.

Within the framework of PT (Grenander 1976-1995), both molecules and expressions, whether verbal, written, or pictorial, are **configurations** made of atom-like primitives, called **generators**, with potential **bonds** capable of locking into actual **bond couples**. Each generator possesses a certain bond structure (number, label, orientation, sometimes, spatial order, and numerical properties) which in chemistry is usually associated with valences.

Since all differences between complex objects can be expressed in terms of elementary blocks and a particular order of their connections, PT is an incomplete but extremely general and powerful view of the world, especially outside the domain of

physics. It is a platform on which a more complete view of a great diversity of discrete and discontinuous objects can be built by adding other mathematical tools. The generative grammar, in which expressions are various combinations of various blocks, falls into the application basket of PT together with chemistry and countless other areas.
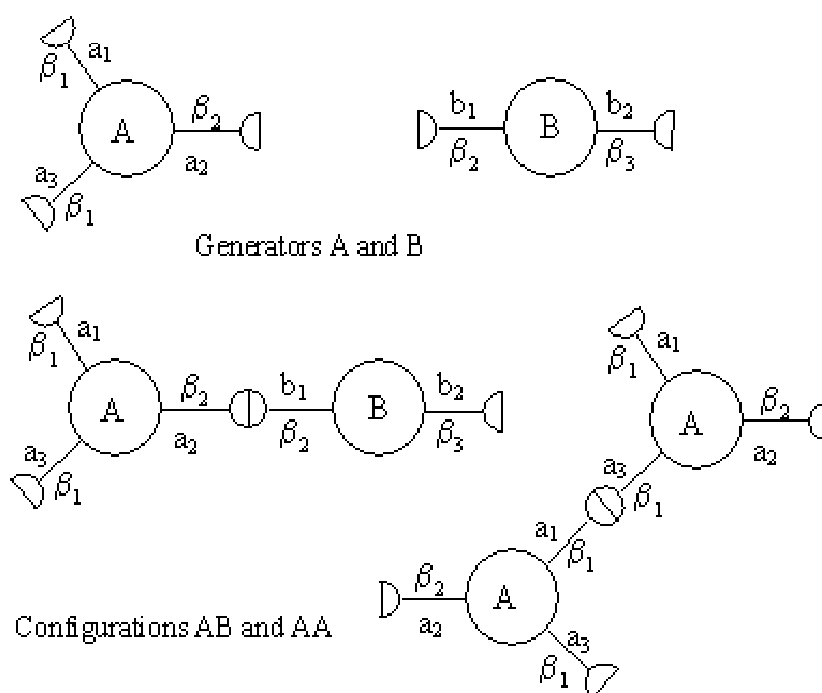

Figure 7.1 Generators and configurations

As an example, two generators from a **global** generator space **G** are identified as **A** and **B** in Figure 7.1. They have bond structures characterized by **local** bond coordinates a and b indexed both individually and by their generator. For example, bonds of generator **B** are indexed as b_1 and b_2. Bond values β (numerical, Boolean, or strings of characters) are attributed to the bonds. Given the generator space and a bond value relation ρ, a configuration space can be defined so that for each pair of bonds with values β and β', bond value relation ρ is either TRUE—and bonding is allowed—or FALSE—and bonding is forbidden. A configuration built "by the rules" is *regular*.

Appendix 2 presents some large real-life configurations that can make you speechless.

The realism and power of PT comes from its flexibility. All the rules can be relaxed or made stricter. Thus, the bond value relation ρ can be characterized by a real number between 0 and 1, which is neither TRUE nor FALSE, but just a probability of the bond. For example, the Greenberg tables—the example I will exploit *ad nauseam*—describe a space of regular bonds between the letters of the Arabic roots in various positions. They also fuzzily divide the entire space of possible letter doublets into regular and irregular. Thus, two identical letters in positions I and II are highly irregular.

The regular Arabic roots are triplets of consonants. This **pattern** is defined on the following generator space ($w2_1$ and $w2_2$ are bond coordinates of generator W2):
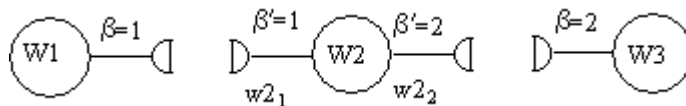
The regular configurations are those for which $\rho = $ TRUE if $\beta = \beta'$, which gives the single possible configuration W1-W2-W3:


This configuration is, in fact, a pattern, i.e., a class of configurations that can be obtained one from another through the same **similarity transformation**, in this case, **permutation** of consonants. It means that W1, W2, and W3 are not individual consonants, but the entire consonant alphabets. To push this idea to its limits, the combinatorial machine that Gulliver saw in Laputa, is nothing but a pattern of all possible strings of text, generated by permutation of letters.

This property of combinatorial systems is characterized in practically all general courses of linguistics as combinatorial infinity. In PT—as well as in chemistry—the "infinity" could be drastically cut by the properties of generators, as it is the case in

Semitic verb roots. Out of $29^3 = 24389$ possible triplets, only 3775 roots (15.5%) were listed in the source dictionary. The probabilities of the consonant pairs significantly differ.

The Greenberg tables describe an artificial object: the list of all Arabic verb roots. They have nothing to do with any actual conversation or text, in which only a small part of them could be used. The power of the approach is that the statistics of the list of verb roots is similar to a DNA analysis: it can be used for the kinship analysis between Semitic languages and dialects, both synchronically and diachronically.

From letters we switch to words, where a peculiar situation arises. Two-word generators (W1—W2) and (W2—W3) offer a natural choice of bond values identical with the constituent words. For bond relation ρ = TRUE if $\beta = \beta'$, the following doublet of composite generators is regular:
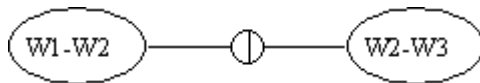
It directly translates into W1—W2—W2—W3, which turns into a triplet if the rule of **haplology** is applied. Haplology was in the focus of Zipf's theory of the least effort.

> Witness the phenomenon called haplology; when two similar syllables—they need not be identical—are adjacent, one may become permanently truncated. (Zipf, 1965, p.85) Abbreviation is then actually a short-cut; and moreover, since the stream of speech knows no other arrangement than that of time, *an abbreviation of speech is a short-cut in time* (ibid., p 284) .

Haplology—the word itself is ripe for haplology—results in two identical neighbors rewritten as one:

$$W1—W2—W2—W3 \rightarrow W1—W2—W3$$

If we are not in the mood for hairsplitting, this **single rewriting rule** already constitutes the smallest transformational **mini-grammar**. Unlike a grammar of a full

language, this grammar is strictly local and does not require any extended memory for generating an output. The speaker does not need to keep a large part of the entire image of the source in mind to express the thought, as it might be required for using a German verb with a separable prefix or for expressing a complex thought in any language. Only a recognition of two adjacent generators as different or identical is required, which is an effortless task for human mind and, probably, a foundation of all animal intelligence.

Clearly, there is an even simpler grammar, used by animals, in which the configuration is a single isolated generator, i.e., a sign, but this is no grammar at all, but a naked lexicon.

The first step in the evolution of grammar is the language with utterances connecting two generators.

A slogan in knowledge graph theory is that "Thinking is linking somethings" (*Sic*!).

Hoede (2003).

I call this language **Nean,** as a tribute to our cave-dwelling ancestors. I do not mean that the Neanderthals really spoke it.

To start with a binary relation is the most natural thing for a mathematician. The idea that protolanguage starts with a pair of connected words was convincingly discussed by mathematician Keith Devlin in his remarkable book (Devlin, 2000). He even gave a formal Chomskian structure for it as an elementary tree with unlabeled nodes (p.170):

This idea is completely in line with the general concept of protolanguage by Bickerton (1990, 1995, 2003), who had created the entire area, previously avoided by linguists, and whose work had an impact on Devlin (as well as on myself, of course). On my part, starting with Bickerton's idea, I am also strongly influenced by Pattern Theory and the concept of configuration in which both generators and bond couples contribute to stability. What I consider the main product of this influence is the assumption that Nean

is capable of expressing much more than a simple link: it expresses a source up to a significant complexity.

Nean directly translates binary links in the source into doublets in the utterance. Nean looks like a sequence of doublets. Of course, it can be peppered with singlets. A conversation on Nean looks like:

"ab a b a   bd ad d d ad !"
"d d d ab c ab c!"
Or:
"Ug big big big Ug big Ug ."
"Og Og hungry Og hungry hungry"

Nean may require some phonological means to mark word delimiters and stops, but the Ug-Og dialogue shows that it is not absolutely necessary. The change of content or a natural pause caused by an external event would do.

I claim, without any proof, that this is where the human language starts. A dog can say "I am hungry" or "I want to go for a walk" by whining at the table or sitting at the door because the dog's " I " is clear from the context. Probably, literature on animal communication can provide more complex messages, but I am not familiar with it.

On the example of this primitive grammar we can see the distinction between the formal grammar, in which an unlimited number of generators can be combined in a doublet, and the actual utterance generation, in which the output is dictated by the source. For a formal linguist, any doublet is as good as any other, while for the linguist who looks at the language generation as natural process, only the doublet that preserves the connectivity of the source is good, and the one that does it best **is** the best. We have a parameter for comparison—called *fitness* in Artificial Life—which is a condition of "naturalness" in terms of thermodynamics as well as **selection**.

To make the next step and launch the evolution of grammar, we take a random sequence of doublets and compress ("zipf") it, using the following set of rewriting rules:

ab + bd → abd

ab + ad → abd OR adb

ac + bc → abc OR bac

This "haploid Nean" is an example of a mini-grammar of a somewhat larger size.

Given the list of generators from generators space **G**, the probability of a configuration is defined by the probability of its bond couples. Instead of products of probabilities, sums of logarithms of probabilities ("energies") can be used, which makes a configuration additive regarding the "energies" of its bond couples, as it is typical for chemistry. The "energy" is a quantitative measure of what can be approximately characterized as the strength of the bond couple or the mutual affinity of two generators.

A series of mathematical and practical problems arise when we apply the concepts of probabilities and energies to non-physical systems such as language. I am not qualified to analyze this aspect of theory and can only outline the essence of the difficulty as well as a way out.

To use probabilities, we need a complete system of possible outcomes in a dynamic system, as well as a clear criterion of which outcomes are independent. This is a very tall order.

Dynamic systems are ensembles of large number of entities that randomly exchange a certain additive (conservative) value, statistically distributed over the entities. In general, probability and energy are convertible currencies in closed (i.e., non-existent in reality) dynamic systems: the lower the energy of the state of the system, the higher its probability. The lower the overall energy, the more stable the configuration.

This relation, completely transparent in statistical mechanics, which deals with idealized closed systems, is not so clear when we deal with real complex systems.

To use energy, we need a scale with a zero point (or the so-called partition function in statistical mechanics). We meet neither condition in the chemistry of even slightly complex molecules. When we discuss **evolving** systems of life and Artificial Life (ALife), the physical energy becomes meaningless. Instead, entropy multiplied by

statistical temperature could be used—if only we had really well-defined dynamical systems, which we have not.

The problem described here is very general. Probability theory is one of a few areas of mathematics where we can find debating sides (another is—no wonder—the foundations of mathematics). **Bayesian theory** , which is the subject of many sites on the Web and can be found in any, even low-level, course of probabilities, points to the way out of the incompleteness of our knowledge about real systems, which we can complement by additional observable data. An avalanche of works has been triggered by the community of the worshippers of Reverend Thomas Bayes (1702-1761), but it safely passed over the unsuspecting chemists.

In chemistry, the problem of the incompleteness of the data finds a very simple solution. Let us assume that only two structures in equilibrium are of interest, while the rest of the system remains the same. In this case the theory gives us not the absolute values of probabilities of two states, but their relative probabilities in equilibrium, and this is all we need for most practical purposes. The logarithm of the ratio of the probabilities is proportional to the difference of their energies. In general, when we do not know all possible alternatives, let us take only two. We can tell which one is more probable by comparing their energies, provided the rest of the circumstances the same. Equilibrium and evolution of a complex natural system, however, are two incompatible things.

From somebody with a chemical frame of mind, like myself, looking outside chemistry, Pattern Theory offers a general solution for any process: let us judge configurations by their regularity. Irregular configurations are strained, stressed, unstable, and short-living. The regular ones are stable, normal, uncontroversial, legitimate, and dominating. But what is irregular? It is what bends the rules. And what is regular? What conforms to the rules. Well, what is the origin of the rules? They are taken from the observations. Pattern Theory is an unusual kind of mathematics that restricts the freedom of imagination by the observable reality. There is no standard way to design generators, some models could be better than others, and a lot of human intuition must be involved in a development of an application. One might draw from the contraposition of regular and irregular structures an important conclusion that if a stable structure can transform into

another stable structure, it can do it, by definition, only through an unstable one. This is what chemistry is about. Otherwise all **relatively stable** structures on earth would immediately turn into the **most stable** equilibrium combination of the most stable individual structures and time would stop.

The above peek into PT was extremely limited and superficial. One has to look into the original work to appreciate the richness of the subject. One aspect of PT that I haven't even touched upon is the concept of the pattern as a **group** (in mathematical sense) of transformations that define a geometry of the configuration space. The second aspect is the idea of template, i.e., a representative, "typical" (Aha! Here comes the mysterious "typical") configuration of the pattern. It could be the most valuable treat for cognitive sciences: we all have mental templates for "cat," "dog," and "politics." What are they? The centers of gravity of a pattern.

It will do for my purpose to point to PT as a mathematical higher ground, strangely overlooked, for complex systems from which both linguistics and chemistry could be seen side by side as affectionate sisters.

Much more important, the higher ground of PT allows me to represent both the source and the output of speech generation in the same philosophically neutral mathematical terms: as configurations.

I will return to Nean in Chapter 12.

8. SOME RISKY IDEAS ABOUT MATHEMATICS AND LIFE

I must make here a risky digression—it could be safely skipped by the reader—addressed to all those interested in the big picture, including myself. This digression is intended to once again draw attention of all researchers of complex systems to PT. I venture to express an outsider's opinion (anyway, mathematics is a language, too) on the applicability of mathematics to complex real phenomena.

The fundamental *sine qua non* component of any mathematical system is the **theorem of closure**, which requires the set of terms to be well defined and closed to any uninvited visitor during reasoning. It can be traced back to Aristotle and is a condition of logical thinking. In other words, mathematics is poorly equipped to deal with the notion of **novelty**. Once the world of a mathematical system, i.e., its terms and axioms, has been created, whether in six or 666 days, there could be nothing new under its skies. This is definitely not the case under the skies of the planet Earth, as geological, biological, and social evolution—and anybody's personal life—clearly manifest. This is why mathematics, as well as physics, is not adroit enough to deal with evolution and the open set of terms. Real systems are inherently open to new and unanticipated terms, i.e., **new** generators of PT.

Mathematics, however, would not be mathematics if it failed to design a mathematical system for formalizing something. As it seems to me, the seeds of a mathematical theory of novelty were planted in a very little known segment of Bourbaki's monumental *Elements of Mathematics* about the **scale of sets** (Bourbaki,

1968, p.259), to which I would refer the curious reader. It describes how a **new** set of terms is formed from the **old** one by converting **new** combinations into **old** elements for the **next** combinatory step. It seems tricky, but is in fact simpler than it seems.

As far as Pattern Theory is concerned, its basic sets are remarkably open to novelty. PT is just born to be a mathematical tool of evolution because it can construct and implant **new** generators. It does it the same way the scientists express new and groundbreaking ideas by using the strings of old letters and inventing some new symbols and terms. Moreover, by using the convertible currencies of probability and energy, PT stores in its shed all the necessary tools to model the **realistic** evolution of complex systems. Being capable of distinguishing between what is possible and what is more probable than something else, PT straddles the fence between the camps of Chomsky and Greenberg.

Energy or quasi-energy, fitness, etc., can be considered a particular case of a more general concept of the natural world—**stability**, which is the other side of **regularity**.

In real life some scientists (especially, chemists) deal with differences in energy and others (especially, in computer simulation, economics, and cognitive sciences) deal with conditional probability because those are the **observable** source data.

It must be noted that although chemists freely convert relative energy into probabilities (in the form of equilibrium concentrations) and back, they are always aware of inherent uncertainty whether all possible outcomes are taken to account.

This is not the case in the closed system of three-letter roots, where all combinations are completely countable. The roots of Arabic are examples of linguistic molecules. Greenberg's results can be reformulated in terms of energy or affinity, i.e., attraction of root letters to each other—yet another reason for chemistry and linguistics to listen to the call of the common blood in the lush jungles of complexity. So are the giant molecules of musical compositions with the regularity consistently relaxing through the last two centuries. This is what, probably, keeps the drumbeat of pop in demand. Similarly, the statistics of marriage and divorce, to take a different angle, tell about the strength of marital bond in a social molecule and the position of the equilibrium between its two basic states of association and dissociation.

Big ideas have a curious property of staying unnoticed just because of their large size. Coming to the end of my digression, I do not want to miss an opportunity to point to one such idea. I have nothing to lose by mentioning a remarkable way to circumvent the problem of "physicality," which cuts a deep chasm between physical sciences and humanities, as well as between real and simulated systems. I realize that it may seem paradoxical and even pervert to both separated sides longing for an embrace.

Suggested quite casually by Prigogine and expressed in a very general form by Rosen (1991, 2001), the idea consists in regarding the physical, closed, and calculable world a particular case of a more general world where the vague, large, complex, hard to catch and itemize systems, like life, are a more general case than any physical model. They do not need an explanation through anything else. In a way, what Robert Rosen heretically suggested was a version of the Copernican revolution: the physical earthly knowledge is a satellite of a larger conceptual body of life sciences and evolutionary ideas in general. The "normal" reductionist angle of vision has been: life, society, and mind exist because there is some physical foundation for them. Robert Rosen reversed this relation: no, physics and chemistry exist because there is life on earth. No wonder, the first step that physics made in explaining life was to declare it impossible because of the improbability of a spontaneous assembly of DNA or RNA (Wigner, 1961).

For more about the scale of sets, chemistry, mind, and some illustrations, see Tarnopolsky (2003).

9. CHEMOLINGUISTRY: A CHIMERA

In this chapter I discuss a very general paradigm which does not seem to follow from any other and is one of the relatively recently discovered foundations of modern chemistry: the **theory of transition state**.

Honestly, a linguist has no need to know any down-to-earth chemistry. I will attempt to present here some almost unknown outside chemistry ideas as chimeras combining the properties of both chemistry and linguistics. Atoms will be labeled as meaningful words but will behave like atoms. We will use such familiar terms as atoms, energy, and even probability, in an intuitive and not in the strict physical and mathematical sense. We will try to find some criteria of checking the configurations for a measurable property. We will try to arrange our atoms into more or less realistic configurations. Coming back to Mark Baker's "*The Atoms of Language*," what if words were indeed atoms? Then both the expressions and their sources were molecules and the way from the latter to the former would go through a transition state, as it happens with molecules.

The term "transition state" is somewhat misleading. As all chemists know, it is not a state but a process. It is the state of flux.

The concept of transition state (Eyring and Polanyi,1931) is a very general concept of **dynamics**—the science of things in motion. It is used in **kinetics,** the part of

dynamics that studies the speed of the motion, but only if the motion is **discontinuous**, i.e., **catastrophic**, in terms of René Thom. In essence, transition state theory regards is as continuous. Can you imagine that?

> Transition state theory (TST), introduced by Eyring and Polanyi … in 1931 as an early attempt to determine absolute reaction rates, is too often considered the domain of the chemist or chemical physicist. However, the transition state (TS) is actually a general property of dynamical systems which involve an evolution from "reactants" to "products." Such processes include, but are by no means limited to, the ionization of atoms, the dissociation or re-action of molecules, and even the escape of an asteroid from its orbit (Jaffe et al, 2000).

Surely, any asteroid is a pending disaster.

The theory of transition state elucidates why things happen by explaining why they do not. It postulates that if one stable state of a system can turn into another, there is an ephemeral and unstable transition state between them. Its energy (stress) is higher and, therefore, stability is lower than that of both stable states. It is irregular, unlawful, and cannot be portrayed by common chemical formulas.

The transition state sets a **barrier** on the way of transformation. The lower the barrier, the more probable the transformation. This is why a sheet of paper does not ignite spontaneously and needs a burning match to push it over the barrier. For the same reason explosives can be safely stored and transported: they are protected by the barrier of the transition state on the way to the products of explosion. The detonator jolts the substance over the barrier.

Some general notes on transition state will be included into Chapter 11. Illustrations of mechanisms of chemical transformations, in both chemical and metaphorical terms, are given in Appendix 3.

As an introductory example, let us turn to the splendid collection of stressed life configurations left by Shakespeare. When we speak about a good chemistry between persons, we mean that their relation is stable. The initial relations between Othello and Desdemona have a negligible stress, low energy, as a chemist would say. The malicious energy of Iago initiates the explosive transition like a detonator. The highly stressed

situation dissipates its energy along the way to another stable but unfortunately tragic state. The short-living process, which is the core conflict of the tragedy is what a chemist can associate with transition state. The measurable parameter is stress synonymous with instability.

One may visualize atoms as ping-pong balls with a word on each, including symbols of chemical elements, written with a soft-tip marker. We attribute to them the ability to form bonds of various strength measured by the energy: the lower the energy, the stronger the bond and the more energy is needed to break it. Energy, therefore, is a measure of instability and, on a different scale, of improbability in a dynamical system. We will use also the terms **stress** or **tension** as the opposite of stability: the stressed configurations are those with high energy and low stability.

The words can themselves be labels of real or imaginary objects and events. The theory of meaning is one of the most confusing intellectual areas and we should better avoid definitions, preferring models instead.

Model 1

The model illustrates what a chemist could expect from three atoms labeled as words: Tom, Tim, and book. The name of the transformation is GIVE. Its detailed description is: Tom GIVE[s] book [to] Tim. We can see the extremely stable Tom, Tim, and the book, but no such thing as "s" or "to." But what is GIVE?

We observe the following states:

1. The initial state: Tom in contact with the book and Tim nearby.
2. The unstable and short transition state of transfer, which is GIVE.
3. The final state: Tim in contact with the book and Tom nearby.

Figure 9.1 shows the GIVE transformation as a chemist could depict it. It turns out that there are two possible **mechanisms** of the process, for which **A** and **B** are two possible transition states.

Since the initial and final states are relatively stable, the book can be in Tim's or Tom's possession indefinitely. In the act of giving, along mechanism **A**, Tom holds the book and offers it to Tim, who also touches it. For a short while the three atoms are locked in the ephemeral and unstable transition state, which is, in general, reversible. Mechanism **A** corresponds to a smooth continuous transfer. Either Tim or Tom, or both can change their minds concerning the transfer. But they can also fight for the book. Along mechanism **B**, Tom can simply leave the book on the table, after which Tim takes hold of it. The transition state for this version of GIVE is just three disconnected atoms.

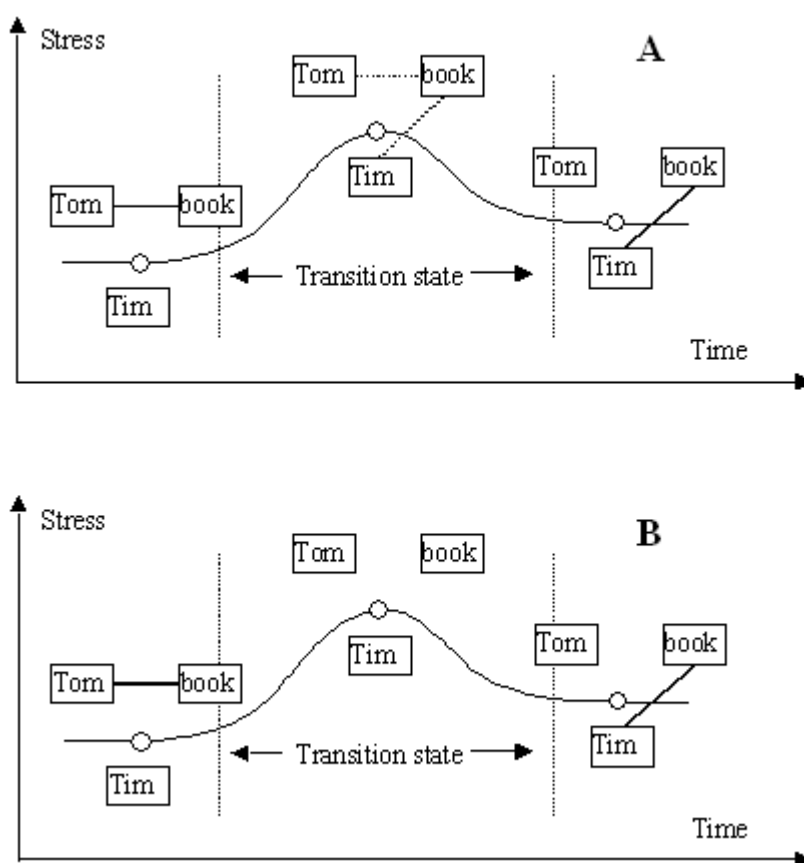**Figure 9.1   Two mechanisms of transformation GIVE from the point of view of chemistry.**

From the point of view of a chemist, GIVE is a name of a chemical—better to say "chimerical"—reaction between molecules Tom-book and Tim, resulting in the molecules Tom and Tim-book.

How fast this reaction could run depends on the energy (height) of the transition barrier. To chimerize more, we can even say that while Tim and Tom can do well without each other, a book without an owner is an irregular and therefore unstable configuration, while a book with two owners is also a potential source of conflict, unless there is a stable bond between the owners. For the forceful transfer, the transition state can be highly stressed and its outcome hardly predictable unless we know which stable state is less stressed, i.e., who is stronger, Tom or Tim.

The transition state starts with some really invisible processes in Tom's mind. To make any predictions about the direction and speed of the transformation, we have to evaluate an open set of other circumstances not in the slightest way reflected in Figure 9.1: relation between Tom and Tim, whether Tom has finished reading it, the availability of other books that Tim could use as substitute, the influence of other persons who have their opinion about Tom, Tim, and book, and even if the weather is better for reading than for sailing. This is exactly what a chemist does to study a chemical transformation and optimize its course by increasing the speed of the beneficial transformation and suppressing all the competitive directions.

Catalysis is the standard tool for speeding up one direction of structural change at the expense of the competitive ones. Catalysis can be compared with the role of a parent in a smooth transfer of a book from one child to another: the parent forms bonds with all the participants and decreases the stress of the transition state.

In fact, both types of transition in Figure 9.1 are known in the process of chemical transfer of an atom from one place to another, with some additional considerations. Once again, the transition states in chemistry, like thoughts in the brain, are not typically observable. They are, actually, chimeras of imagination, although some progress has been recently made in catching them. But they leave evidence which is observable and can be identified by the same detective methods as a crime without a witness, including forensic experiments (Appendix 3).

The above chimerical mechanisms are popularizations of what happens during the chemical transformation: it goes through a transition state the energy of which determines **how fast** the transformation will happen in the **short run**. Tim and Tom exist in single copies, but molecules are numerous. In the **long run**, since the passage through the transition barrier is reversible, there will be an equilibrium depending on whether the initial or final state has a lower energy. My illustration is intended to draw attention to the **kinetics** of transformation, i.e., the short-run outcome. Chemistry is interested most of all in the speed of the transformation. If there are more than one direction of a process, chemistry answers the question "what is going to happen when the process starts?" in the typically chemical manner:

In the short run, what happens is what can happen faster,

i.e., what goes through the least stressed transition state, and

in the long run it is the less stressed stable state.

The universality of this principle can be seen on any modern war: it is easy to start but hard to finish.

The aspect of speed (usually called **rate** in chemistry) has attracted little attention of philosophers used to the fleshless creations of the mind where everything that can be put into words is possible and it pops up in all minds with the same speed. Instead, endless debates about body and spirit, form and substance, semantics and syntax, meaning and sign, thought and utterance, all within the static framework of frozen structures, have been rolling through centuries, apparently, without any barriers of any kind.

Why do structural transformations happen at all?

In chemistry all particles naturally follow a distribution of energies (Maxwell distribution), so that most molecules have energy within a medium range. There are always particles with energy above that level and their collisions result in passing the transition barrier.

There is the fundamental thermodynamics of open systems, but no waterproof theory that could predict or explain their design. Examples of open systems are the

atmosphere with weather, life with evolution, society and language with history. It is not clear whether such theory is possible because any general theory would be silent regarding most interesting problems which are always concrete. Moreover, evolution of a large open system is an interplay of chance and necessity. The best way to observe transitions of large complex systems is to read history, which is a roster of long time stagnations interrupted by short term periods of turmoil that not necessarily lead to a new structure, as well as of long term drifts. The Imperial China, Roman Empire, French Revolution, and Industrial Revolution are classical examples. Lightfoot (1999) provided an excellent concise review of this entire area not only for linguists but also for chemists and anybody else who is not shielded from the world by a TV screen.

For a chemilinguist, the development of pidgin and creole languages (Bickerton, 1981) are examples of the resolution of an initial stressful situation of the Babel Tower type. The mass import of West European words into the old Russian language under Peter the Great (the pattern that repeats itself today, so that two alphabets, Latin and Cyrillic, could be used intermittingly), the mass invention of a new lexicon on the base of the old language in China, and the patriotic defense against imported words in Hungary two centuries ago parallel historic processes and are all as much answers to a stress as many scientific shifts and technological inventions. One can only guess whether the theory of punctuated equilibrium in biological evolution is just a spoof of human history.

The notion that an utterance has a short-time history and displays in a discrete time as a mechanism, i.e., a sequence of states (derivations) was to Chomsky's unquestionable credit. But the main revolutionary discovery made by George Zipf (Zipf, 1949, 1965), whose name is conspicuously absent from many recent works on linguistics, even those taking a comprehensive view of the field (from Jackendoff to Uriagereka, for example) was indigestible by static formalism.

If you call something "move," how fast it is? Zipf had no direct measure of effort whatsoever and identified effort with the word length. While Zipf's theory remains controversial, his results are not. A chemist would certainly reformulate the principle of the least effort as the principle of the fastest transformation.

Model 2

In Model 1 only one direction of transformation was possible. Chemical reactions usually run in several possible directions, theoretically, in all of them, which is an incredibly dense branching, fortunately, as implausible as the linguistic "discrete infinity." Let us take a more ambiguous case of the direct-indirect object tandem: GIVE (ACTOR, RECIPIENT, OBJECT) with several alternatives. It corresponds in English to at least two expressions (excluding a bunch of Passive forms):

1. ACTOR GIVEs an/the OBJECT to-RECIPIENT.

2. ACTOR GIVEs RECIPIENT an/the OBJECT.

The first expression can mean that ACTOR gives to the RECIPIENT one of several available types of OBJECTs, selecting an apple and not a book or a flower. It can also mean (a perilous situation in mythology!) that the ACTOR gives the single available apple to one of several RECIPIENTs. Or, it is one of several ACTORs who GIVEs the only available OBJECT to the only RECIPIENT. It may be a combination of various situations. It may also mean that the object is taken from the ACTOR by force, etc. It can be anything observable by an individual or a group. To reveal the ambiguity in the act, let us analyze the act itself, representing it as a transformation of a configuration, Figure 9.2, where the usual transformation **A** follows by a more complicated situation **B** . We can see a highly uncertain transition state with several outcomes.
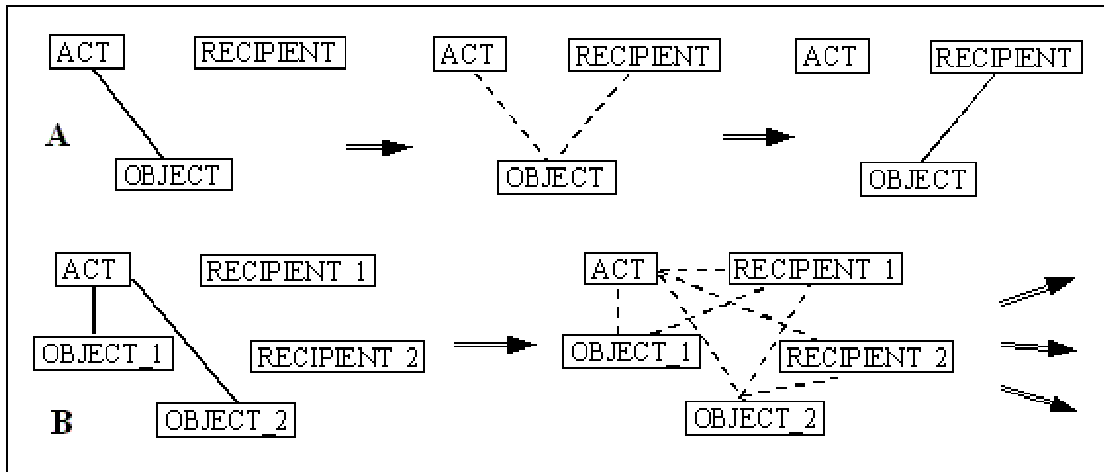
Figure 9.2. Action GIVE represented as configurations

In the initial state of transformation **A**, which can be reversible, the actor is connected to the object. In the final state, there is a bond between the recipient and the object. Between them lies the ephemeral and fleeting transition state in which both actor and recipient retain the connection with the object. The transition state has a higher level of uncertainty than both stable states because the outcome is not known: the transfer of the object can be delayed or cancelled, or the situation can turn into a fight.

We avoid the term entropy, using *uncertainty, stress, ambiguity*, and *irregularity* instead, because we cannot calculate entropy without a closed set of outcomes and a probability distribution over it, which is so easy in computer models but hardly ever is achievable in real life.

In transformation **B**, where at least two objects and two recipients are involved, the uncertainty is significantly higher and there is a whole array of outcomes regarding who gets what, if any. A web of relations of different *strength* connects the actor and the recipients, so that a combinatorial space of transition states can be described by a matrix of bond strengths between its components. The matrix can be strongly influenced by preceding states and memory traces. The relative affinities of the actor toward the recipients, as well as to objects, may be the decisive factor in choosing one recipient (object) over the other, as it in fact happens beyond mythology.

The outlined picture is consistent with the chemical paradigm, which can be roughly generalized as: the transformation through the least stressed (less ambiguous, uncertain, and irregular) transition state is the most probable one—**in the short run**.
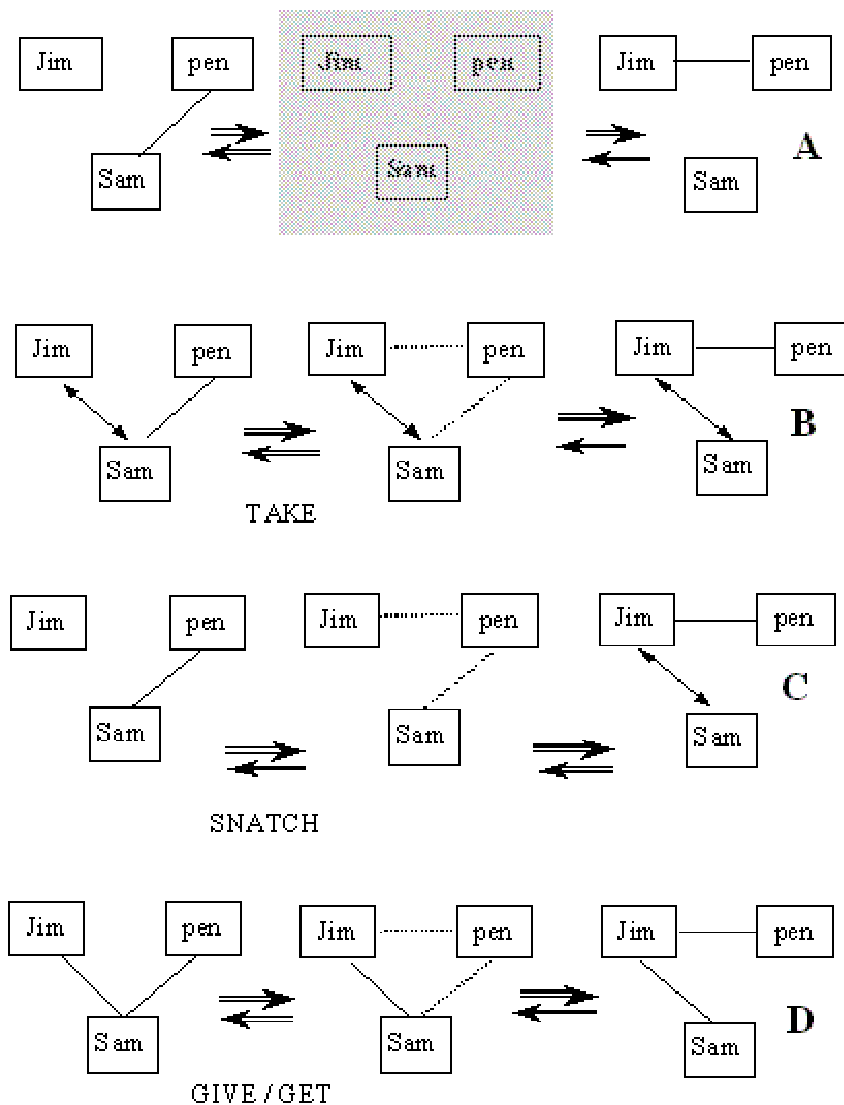

Figure 9.3   Variations of object transfer

The somewhat "thomistic" Figure 9.3 starts with the general situation **A** of transfer of a pen from Sam to Jim where the exact mechanism of is shielded by the gray square. Transformations **B** to **D** reveal a variety of subtleties in the abstract transfer:

B. Antagonistic relations shown by the two-head arrow suggest a forcible takeover.

C. Unexpected takeover generates antagonism.

D. Willful transfer against the background of friendly relations.

These subtleties reveal an important property of the real world: the source can have different images and the configuration of the source can be reliably reconstructed only from the totality of images.

The purpose of my illustrations is to show how the circumstances of an observable event, including the often complicated relations within its human and material participants, influence the outcome. The relations constitute the social framework and they are socially meaningful. We sometimes overlook, as with every big picture, that for some reason, human mind, language, tools, and society had appeared all together and, probably, are just the extrabiological aspects of *Homo sapiens*. If the tools go back to *Homo habilis*, about two million years ago, so may language. The questions of this kind are difficult to answer in any substantial way because the spoken language does not leave a material trace. All we know from observations of ourselves and animal societies is that communication of any kind is a *sine qua non* of social life.

It is not up to a chemist to engage into such discussions. What chemistry demonstrates, if abstracted from the material nature of its generators and bonds, is the fine structure of the transition from one pattern to another. It portrays **a discontinuity as a continuous process** and this is where it goes farther than René Thom.

In the chemistry of reversible transformations, once the system is initialized and brought into motion, the final **equilibrium** is defined in the long term by the energies of the initial and final state, while the short run **process** is defined by the height of the transition barrier. In the chemistry of open systems, to which life, mind, and society belong, the system can be maintained far from equilibrium as long as it is supplied with free energy and can dissipate heat, but in what particular form? In the form to which it has arrived through its preceding historical **odyssey**, with all its trials, errors, accidents and choices, following the beautiful metaphor of David Lightfoot.

The transition states are typical for the open non-equilibrium systems that are searching for a steady state after having been knocked out of the previous steady state. The almost forgotten profound analysis of this phenomenon of **homeostasis** belongs to Walter Ross Ashby, one of the founders of modern artificial intelligence (Ashby, 1960, 1964). The significance of his analysis for the problems of emergent properties comes from the fact that his homeostat was homunculus-sterile and algorithm-free. Regarding language origin, there was nobody to teach protolanguage to the first speakers. Language had to emerge from the spontaneous interplay of configurations in the mind with the configurations of the social life.

There is a deep parallel between a molecular system and the homeostat : the interacting molecules and the blocks of Ashby's machine spontaneously find a state of the lowest energy through a series of short-living transition states of an increased energy. The difference is that in the molecular system the number of blocks is very large. The most important similarity is that each molecule and each block have all the other kinds of blocks in their topological neighborhoods.

10. TIKKI TIKKI TEMBO: LANGUAGE AS A FORM OF LIFE

The molecular matrix is a mathematical object of the same nature as Greenberg's table of Arabic roots. Instead of the number of bonds between pairs of atoms, the distances between them or the bond strength values could be entered. Bond energy is a measure of improbability that the bond will be spontaneously broken. To somewhat vulgarize the chemical reality, the probability of bond breakup is extremely low at room temperature but goes up with temperature or irradiation. This vulgarization is minimal for hydrocarbons, such as the components of mineral oil.

In general, a simple skeletal graph is completely represented by a matrix of incidence consisting of only zeros and ones. Graph **is** the matrix, and its picture is only a visualization. The square matrix, however, can be made as realistic as we want by adding qualitative and quantitative flesh to the bare topological bones and inserting numerical values into the cells of the matrix.

For enthusiasts of long mental leaps, graph as a mathematical object is a beautiful example of an **extralingual language** universal, just look at Appendix 2. If it sounds oxymoronically like "the empty category is not empty," it is only because such terms as language and life each had acquired double meanings after the advent of computer science and molecular biology. There are human language and biological life, but there are also formal languages of programming and mathematics.

As soon as we have a population of interacting configurations, there is a possibility of an Artificial Life system.

In the most general terms, life—it could be called **meta-life**—is a system capable of (1) replicating itself with (2) errors (3) while using a limited resource of energy. The computer that uses the same tiny energy to display any picture on its monitor is an extremely misleading device for those who wants to live in the real world. People will not switch from despotism to democracy just because democracy and despotism are the words of the same length and require the same effort to type them on the keyboard.

It is impossible to review here the modern ideas about life-like systems, and the following remarks will be fragmentary. The major problem with any large picture is that you cannot find a large enough frame to hang it on the wall. You have to reduce it to a size where many important details are just specks. The picture painted by Devlin (2000) is more a window than a picture: it allows you to see the large world through its modest frame with the help of the optical abilities of mathematics.

For the Darwinian biochemical life, duplication and errors (mutations) are obvious, while the limited character of resources of matter and energy are not always kept in mind.

Heat alone cannot be directly utilized by life unless it is processed by power stations and internal-combustion engines and other thermal machines. Only humans have been capable of doing this, albeit only after 1700. This brought to life modern civilization, which is also a form of meta-life. Technology, the most conspicuous part of modernity, uses blueprints and descriptions the same way living cells use DNA for their replication. The limited resources of technology are capital, labor force, matter, energy, and consumption. This line of discourse would certainly detract us from protolanguage and I have to stop here.

There is a trend in linguistics that views language in terms of population dynamics, i.e. as a form of generalized life. For initial references, see Komarova and Nowak (2003), . Similar or related ideas can be found in works of James Hurford, Simon Kirby, Angelo Cangelosi, and others.

Language reproduces itself with mutations within the social communication. But what is its limited resource that exerts a selective pressure? It is not energy, because verbal exchange does not take a lot of energy. I believe, not claiming any originality, that

it is time for both expression and understanding. If communication is too slow and far behind the pace of events, it fails to perform its function and the bottleneck linguistic phenomena die out. The fast development of a distinctive language for wireless text messaging by teenagers is a supporting example.

The situation can be illustrated by the Japanese folk tale *Long Name*, known in USA as the Chinese folk tale *Tikki Tikki Tembo* (Tikki, WWW, A) .

As I remember it from my childhood, the parents gave their boy a very long name to ensure his happy and long life:

Tikki Tikki Tembo No Sarimbo Hari Kari Bushkie Perry Pem Do Hai Kai Pom Pom Nikki No Meeno Dom Barako.

When the boy fell into a well, the children around started to call his parents but could not chant his name correctly and had to start all over again. While they kept trying, the boy drowned.

NOTE. There was quite a discussion on the Web about whether *Tikki* was a Chinese or a Japanese tale (Tikki, WWW, B). I remember it as a Japanese tale in a Russian translation, and Ariko Kawabata, a participant in the discussion, confirmed it, which does not exclude its Chinese origin.

There are two "biological" approaches to language. One regards language as a classical biological adaptation (Pinker and Bloom, 1990), while the other one is to see it as a form of meta-life (general reviews: Komarova and Nowak, 2003, Hurford, 2003). The second trend opens the Pandora box of computer simulations, but the approaches do not conflict.

I believe it is appropriate to refer here to Cavalli-Sforza (2000) with his panoramic view of human evolution. He drew close evolutionary parallels of genome and language—parallels that did not necessarily intersect.

The problem with adaptation in Darwinism lies in the circularity of the concept of fitness. Cavalli-Sforza had a clear view of fitness as the rate of reproduction: the genome that reproduces itself faster is more fit. This approach, being essentially Darwinian,

breaks the vicious circle by offering a quantitative measure of fitness regardless of the material nature of meta-life and identifies fitness with nothing but the observable speed. It is accepted in ALife.

The entire treatment of life as competition of replicating sequences for a limited resource, it must be noted, comes from Manfred Eigen, a Nobel Prize recipient for chemistry (1967). His meta-chemical ideas were originally expressed in Eigen (1971-1978) and later in Eigen and Schuster (1979). Some elements of Eigen's theory, in the form of population dynamics, can be also found in Komarova and Nowak (2003). Eigen used a linguistic example for illustration (Eigen, 1977). This is how the initial sentence evolved after a number of reproductive cycles in a life-like model:

1   TAKE ADVANTAGE OF MISTAKE

5   TAKF !DVALTAGE OF MISTAKE

10   TALF ADVALTACE OF MISTAKI

70   TAKEB ?VALTAGI LV MIST!KE

The main thesis of kinetics is neither chemical nor physical. It is based on common sense: the more molecules are present in the unit of volume, the higher the probability of their collision, which increases the probability of the successful collisions that lead to transformation. The limited resource in chemistry is strictly material: the fastest reaction pulls the rug under the feet of competing reactions by consuming their starting material, common for all, so that the slower reactions are increasingly suppressed with time.

 Eigen's works started the entire area of Artificial Life. They are too technical to be explored here, but a whole host of beautiful and general ideas on evolution, fitness, language, and music, against a rich cultural and philosophical backdrop, can be found in his popular book with Ruthild Winkler, originally published in 1965 (Eigen and Winkler, 1993) well before his detailed works on molecular evolution. It contains a chapter on molecules and language.

I cannot resist the temptation of quoting Eigen on Chomsky: "..we could say that Chomsky's linguistics applies to language in the same way that thermodynamics does to

the weather," which is hardly a compliment from a natural scientist. For the context, the reader should look for the page 269 of Eigen (1993). In the same book one can also find a discussion of the frequency analysis of intervals in musical compositions very similar to Greenberg's analysis of Arabic roots. Of course, more modern sources can be found.

The forms of meta-life, other than biochemical life, are language, culture, customs, society, technology, science, art, even some games, and the list is open-ended. All of them originated and have been evolving on the platform of biological life, under which we find nothing but chemical reactions. The term Artificial Life, which originally echoed Artificial Intelligence, can be used instead of meta-life in both meanings: as human simulation of life and as forms of meta-life created by human reason and hands, whether intentionally or as the game of chance and necessity. In the context of Darwinism, the non-biotic life was characterized by Dawkins (1989) as life of **memes**, the cultural and mental counterparts of genes. Note, however, that it was Zipf who first used the expression "genes of thought."

If we accept the kinetic concept of evolution, then the significance of the main linguistic idea of Zipf (1949, 1965) becomes obvious. The patterns of live speech are selected along the criterion of least effort, comprising the shortest length, fastest utterance, and its fastest subsequent understanding.

To complement the ancient Chinese tale, here is a modern example. The Russian post-Communist revolution brought the abbreviation **MREO UGIBDD GUVD**, which itself is so long that it asks for another abbreviation, probably, **MUG**. It means:

The local office of the inter-regional office of technical inspection of transport of the department of state safety inspection of road traffic of the top directorate of internal affairs.

This *Tikki Tikki Tembo* of a kind corresponds to the American **auto registry.**

11. ZIPFING THE CHIMERA

In order to describe a source, we have to flatten, turn into a string, and squeeze its configuration through the bottleneck of speech. This may require breaking some bonds and forming some new ones, which is exactly what chemistry is about. Speech generation, from the point of view of a chemist, is a "chemilingustical" reaction. As any natural process in a non-equilibrium system, it requires **free energy**, which is physical energy in the form convertible into work. It is supplied to humans with food. Naturally, if language helps to get more food, it will survive in a population. There is nothing in thermodynamics, however, that predicts the origin of language. Not intending to engage further in a discussion of this large and difficult topic (see Pinker, 2003), I mention it just to point to a large area of the big picture: the adaptive aspect of language. The very fact that language is an adaptation in no way predicts its structural properties. In a sense, everything that exists in living systems is adaptation.

Figure 11.1 depicts the essence of the process of linearization of a non-linear configuration. The initial state is what is called **image** in PT. It is the configuration of the source processed by organs of perception—another vast topic of cognitive sciences.

The configuration space is the first of the *regular structures* that we are building. The next consists of *images*, a concept that formalizes the idea of observables. In other words, a configuration is a mathematical abstraction which typically cannot be observed directly, but the image can (Grenander, 1996, p. 91).

Unlike thoughts, images in individual minds can be compared and shared by comparing and sharing their verbal descriptions, drawing pictures, imitating, pointing, etc. Two people almost always agree whether the animal in sight is an elephant or a mouse.
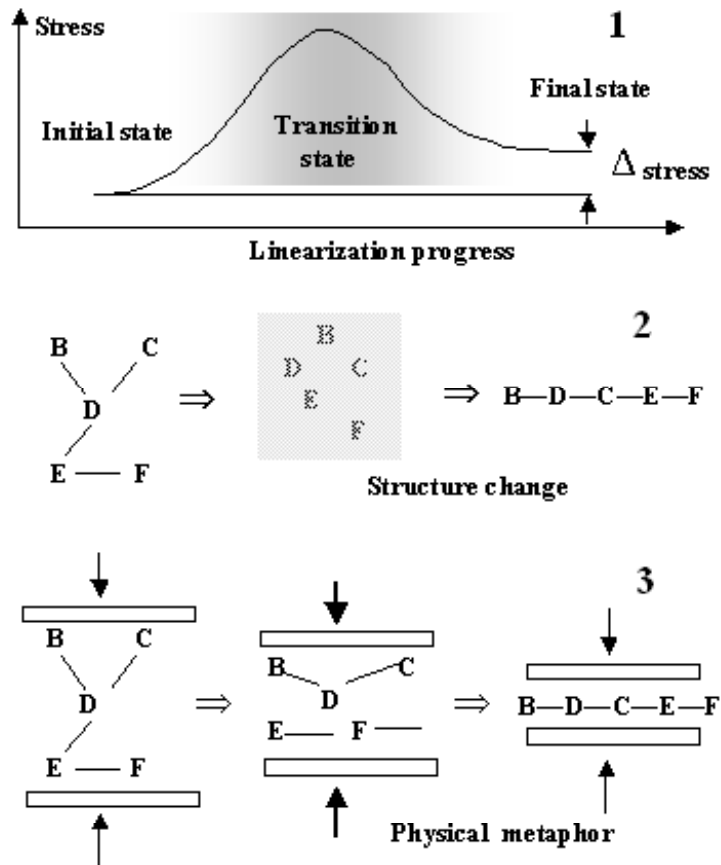

Figure 11.1 Linearization of the source into the output string

A lot of legitimate questions of epistemological character could be asked here: what is the relation between the source and its image? How can we know anything about the source if we have only its images in the brain? Do we really have an image of somebody's image, etc. I will not respond to any of them because, as a chemist with an upbringing of an experimentalist, I have to abide by two principles: to draw the distinction between observable and hypothetical and imaginary objects and the distinction between an established consensus and arguable approaches. This does not mean that I consider philosophy worthless, just the opposite. I believe it is still not only waiting to be

called science but also longing for it. I can only hint that the relation between the "*an sich*" (Kant) configuration and its image in PT may point toward a promising direction of investigation because both are represented in the same mathematical language.

In Figure 11.1 linearization is shown in three aspects:

1. Transition over the stress barrier from the initial to the final state;

2. The structure of the initial and final states with fuzzy intermediate state;

3. A physical metaphor of the process as a mechanical squashing.

In this way the unyielding porcupine of the source is pushed feet first through the apparatus of speech. I like to call this process **zipfing** to emphasize that it must be done with the **least effort** in order to compete with other modes of sound or sign communication. The least effort is required at all stages: to understand the image of the source, to break it up into singlets, doublets, and larger fragments and to align the fragments for the output. Note, that these stages form a cycle because the next stage will be again the understanding by a listener—the situation explored by Simon Kirby, James Hurford and others, see Kirby and Christiansen (2003). This brings us back to the ideas of Manfred Eigen about hypercycle and right into the identification of language as a form of life.

The residual stress (Δ_{stress}) is the measure of the irregularity of the output as compared with the source.

NOTE. A nit-picking chemist, as well as linguist, may inquire about the material balance of generators in linearization. How can we form two doublets with the same generator if there is only one generator of each kind? Here is a hint: there is a population of generators in the mind, similar to a population of molecules in a volume of liquid, but it is a population in time, not in space. For example, stuck in traffic, we repeatedly return to the idea of being late to a meeting, although any idea exists in a single copy. The time population of this idea is much larger under the circumstances than the idea of the world energy crisis.

Figure 11.2 further illustrates the idea of chemilinguistry. In **A**, the resulting triplets have the same degree of stress and are equally probable. The transition state can go either way. In **B**, the resulting **SOV** (Subject-Object-Verb word order) is, arbitrarily, more stressed and less probable than **SVO** because **V→O** in the source changes its

orientation to the irregular **O➔V**. The least effort condition is partially satisfied if the topology of the source is maximally preserved during linearization. I have to remind that we are dealing here with protolanguage, which has no grammar except the preservation of topology.

In chemistry, the structure that forms behind the lowest transition barrier dominates the final state. This principle, if applied to language, transforms into the following hypothetical definition of grammar in a meta-chemical but by no means metaphorical sense: Grammar is the catalyst of language generation
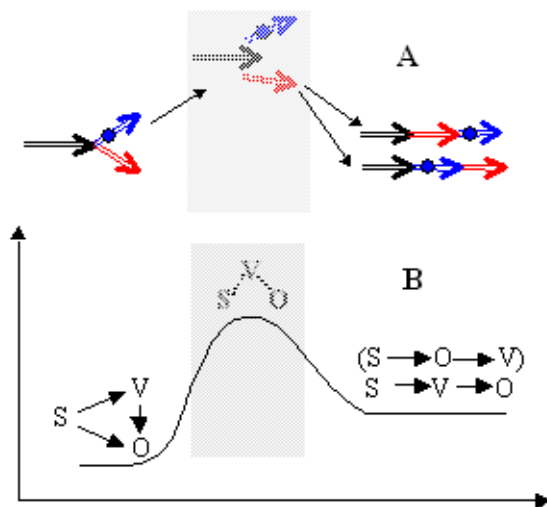
**Figure 11.2 Examples of linearization
A: Fork-like source; B: Triangular source**

Catalyst in chemistry does exactly this: it decreases the stress (energy) of the transition state. **Grammar** decreases the stress of transition state because of the preservation of linear fragments longer than doublets and because of introducing syntax, i.e., the means of topology preservation other than simple adjacency.

It seems to me that what Noam Chomsky has been searching for is the formula of this catalyst. Of course, this catalyst must be innate and, of course, it must be part of a larger picture. And of course, we still do not know what it is.

The short run situation, controlled by kinetics, applies only to the spoken language. An assiduous writer who has plenty of time to think and to rewrite the text and is not bothered, unlike Hemingway, by physical ailment, can really put the patience of the reader to a test by exhaustingly long sentences. An elitist writer and elitist reader will find each other like a sadist and a masochist.

As a self-illustration of the Zipf Principle, having little patience by nature, I prefer the short term **stress** to the somewhat longer but more accurate terms **irregularity**, and

improbability. I could also say **energy**, but stress and energy are opposites in pop psychology.

In order to legitimize the use of chemical analogies beyond the general pattern approach, we have to explain in what exactly way the chemical systems are comparable to the language generation systems. And what is catalysis, anyway?

Procreation, cooperation, competition, and social order need acts of contact and exchange. In the molecular dynamical systems the random events of exchange are collisions of particles. Collision is predominantly a binary event. Molecular system has no memory of its previous state, unless it is alive.

Switchboard systems are an alternative kind of dynamical systems. Modern telecommunication makes a physical contact at home or a marketplace unnecessary. Telephone communication is one example. Mind is another one: it acts as a switchboard system on short segments of time and the neurons do not dash around inside the scull.

Figures 11. 3 and 11.4, saving a lot of words, compare molecular and switchboard systems .

The switchboard (SB) dynamical systems are not the same as the **connectionist systems in** Artificial Intelligence, but I cannot refer to any source other than my own description. As far as connectionism is concerned, this is another contentious area from which I would like to stay away.

The SB system is just a mathematical image of a certain spontaneous activity which can be mapped onto molecular processes, while the connectionist networks are through-flow processing systems, usually, with feedback or under external control. The events in SB systems are momentary connections and disconnections between elements of the set of sites presented as small circles on a larger one. The physical and biochemical mechanism of the switching is irrelevant. The contacts involve a certain medium which can enhance the connection or hinder it, as well as remember it for some time. The fading memory of previous connections is represented by dotted lines. There are ephemeral connections, as well as long-lasting ones. The "movement" in the switchboard is spontaneous and has the properties of ordered chaos. This is completely opposite to any computer, even the one simulating ordered chaos.

With all the differences, both molecular and SB systems are, to some extent, conservative. While the isolated molecular system maintains its energy, the SB system requires external **free energy** (i.e., energy in the form capable of introducing order, unlike the chaotic heat) to make and unmake connections. It is thermodynamically open, but the constraints on the supply of energy limit the activity in the same way as temperature limits the average number of collisions in molecular systems. The SB system is thermodynamically similar to life because it stays far from equilibrium until the supply of energy lasts.

I am not aware of anybody working with a computer on a limited energy supply in such a manner that the computer itself is tweaking its software to do maximum computations per unit of energy. Such a computer would be a true model of the mind, but I believe that it could be created only within a population of similar computers exchanging segments of software and capable of errors. The open source software in the community of programmers is close to this form of evolution.

> I am sadly aware of a large and menacing boulder of software capable of unpredictable errors but protected by monopoly from selective pressure: I am using it for typing this paper.

Figure 11.3 needs little explanation. The molecular system is a series of collisions and the SB system is a series of connections fading with time. The dramatic difference is that the first one is completely chaotic while the other one is a **dissipative system**. The significance of this fact is enormous, but this is not a good opportunity to expand on it. Instead, the reader should turn to the numerous works, many of them popular, of Ilya Prigogine, for example, Prigogine and Stengers (1984). For a linguist interested in a larger picture, this area, however difficult, could be stimulating. It could suggest, for example, some against-the-tide research aiming at the connection, denied by many, between material culture and the character of a particular grammar. The production of free energy in the form of food is a necessary condition for an evolution of a complex living system. Some languages may be better than others for this purpose and they can themselves evolve faster.

The six images of Figure 11.4 illustrate the concept of catalysis. I will speak about it mostly metaphorically, but excellent popular sites can be found on the Web. First three images portray the molecular systems where particles are chaotically moving. The

catalyst limits the freedom of movement of colliding particles by forming fleeting bonds with them. It works as a biased switchboard operator who tends to connect his friends at the expense of all the others. When the catalyst is immobilized, for example, on a solid surface, the chemists speak about heterogenous catalysis.
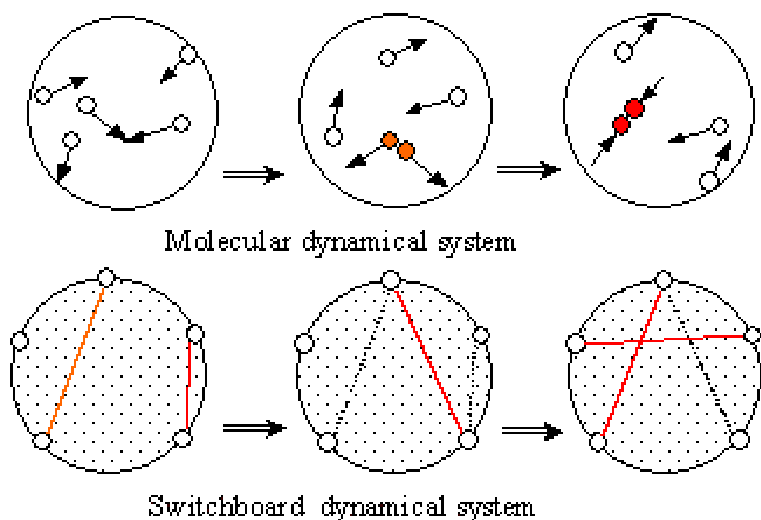
Molecular dynamical system

Switchboard dynamical system

Figure 11.3 Molecular and switchboard systems

In images 5 to 6 the space is topological but non-Euclidean (the latter is also topological). There is no movement in this space and all generators are immobilized. It is the configuration space of PT.

When a connection associates with a **new** generator (here comes **novelty**), it can symbolically represent the connection in any further configuration with its participation. Conversely, the symbol evokes its original meaning (image 5). In the case of learning (image 6), the frequent connection becomes permanent.

Figure 11.5 illustrates the template catalysis, which is employed by life for most important and intimate biological functions. This catalyst is a very large molecule, comparable with text, and it is active only at a small and moving area, as with reading a text. Moreover, as any text, it is capable to copy itself, which is what life means from the point of view of chemistry.
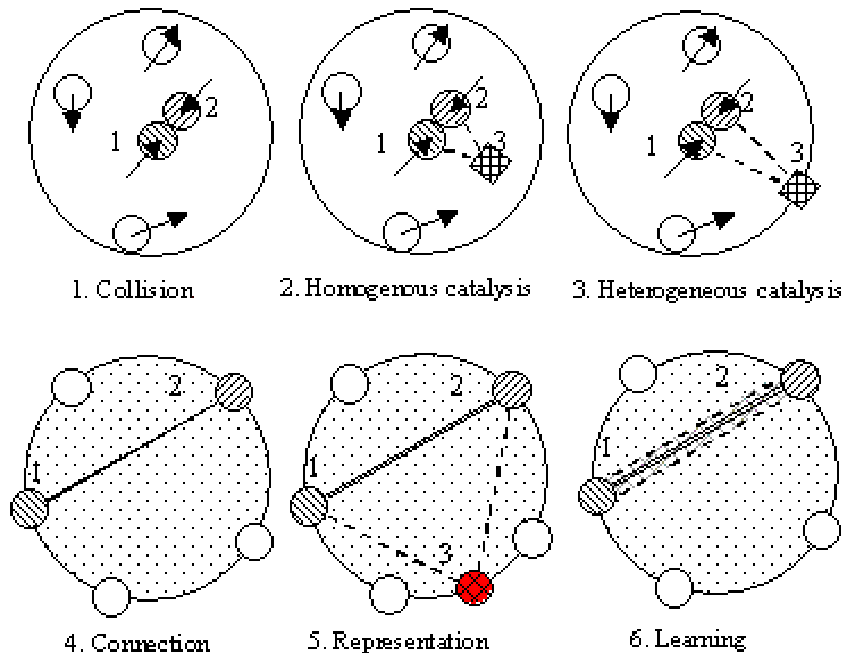
Figure 11.4 Catalysis and learning

The concept of **transition state**, which I touch upon only superficially in Chapter 9 and here, but more in depth in Appendix 3, is still little known outside chemistry and politics. It has been spreading recently as a new domain of complexity. Transition state begins to look like a paradigm of a universal importance of the same magnitude as thermodynamics and basic laws of nature. It may be underlying a lot of various phenomena, from geological activity such as earthquakes and volcano eruptions to "punctuated equilibrium" in biological evolution to the flight of albatross. It may turn enlightening in the study of evolution of language and society.

Figure 11.5. Template catalysis

The concept of punctuated equilibrium (Eldredge and Gould, 1972, Eldrege, 1985), sharply debated in biology, found a more hospitable soil in sociology, for example, in theory of organizations. The reason for that seems to be rather trivial: social institutions, unlike biological evolution, are directly observable. What follows from the observations in history and sociology—and even from individual human experience—is
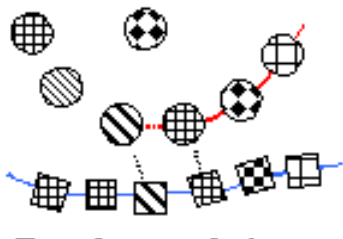
the alternation of long (most probable) periods of stability and even stagnation with the turbulent (less probable) periods of intense transition from one stable state to another. In mathematics this pattern is known as Lévy process or Lévy flight.

On a smaller scale, a similar pattern can be seen in the behavior of birds and animals, such as albatross (Viswanathan, *et al* , 2002) and jackal. The imitation of a typical pattern is shown in Figure 11.6.
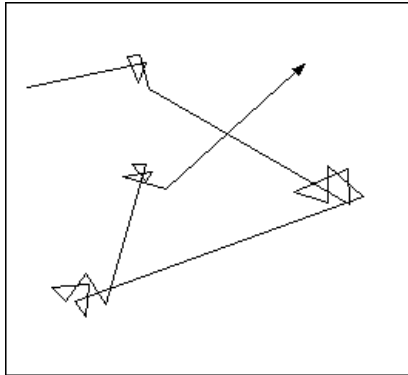

Figure 11.6. The 2D pattern of Lévy flight

The already mentioned work of Ashby (1960, 1964) on modeling homeostasis with a set of interconnected electric devices adds another certificate of generality to the concept of transition state. In his experiments, Ashby used a system of interconnected mechanic units exchanging electric input and output signals with each other. The units could be in one of a series of states each. The system could ultimately find its state of equilibrium. When disturbed by the experimenter, the system began a frantic and apparently chaotic search for a new state of equilibrium. Having found one, it could stay in it until the next perturbation. Ashby characterized this property of the system as homeostasis. He considered it the basic property of natural and artificial dynamic information systems. This alone was not new. The novelty was in the observation of the excited, unstable, and short-living transition state in a system of a switchboard type. The units, sitting on a bench, exchanged not by collisions but by information. Moreover, it was a dissipative system.

There is a wonderful discussion of punctuated equilibrium and related topics in Lightfoot (1999), who, by the way, mentions both Darwin and Marx as the founders of general evolutionary theory. It is hard to dispute this juxtaposition: in fact, Marx complements Darwin by his concept of history punctuated and driven by revolutions, but he borrowed the idea from Hegel. Anyway, both Darwin and Marx have no peace in their graves.

What is the truth? The truth is a consensus. This is why there is no eternal truth.

More about transition states in society and mind, see Tarnopolsky (2003).

12. A CHEMIST AND A CHIMP SPEAK NEAN

Nean is a sequence of single words and, mostly, word pairs that represent the bond couples in the source by word adjacency. The relative significance of generators in the source (i.e., topicalization, also called focusing) is expressed in Nean by their frequency (this is very PT). An utterance in "English Nean" would look like this:

Ug Og Ug meat

Nean is, actually, a pure grammar and it can be applied to any lexicon, assuming that the words are just labels for generators. We can speak, write, and sign in Nean using words of any language, as well as arbitrary symbols, as long as we have a source and do not fret over the philosophical relation between the "real thing" and its symbol.

I see protolanguage itself as a transition state toward a grammaticalized spoken language. On an evolutionary timeline, Nean sits in the very beginning of the transition to language. Right before the stage of developed language, Nean may take a somewhat more sophisticated form of the "haploid Nean" (see Chapter 7).

Beginning to speak, the children recapitulate the stages of Nean by moving from single words to doublets, triplets, and multiplets.

The subsequent development of grammar is beyond the scope of this paper. Plenty of good ideas can be found in Language Evolution (2003) and large literature, to

which a chemist can hardly add anything. The chimeric approach, however, may illuminate the problem from a particular angle. Thus, we can hypothesize, within the grammaticalization theory, that the inflections develop from words (see a short perspicuous overview by Hurford, 2003), especially those that signify large classes of objects. We can also hypothesize that verb is a later invention (from words like hand, foot, and mouth for verbs do, go, and eat) and this is why the **SOV** word order dominates at earlier stages of language evolution. If we regard the inflected word, like (he) SPEAK-s, as a former doublet, later collapsed, the genesis of morphemes with syntactic function becomes a natural continuation of Nean. Uniformitarianism—the idea of deep similarity of all languages—may have a point after all. This is exactly what I mean by saying that all languages perform the same function—linearization of the source configuration—by the same means of preserving binary connections during zipfing (a.k.a., trash-compacting).

My own experience of a translator tells me that it is impossible to render a technical or in any way special text unless the translator has reconstructed a mental image of the source. Moreover, such languages as Legalese and Patentese may require translations from English into English. The situation is more subtle, but basically the same with poetry. Understanding is the ability to convey the source to somebody else. The language cannot be outsourced.

It is sufficient to turn to the reviews by Tomasello (2003), Hurford (2003), and others from the same source to see that my chimeric vision of Nean is by no means revolutionary. There is a lot of hands-on work going on in linguistics that would please a chemist-realist, who would, however, still distrust any computer simulation in which no **thermodynamics** (with an energy-like parameter) and **kinetics** (with a transition state) are involved. What Nean could mean against this positive and energetic background is, paraphrasing Michael Tomasello, an "evolutionary fairy tale with which to **begin**" (instead of "**conclude**," Tomasello, 2003, p. 108).

In this Chapter, I illustrate the transition from a source to the utterance with examples. If only the meaningful words are left, the **previous sentence** of this paragraph can be in various ways decomposed into doublets, for example:

I illustrate	source utterance
illustrate transition	illustrate example
transition source	Chapter illustrate
transition utterance	(and others)

In a context, the shredded sentence still tells something.

Using haplology, we can obtain some triplets and longer oligomeres:

I illustrate + illustrate transition = I illustrate transition

transition source + transition utterance = transition source utterance

(if there is no doublet utterance source, then transition source

utterance is more probable than transition utterance source).

Chapter illustrate example

etc.; one can play with this linear LEGO.

As it can be seen, the choice between "triangular" structures like transition source utterance or transition utterance source depends on the orientation of the bond between utterance and source. We could see that in a very general form for the basic syntactic formula of word order, Figure 11. 2 (Chapter 11). The linearized configuration will depend on the direction of the **V,O** bond. This is, of course, pure speculation, but it could be up to a linguist to confirm or reject it by observable facts, and, quite possible, such facts are already somewhere in the literature. There could also be a proof that **SVO** corresponds to a better zipfing in a faster moving world with longer utterances. By the way, the growing string length per source applies a harsher selective pressure regardless of the pace of cultural progress.

Figure 12.1 presents further illustrations to the case of linearization of a triplet in the case of the basic word order. The fuzzy non-oriented transition states in the middle column are obtained by breaking a single bond of the source. They lead to the final states in the right column. We can also see how the factor of the relative strength of a bond could work. But what can we say about the invisible transition states, whether in chemistry or in linguistics? Isn't it the same as fighting the Gothic monsters of formal

theory? Well, I can't deny it is, to some extent, but there is an important difference, better seen from the chemical side.
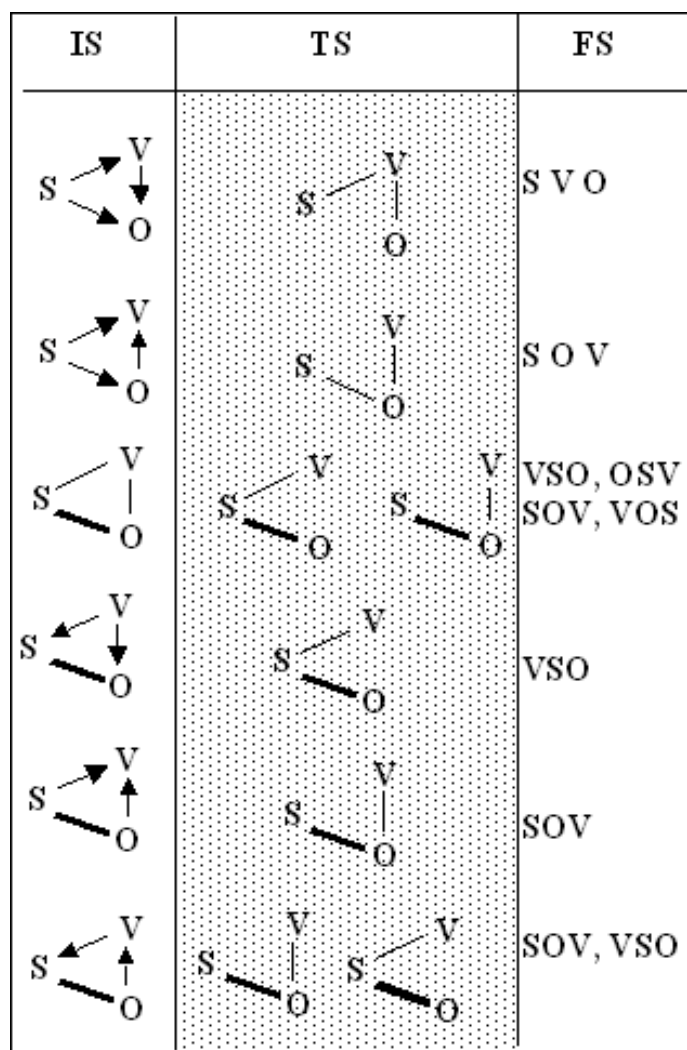

Figure 12.1 Initial (IS), transition (TS), and final (FS) states for S,V,O order. Stronger bonds are marked by thicker lines.

The so-called Hammond postulate in chemistry—one of a few most fundamental ones—illustrates how chemistry handles transition states. George S. Hammond (Hammond, 1955), being perfectly aware that his postulate was unprovable at the time (it could be proved or refuted someday), thus formulated his idea in the language comprehensible by a chemolinguist:

> If two states, as for example, a transition state and an unstable intermediate,
> occur consecutively during a reaction process and have nearly the same energy content,
> their interconversion will involve only a small reorganization of the molecular structures.
> (Hammond, 1955).

It means that if two structures along the transition pathway have similar stability, they have similar structures. One of the corollaries is that in the beginning, the transition state is closer to the initial and in the end to the final state.

What Hammond did was taking the general idea of transition state formulated in terms of energy from physics and adding to it the structural aspect alien to physics. Hammond postulate is so general and so extra-chemical because it is stated in two universal scientific terms that belong to all natural sciences: energy and structure. It is the "lack of energy," sorry for the pun, that distinguishes all formal linguistic theories.

I have no firm ground under my feet, however, when stepping on a linguist's turf. It is quite possible that the formal theory is compatible with an adapted Hammond postulate in one form or another.

What Hammond postulate itself lacks, sharing this shortcoming with formal linguistics, is the evolutionary aspect. It is of no importance whatsoever in the chemistry of simple systems, but is crucial in the study of any complex natural system, such as language or biochemistry. All those systems have evolutionary memory. The generalization of the concept of transition state over the evolutionary transformations, which I suggest, would mean at least a temporary completion of the universal theoretical framework for complex systems, provided the non-equilibrium thermodynamics is also included in the picture.

The order in which the generators of the source configuration historically appeared—the heredity and genealogy of the language—is an important part of the explanation of its current state.

> And in some cases this new function of the word is the first instance of this function
> being fulfilled at all, in the language concerned (Hurford, 2003, p. 52).

In other words, the present of the complex system depends on its past. Chemistry is based on the opposite principle: the properties of a molecule do not depend in any way on its origin.

To be more exact, the distant future of the complex system depends on its close past.

To complement the above quotation from George Hammond, here is a quotation from Nim Chimpsky, an educated chimp:

Give orange me give eat orange me eat orange give me eat orange give me you.

(Terrace, 1970, p.210)

This is a perfect Nean.

Still, I will have the last word, quoting myself, see page 29:

$$C—H_1 \ , \ H_1—C \ , \ C—H_2 \ , \ C==O, \ H_1—C \ , \ O==C \ , \ C—H_1 \ , \ C==O \ !!!$$

Hey, my Nean is as good as Nim's.

13. SCENES FROM THE CAVE LIFE TOLD IN NEAN

Hurford (2003, p. 53) notes that the study of grammaticalization requires going backwards from the modern to the earlier and simpler stages of language. Thinking backwards from the products of a transformation to its initial state plays an important role in the daily work of the chemist. For example, thinking about the best way to synthesize a naturally occurring drug, an industrial chemist develops a converging tree of routes leading to the goal. Another chemist may imagine a different tree of pathways leading to the natural synthesis of the drug in the plant. As the next step, both have to select a few most probable transformation chains, which is done using different selective criteria. The industrial chemist uses the overall cost as the criterion, trying to involve a minimum of intermediaries and byproducts, while the biochemist looks for particular intermediaries and byproducts. Let us take note of both completely compatible investigative approaches. In criminal investigation, the first is echoed by *cui bono*, looking for the one who profits most from the crime, while the second corresponds to collecting material evidence.

Figure 13.1 shows three kinds of problems arising in chemical research that are common for all studies of structural transformations, including human history and politics. Another general question can be added to them: what else could happen?

It is impossible to reconstruct the origin of language in the absence of any however fragmentary data, i.e., intermediaries and byproducts, in the chemical parlance.

What may be possible is to understand the principles that guide the evolution and find or reinterpret some evidence available at its later stages.
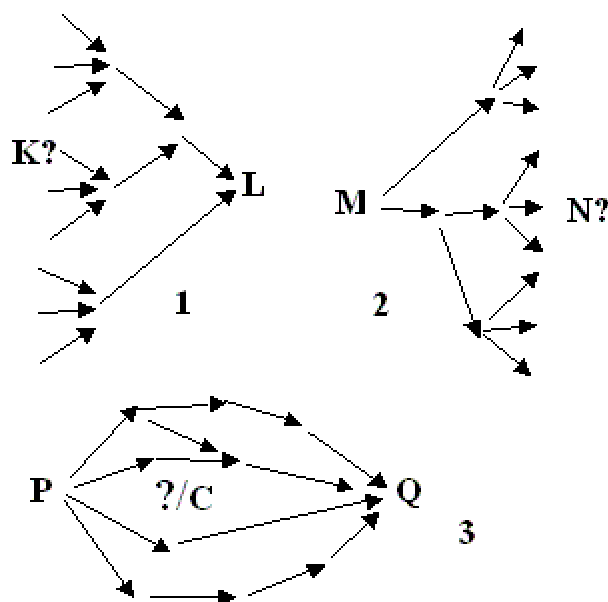

Figure 13.1 Three kinds of chemical investigation:
1. Can L come form K?   2. Will M generate N?
3. Which pathway from P to Q is optimal regarding
 condition C?

In this Chapter, I would like to recapitulate some ideas and expand on the chimeric principles of historical investigation of protolanguage, without claiming any positive results, but using positive examples.

Since we are interested in the development of language from protolanguage, let us, as it is appropriate for a chemist, go back from the following **full-language** expression:

In the cave Ug grimly gives a bone to Og.

The source of the expression is an image of a situation in the (non-Platonian) cave. The complete situation, which may not be seen in all detail by each witness, can be more complex, regarding the background, participants, objects, time, weather, motives,

physical health, social tensions, etc., and a history, in which the current source is just a limited projection (image) of the most recent episode.

Our strategy is to **hypothetically** reconstruct the way from the source toward its transformation into the linear expression, using our knowledge of the caves, inanimate objects, animals, humans, and their interaction. We assume that the laws of nature were the same in the ancient cave as they are today.

We can attempt the reconstruction by comparing various sources that can lead to the same expression. We can do that only by using the same language for both source and expression, which is the language of PT: our objects are configurations with a connector graph and numerical properties of the generators and bond couples. These properties are similar to the numbers in the Greenberg tables, but here we assign them to a large extent arbitrarily and just see what comes out of it. The following play with various sources is similar to chemical experiments in which we change conditions, observe the results, learn something, and design new experiments to learn more.

To portray the source configurations, we will use a large circle with generators positioned on it as small circles, symbolizing the switchboard system. The bonds may have direction and be of negative strength, corresponding to repulsion.

It is necessary to establish some principles of attributing direction to a bond. We can assume that there is a certain order of precedence between generators. For example, both Ug and Og have, probably, existed for about twenty years, but the bone appeared only today. This is why we have Og → bone orientation. Similarly, the good old Ug gives the freshly cleared bone only today and not everyday: Ug → give. By the same logic, we should direct give to Og: give → Og. The configuration of **Source 1** is largely arbitrary. Our goal is not the right configuration but to see how the change of the source influences the transition state and the output.

We don't know what is true. Again, the truth is a stable consensus, always temporary. More on this subject, see the famous Kuhn (1962), which is another illustration to the theory of transition state.

I need one more digression. Note, that this approach has to keep in both the source and its representation in the same frame. Therefore, **hypothetically**, knowing something about the evolution of the source, for which we may have some hard archeological data, we may draw conclusions about the language change. This is not a popular point of view, but Newmayer (2003, p. 73) tends to go against the (political correctness) tide and there is not a single principle of nature to support the tide.

Thus, it seems probable that the English language experienced one of the most dramatic changes because of the active, turbulent, and charged with energy social evolution, plus the ethnic mixing of England during its early history. The cultural factors (education, press) later stabilized the evolution.

The long and well documented history of Sumerian language is a great source of evolutionary facts (for a quick look, see Halloran, WWW), although not about the spoken form. Lightfoot (1999) assembled and analyzed rich and intriguing material on language change in his both detailed and "high-ground" book and placed it on a large picture with history as natural process in its center. I believe that the research in the co-evolution of economy, social order, culture, and language may bring some positive results. Language, like tea in England, beer in USA, vodka in Russia, and maté in South America are means of social bonding.

Bottéro (1992), in his breathtaking book, gives a remarkable opportunity to peek into the mechanics of the ancient mind through the tablets with interpretations of dreams. We see how words become the source of thoughts, reversing the usual order.

Let us start with the following source:

Source 1
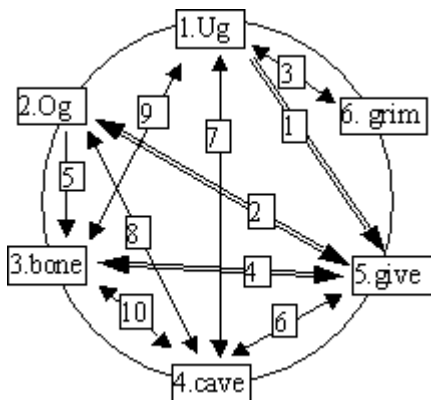
This configuration reflects our belief that it is Ug who looks grim (bond 3 between generators 1 and 6), Ug may or may not give and is, therefore, a primary generator regarding give, bond 1 directed toward generator 5 and is especially strong, etc. We direct bond 5 toward bone and just see what follows. The picture is utterly hypothetical and intentionally complex: everything seems connected to everything.

Of course, it is not: neither **bone** and **grim** nor **Og** and **grim** interact. Unreasonably, however, the bond **Ug** – **Og** is absent.

The generators of the configuration are labeled by six numbered words in the list:

words = Ug Og bone cave give grim
        1   2   3   4   5   6

The source configuration is coded in the 6 x 6 connectivity matrix, the non-zero elements of which are the bond affinities, taking only two arbitrary values, 1 and 2. Instead of **affinity** I will use a shorter **score.** The bond with score 2 is stronger (more probable) than the bond with score 1 and it is denoted with a double line. The diagonal could, in principle, reflect the weights of the generators. For the sake of experimentation, the scores are assigned to the bonds intuitively and tentatively.

	1	2	3	4	5	6
1	0	0	1	1	2	1
2	0	0	1	1	2	0
3	1	0	0	1	2	0
4	1	1	1	0	1	0
5	0	2	2	1	0	0
6	1	2	0	0	0	0

For practical reasons, we will code the same information in the form of sparse matrix **S**, which lists not the generators but the non-zero bonds.

S =

1	1	5	2	1		7	1	4	1	2
2	2	5	2	2		8	2	4	1	2
3	1	6	1	2		9	1	3	1	2
4	3	5	2	2		10	3	4	1	2
5	2	3	1	1						
6	4	5	1	2		Columns: 1	2	3	4	5

In rectangular matrix **S**, the first column is the bond number, next two columns are a pair of connected generators, column 4 contains the score of the bond and the last column indicates whether the bond is directed (1) or not (2) from the first to the second generator in the row.

The linearization is performed by a simple MATLAB program called, as the language, nean (Appendix 4), which can be easily modified and improved by anybody more experienced in programming than myself. It is probably easier to create it from scratch.

The input data are matrix **S,** array **words**, integer **NN** and real number **score.**

Example: The input of **words** and **S** for **Source 1** is:

words= ['Ug '; 'Og '; 'bone '; 'cave '; 'give '; 'grim '];

S =[1 1 5 2 1; 2 2 5 2 2; 3 1 6 1 2; 4 3 5 2 2; 5 2 3 1 1; 6 4 5 1 2;7 1 4 1 2; 8 2 4 1 2; 9 1 3 1 2; 10 3 4 1 2];

The program does the following:

1. Using the function of **random permutation**, it generates a random linear sequence of generators in **words**;

   Example: Ug bone cave give grim Og

2. checks each pair of neighbors against the matrix **S** of the source configuration.

   Example: Yes, Ug is coupled with bone, …No, Og is not coupled with grim.

3. calculates the sum of all adjacent generator pairs that are connected in **S** and adds it to the overall **score** of the sequence;

4. repeats steps 1 to 3 **NN** times and compounds a list of all different permutations in **words** with the same total **score.**

It is not absolutely necessary, but Nean needs the number of cycles NN large enough to guarantee that all possible permutations are checked against the total **score** criterion for all selected sequences. Translating into Chemicalese, the number of molecular collisions should be large enough.

Important!   The goal of the program is to simulate a completely random process, similar to molecular collisions. It does not code any intellectual activity and contains no algorithm other than calculating (not simulating!) the total score and counting identical strings. Its algorithmic part relates only to packaging the data, not generating them, because no algorithm exists for a random sequence. Computer can only imitate random (actually, pseudo-random) numbers. **The universal grammar for Nean cannot be learned: there is nothing to learn. Chaos is inherent in any large natural system such as the mind.** What cannot be learned has no source and is **new** and autopoietic, i.e., self-emergent. The **learnable grammar** starts with the haploid Nean.

Here is a typical output of program nean:

```
Og  give bone cave Ug   grim          grim Ug   cave bone give Og
bone give Og   cave Ug   grim         grim Ug   give Og   bone cave
cave Og  give bone Ug   grim          grim Ug   give Og   cave bone
grim Ug bone give Og   cave           grim Ug   give bone cave Og
grim Ug  cave Og  give bone
```

```
 number of cycles NN=5000, run time: t=0 min,
 score = 7, score matches 60,  number of strings 9
```

There are no sequences with the score over 7. The lower score input leads to a larger number of sequences. The total number of permutations is 6!=720.

Conclusion (not surprisingly): A complex and confusing source leads to a highly degenerated and confusing output.

Source 2.
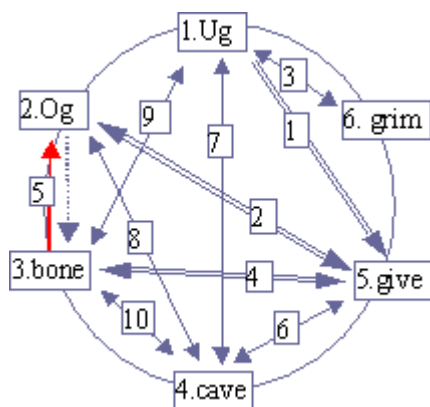
The only change in **Source 2** is the reversal of bond 5: bone → Og instead of Og → bone. We can regard it a mutation in **S**:

S   =[1 1 5 2 1; 2 2 5 2 2; 3 1 6 1 2; 4 3 5 2 2; 5 **3 2** 1 1; 6 4 5 1 2;7 1 4 1 2; 8 2 4 1 2; 9 1 3 1 2; 10 3 4 1 2];

```
Og   give bone cave Ug   grim              grim Ug   cave bone give Og
bone give Og   cave Ug   grim              grim Ug   give Og   cave bone
cave Og   give bone Ug   grim              grim Ug   give bone Og   cave
grim Ug   bone give Og   cave              grim Ug   give bone cave Og
 grim Ug   cave Og   give bone
```

 number of cycles: NN = 5000, run time: t=0.13723 min,
 score= 7, score matches 47, number of different strings 9

The change, as compared with **Source 1**, is emphasized in bold print.

The ambiguity, therefore, remains. Nevertheless, we notice that six out of nine sequences for both sources start with grim Ug, so that the initial position of the subject has a better chance to be generated at random.

Source 3
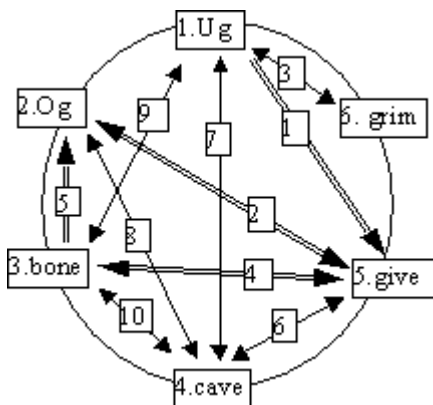
Intuitively we can guess that the ambiguity can be resolved by:

1. Introducing take or Passive, which can be done only through grammar or lexicon.

2. Strengthening S→V or S→O bond

We increase the strength of bond 5:

S   =[1 1 5 2 1; 2 2 5 2 2; 3 1 6 1 2; 4 3 5 2 2; 5 3 2 **2** 1; 6 4 5 1 2;7 1 4 1 2; 8 2 4 1 2; 9 1 3 1 2; 10 3 4 1 2];

The effect is striking: a single sequence at score 8.

 grim Ug   give bone Og   cave

 number of cycles: NN = 5000, run time: t=0.16667 min,

 score= 8, score matches 4, number of different strings 1

For comparison, at score=7 we have a terrible degeneration:

Ug give bone Og cave grim	grim Ug bone give Og cave
Og give bone cave Ug grim	grim Ug cave Og give bone
bone Og give cave Ug grim	grim Ug cave bone Og give
bone give Og cave Ug grim	grim Ug cave bone give Og
cave Ug give bone Og grim	grim Ug cave give bone Og
cave Og give bone Ug grim	grim Ug give Og cave bone
cave grim Ug give bone Og	grim Ug give bone cave Og
give bone Og cave Ug grim	grim Ug give cave bone Og
grim Ug bone Og give cave	grim cave Ug give bone Og

number of cycles: NN = 5000, run time: t=0.169 min,

 score = 7, score matches 129, number of different strings 18

The resolution of ambiguity may come from a simplification of the source: only a few bonds decisively contribute to the score. Or: Don't look around, focus!

As we could expect, the complexity of language comes form the original simplicity of protolanguage. With only a few words in the source, it becomes possible to express some of source configurations in an ordered-chaos way in spite of the completely random generation.

Source 4

Moving backwards from complexity to simplicity, we leave only the strongest bonds of **Source 3** in **Source 4**:

S   =[1 1 4 1 1; 2 2 4 1 2; 3 4 3 1 2; 4 3 2 1 1];
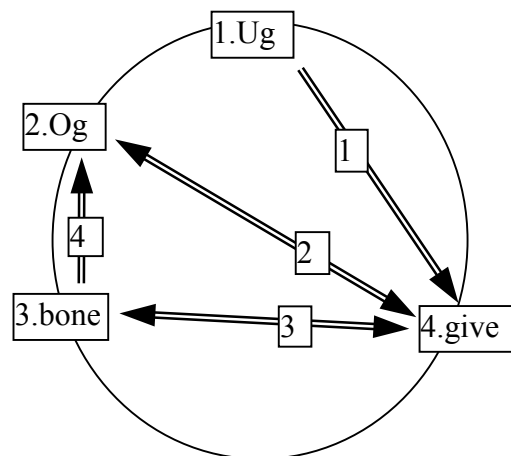
words= ['Ug   '; 'Og   '; 'bone '; 'give '];

The output is singular at a very low score:

   Ug   give bone Og

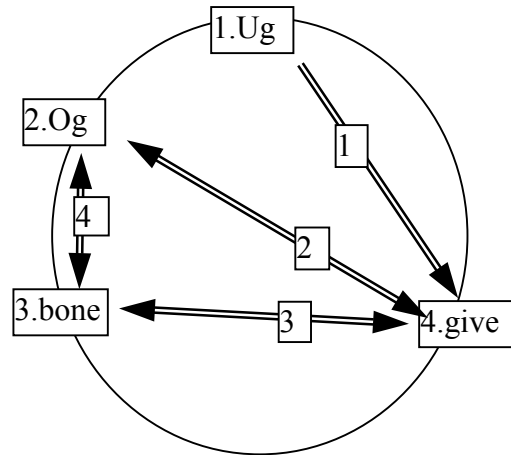 number of cycles: NN = 1000, run time: t=0.0072833 min,

 score = 3, score matches 40, number of different strings 1

Source 5

Unless we know it for sure, there could be an
inherent ambiguity about G2 and G3: first was
Og, then bone or first was bone, then Og?

Lets us make Bond 4 bi-directional:


```
Ug   Og   give bone
Ug   bone Og   give
Ug   bone give Og
Ug   give Og   bone
Og   Ug   give bone
Og   give bone Ug
```

```
bone Ug   give Og
bone Og   Ug   give
bone Og   give Ug
bone give Og   Ug
give bone Og   Ug
```

number of cycles: NN = 5000, run time: t=0.04635 min,

 score = 2, score matches 2326, number of different strings 11

As result a whole array of word orders arises. The asterisk stands for indirect object Og.

SVO
Ug   *   give bone
*   Ug   give bone
Ug   give *   bone

SOV
Ug   bone *   give
Ug   bone give *

OSV
bone Ug   give *
bone *   Ug   give

VOS
*   give bone Ug
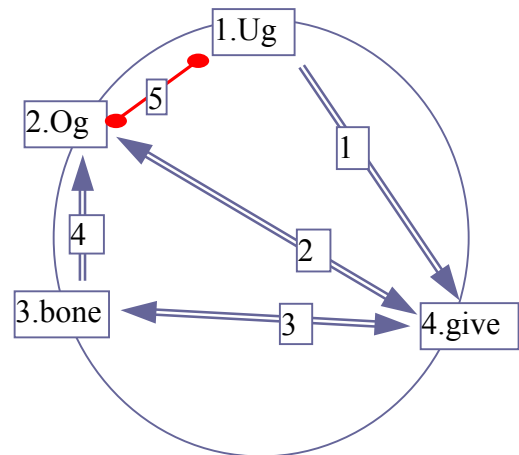give bone *   Ug

OVS
 bone *   give Ug
 bone give *   Ug

Source 6

Let us introduce antagonism Ug—Og by assigning a negative value to bond 5.

S = [1 1 4 1 1; 2 2 4 1 2; 3 4 3 1 2; 4 3 2 1 1; **5 2 1 -1 2**];

words= ['Ug   '; 'Og   '; 'bone '; 'give '];

At score =3:

Ug   give bone Og

number of cycles: NN = 5000, run time: t=0.054 min,
 score = 3, score matches 237,   number of different strings 1

At score =2:

Ug bone Og give	Og give bone Ug
Ug bone give Og	bone Ug give Og
Ug give Og bone	bone Og give Ug

number of cycles: NN = 5000, run time: t=0.056 min,
score = 2, score matches 1259, number of different strings 6

Now let us increase the antagonism:

S =[1 1 4 1 1; 2 2 4 1 2; 3 4 3 1 2; 4 3 2 1 1; **5 2 1 -2 2**];

Ug bone Og give	Og give bone Ug
Ug bone give Og	bone Ug give Og
Ug give Og bone	bone Og give Ug

number of cycles: NN = 5000, run time: t=0.044 min,
score = 2, score matches 1277, number of different strings 6

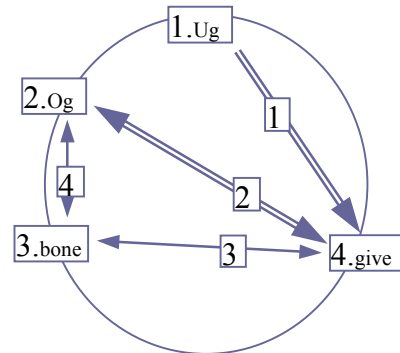Next, we make GIVE predominantly a function of two variables:

S =[1 1 4 **1** 1; 2 2 4 **1** 2; 3 4 3 **0.5** 2; 4 3 2 **0.5** 2];

At score = 2.5:

   Ug   give Og   bone

number of cycles: NN = 5000, run time: t=0.045 min,
score = 2.5, score matches 217, number of different strings 1

At score = 2:

   Ug   give bone Og
   bone Ug   give Og

 number of cycles: NN = 5000, run time: t=0.045 min,
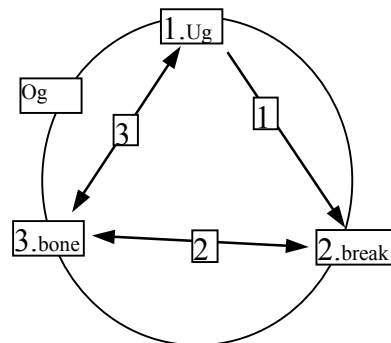 score = 2, score matches 397, number of different strings 2

Source 7

Basic word order.

words= [' Ug   '; 'break '; 'bone '];
S   =[1 1 2 1 1; 2 2 3 1 2; 3 3 1 1 2];

   Ug   break bone
   Ug   bone break
   break bone   Ug
   bone   Ug   break

 number of cycles: NN = 5000, run time: t=0.027 min,
 score = 2, score matches 3309, number of different strings 4

COMMENT: the information about **S** and **O** is preserved. The relationship between Ug and bone is unclear. What bone?

Next we differentiate bond strength:
bond 1, score 2;   bond 2, score 1.5;   bond 3, score 1 .

words= [' Ug   '; 'break '; 'bone '];
S   =[1 1 2 2 1; 2 2 3 1.5 2; 3 3 1 1 2];

At score = 2.5:

  Ug   bone break
  break bone   Ug

 number of cycles: NN = 5000, run time: t=0.025 min,
 score = 2.5, score matches 1664, number of different strings 2

At score = 3:

   bone   Ug   break

 number of cycles: NN = 5000, run time: t=0.024 min,
 score = 3, score matches 880, number of different strings 1

At score =3.5:

Ug   break bone   (now we are speaking English!)

 number of cycles: NN = 5000, run time: t=0.02 min,
 score = 3.5, score matches 819, number of different strings 1

Note a high probability of the expression: 819 out of 5000, or 16%

Here is another try at the distribution of bond scores.

words= [' Ug   '; 'break '; 'bone '];
S   =[1 1 2 2 1; 2 2 3 1 2; 3 3 1 1.5 2];

At score = 2.5:

   Ug   bone break
   break bone   Ug

 number of cycles: NN = 5000, run time: t=0.026 min,
 score = 2.5, score matches 1669, number of different strings 2

At score =3:

   Ug   break bone

 number of cycles: NN = 5000, run time: t=0.025 min,
 score = 3, score matches 866, number of different strings 1

At score = 3.5:

bone   Ug   break

number of cycles: NN = 5000, run time: t=0.026 min,
score = 3.5, score matches 831, number of different strings 1

From the point of view of chemilinguistry, the above experiments are meaningless because they have no real linguistic context. My goal was limited to a demonstration of a possible tool, the application of which is up to nobody but a linguist. This chemolinguistic tool could be just a part of a much larger tool kit combining structural, thermodynamic, kinetic, and evolutionary approaches, and, most importantly, the realistic sources.

A pretty close model approximation to language as natural process is the theater stage where the spectators can see the action and the background, hear the actors, and have different opinions about what is going on. To study language as natural process without approximation, we have to let others watch our daily and nightly life.

13. CONCLUDING REMARKS

For conclusions one should go back to Chapter 2, *Preview*. Here I would add a few notes about the relation between chemistry and linguistics.

I strongly doubt—together with many linguists—that we will ever be able to reconstruct the genesis of language as it indeed happened. We will simply have no evidence, unless we find, as Baron Munchausen did (Raspe, WWW, Chapter IV), some frozen sounds of the primeval chat preserved in the permafrost.

What we can do, however, is to see how it could and could not have happened in principle and to check the theory on practical development of vocal communication between humans and computers, as well as between androids and, God forbid to omit, gynoids. What we ultimately need is the study of language as a **natural process** and not just an insightful computer simulation or the intellectual game of "why not" and "what if." This is where chemistry and linguistics find themselves in the same naturalist society.

As recent American political history shows, we can reach remarkable heights with surprisingly limited language skills. Besides, eloquence does not come from the knowledge of linguistics. A reasonable question from am outsider is: Why do we need linguistics at all?

There is more to language than utility: like air, sea, and mountains, it is a beautiful and delightful medium for humans. It is also the starting point of their mating rituals and the endpoint of the relationships. As observation and study of nature, the study of

language will always kindle the interest of lay people and, probably, even of the future androids and gynoids.

Reading the recent overview of the problem of genesis (Language Evolution, 2003), a natural scientist could feel some professional gratification: linguistics is becoming an exciting **natural science**. Moreover, as one can feel especially from the essay of James R. Hurford (Hurford, 2003), it is part of a much larger paradigm shift or, I would say, transition state, in science as a whole. Regretfully, it is too little known how much chemistry has contributed to the large picture of the world, apart from explaining the most intimate mechanism of life.

There are countless variations in linguistic literature on the theme of Humboldt: "the infinite use of infinite media," see an intriguing discussion in Studdert-Kennedy and Goldstein (2003). This is where the sister sciences tend to go separate ways. For the chemist, the potential infinity of atomic combinations is of no relevance. There is a strict division into the **existing** and the **hypothetical**, on the one hand, and the **known** and the **new** chemical entities, on the other hand. The latter immediately turn into known as soon as their descriptions are published, but there is a daily deluge of new ones.

In some areas of linguistics, any written or spoken sentence is as good as any other, regardless of whether it was repeatedly used in real life situation or not, unless it is "ungrammatical" from the point of view of a language maven. For the chemist, the existence of a chemical compound must be proved by its synthesis and isolation. Nevertheless, not only can we hear ungrammatical sentences all around us, but our entire civilization is built of the daily tide of right and wrong, heresy and orthodoxy and, as chemistry tells us, life itself developed from errors.

The chemical view of the world is part of the general non-Newtonian, non-Einsteinian, and non-quantum (in spite of the quantum theory being the deep foundation of chemistry) paradigm that began to penetrate, first, sciences and then to knock on the door of humanities after the first works of Ilya Prigogine and the first steps of Artificial Intelligence. I believe that the shift took hold between 1950 and 1980. The term "science of complexity," as the new area is called (Kauffman, 1993, 1995), is awkward and calls for zipfing, but is precise. A great course of complexity by Parwani (2002) is available

online. I wish we could say *omnistics*. Just a look at the contents will give the reader the true taste of omnistics: it is about everything but the string theory.

As I would define its major attribute, omnistics is the study of objects in non-Euclidean spaces, namely, discrete topological spaces in which life, mind, and society have developed their overwhelming complexity and which we know not so much through instruments and gauges as through words and countable numbers. The geometry of this world is an open and partially renewable set of points with their neighborhoods. An object, including a sequence of words and an intricate idea, is represented by a sparse matrix. Distance in this space is quantifiable not with a tape measure but with integers corresponding to the minimum of elementary changes from one structure to another. The change of an object is a change in the matrix. It is governed either by non-equilibrium thermodynamics or by human intent. The open character of the matrix is incompatible with fundamentals of the "pre-complexity" physics and even most of mathematics, in whose systems nothing **new** can happen, although a new system can always be invented.

I believe that Pattern Theory is a welcoming portal into this entire area, where the chemist feels at home and so could the linguist. The growing vocabulary of the human race, in spite of the constant loss, is the best evidence that novelty exists.

Moreover, I believe that PT opens the door not only into the chemistry of language, but also into the chemistry of thought, i.e., the evasive and murky transition states of the mind. Grenander (2003) offers a look over the threshold of the mind and enthusiasts are welcome.

Language itself is the portal into Everything, where we can find chemistry, a cookbook, and a story about the origin of the portal itself …if we speak the language of Everything.

15. APPENDIX

15.1   Example of Chemicalese

The structure in Figure 14.1 belongs to C_{60}. The spherical molecule contains only carbon and belongs to the class of fullerenes, which gave birth to the entire area of nanotechnology. Its root morpheme fuller was derived from the name of the famous American architect Richard Buckminster Fuller who designed geodesic domes. The ending –ene means the presence of double bonds. The hexagons have the skeleton of benzene.
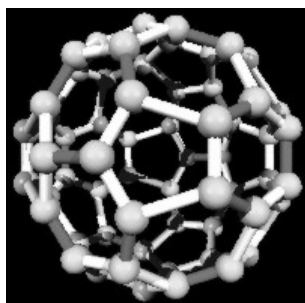
Figure 15.1 C_{60}, fullerene.
The double bonds are shown dark

The nomenclature name for C_{60} is:

Hentriacontacyclo[29.29.0.0.^{2,14}.0^{3,12}. 0^{4,59}.0^{5,10}.0^{6,58}. 0^{7,55}.0^{8,53}.0^{9,21}. 0^{11,20}.0^{13,18}.0^{15,30}. 0^{16,28}.0^{17,25}.0^{19,24}.0^{22,52}. 0$^{23.50}$.0^{26,49}.0^{27,47}. 0^{29,45}.0^{32,44}.0^{33,60}. 0^{34,57}.0^{35,43}.0^{36,56}.0^{37,41}.0^{38,54}.0^{39,51}.0^{40,48}.0^{42,46}]hexaconta-1,3,5(10),6,8,11,13(18), 14,16,19,21,23,25,27,29(45),30,32(44),33,35(43), 36,38(54),39(51),40(48),41,46,49,52,55,57,59-triacontaene [14].

Shortcuts are used also in chemical structures, as they were used in all hieroglyphic systems of writing, for example, *n*-Bu stands for $CH_3CH_2CH_2CH_2$— and A stands for adenine in nucleic acids and their fragments.

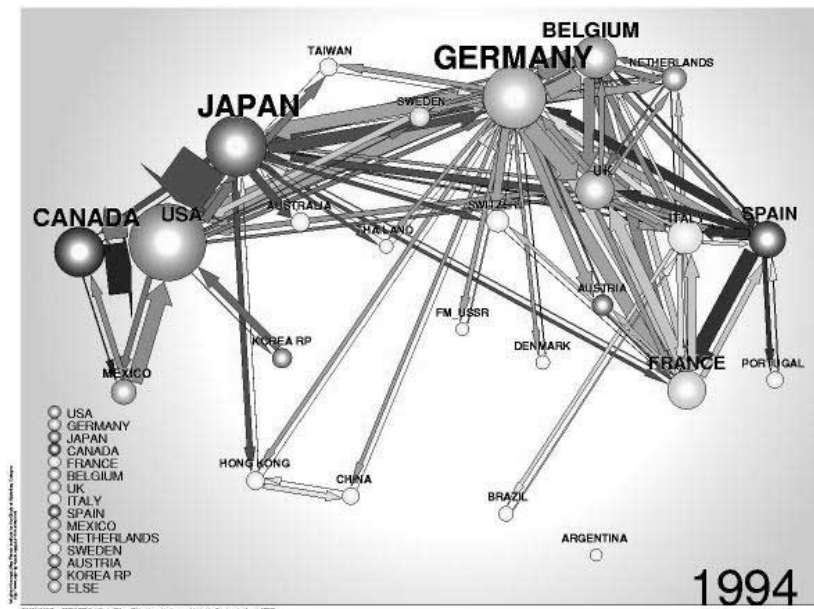## 15.2 Examples of real-life large configurations

Figure 15.2. World automobile trade in 1994. From Krempel (1999)
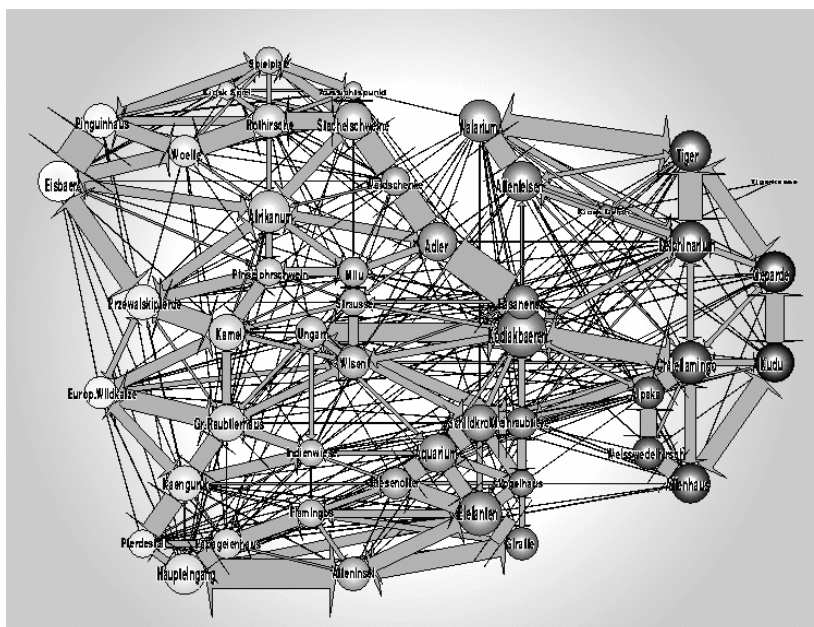

Figure 15.3. Visitors' traffic through Duisburg. From Krempel (WWW)

15.3 The chemical view of the world

The chemical view of the world is very much different from that of a physicist or a computer scientist, but it is only recently that chemistry began to realize its own extra-chemical abilities (Bhushan, 2000). Even before the advent of computers, chemical analogies inspired some landmark works in sociology and social psychology. In computer models of modern economics, an agent looks very much as an upgraded, animated, educated, greedy, and optimistic molecule. What a contrast with chemistry where any molecule dreams only about losing its energy.

Chemists have a simple view of complexity: it is built gradually, step-by-step. A large natural complexity can be built only as result of long evolutionary *history*. The reasons behind this belief are of kinetic nature: a collision of more than two particles is very rare. Nevertheless, complex proteins, as well as minds and societies, manage to assemble. The concepts of chemical mechanism and stepwise concatenation of transformations constitute a historical dimension of chemistry.

Since each elementary transformation is local, the simultaneous occurrence of a significant number of elementary transformations is improbable. In other words, the history of a natural complex system is Poissonian rather than Gaussian. The Gaussian system, synonymous with non-locality, in which any state can, theoretically, follow any other, always comes to an equilibrium while the Poissonian system just drags along from one rare event to another, between which nothing happens, and has no final state.

As far as social evolution is concerned, even wars, which seem to be most common events throughout human history, follow the Poisson distribution (Richardson, 1993), probably, because they are usually initiated by a decision of single person of a limited imagination. From the modern physical point of view, partially influenced by chemistry, all processes in the world are local (Mack, 2001, WWW).

There is a definite appeal in further exploration of the chemical paradigm in the vast area of mind, society, and language. Chemistry possesses the recursivity and generativity that is considered a unique property of human language noted in all general reviews (Calvin, 2000).

It is difficult to categorize chemistry as probabilistic because the large numbers of participating molecules ensure determinism. Yet determinism is absent from the foundation of chemical paradigm. The chemical process, once started either by a chemist or accidentally, runs its course on its own through a series of random events. The chemical system consisting of either large or small number of molecules searches for a new state by random collisions between molecules, "trying on" various combinations and mutual orientations. While classical robots and computers follow a program created by a programmer, the chemical system knows only fast parallel computing based on one operation: drawing a random number. In human reproduction, for example, conception is a reaction between just two molecules and it leads to spectacular results of microcosmic dimension.

Quantitatively, chemistry is mostly focused on time aspects, balance of energy, and irreversibility. At the same time, the meticulous, matter-of-fact representation of chemical events as a sequence of elementary acts, with the behavior of an individual molecule in the focus of attention, brings chemistry on the common descriptive grounds with humanities, especially, history, sociology, and even biography. In his *Selective Affinities*, Goethe (1988) was, apparently, the first to bring chemical symbolism into the chemistry of human relations.

The seemingly shapeless and amorphous appearance of chemistry, which often disheartens non-chemists, may obstruct the view of the mental workshop of a chemist who uses very sharp logical and measuring instruments and exercises a complete freedom of imagination in dealing with immense and incompressible complexity, as well as the experimental rigor, to rein it in. What streamlines the chemical thinking can be formulated as: everything is possible, but most of the possible is improbable and what is probable is local.

Chemistry shares the principles of atomism, composition, and metrics with Pattern Theory. Chemical systems can be regarded as systems of symbolic dynamics where atomic symbols combine and recombine. Therefore, we can hope to design open, evolving, and autopoietic (self-originated) symbolic systems within the framework of PT serving as a kind of meta-chemistry. In order to do that, we must preserve a certain

degree of chemical realism in the symbolic dynamics. In addition, the transition from one state to another must include a random component.

Next I am going to present in a very simplified and vulgarized form some basic ideas of theoretical organic chemistry.

Figure 13.1 consists of three rows (A, B, C) and three columns (I, II,III). The upper row of the Figure 13.1 (A) shows a typical example of a reaction mechanism taken from a chemical textbook and known as SN2, which means Substitution Nucleophilic Bimolecular.

A linguist is as little expected to be familiar with chemical theory as a chemist with transformational grammar. Nevertheless, some general properties of **A** can be seen on the surface.
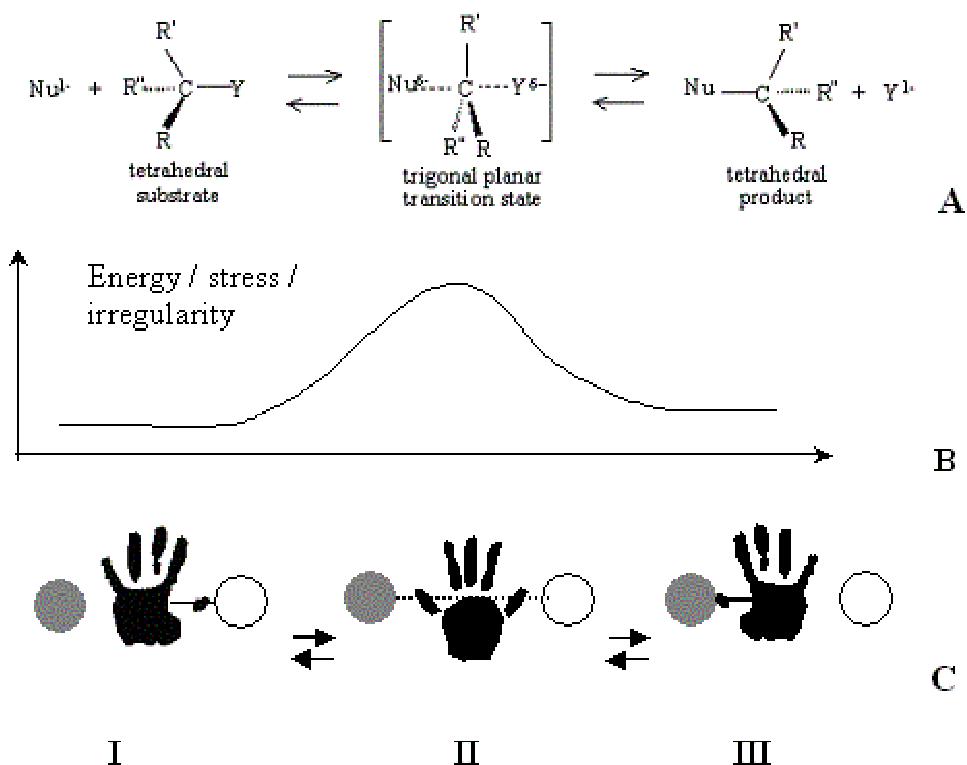

Figure 15.4 Substitution SN2

1. Except carbon **C**, no symbols of chemical elements can be seen. The symbols are of abstract nature, similarly to algebra and generative grammar. Symbols Y, Nu (nucleophil, negatively charged particle), R, $R^{'}$, and $R^{''}$ stand for particular combinations of atoms. The superscript indexes $1-$ and $\delta-$ stand for a unit and fraction of negative charge, accordingly. Charged particles usually have a much higher energy / instability than neutral ones, but vary in stability among themselves. The double arrows $\leftrightarrows$ symbolize the reversibility of the transformation: **I** $\leftrightarrows$ **II** $\leftrightarrows$ **III** . The stable states **I** and **III** are in an equilibrium with the transition state **II**.

Structures in **A** are 3D. The black wedges in **I** and **III** indicate that the bond is oriented toward the viewer and the broken wedges indicate the bond behind the plane of the drawing. All the other bonds lie in the plane.

2. The large square brackets around the transition state mean that it is in the process of change and is neither observable, nor stable.

3. Carbon normally has four valences, sometimes, two. The carbon atom in the transition state II has five bonds, two of which are shown by dotted lines to emphasize that they are irregular and temporary.

The middle part (B) visualizes the change of energy (stress, instability, irregularity) along the trajectory of the transformation. The initial and final states commonly (but not always) have somewhat different energy, so that the equilibrium is shifted toward the more stable state.

The lower row (C) is a 2D pictorial metaphor of what is going on during the 3D chemical transformation. All the events occur in the plane. The gray circle approaches the hand holding the white circle. In transition state II, the deformed (stressed, irregular) five-finger hand is in a precarious position, holding both circles. In the final state III, the hand holds the gray circle, but it is already a different hand.

The theory of transition state asserts that the speed of the transformation from one stable state to another decreases with the energy of the transition state. The latter forms a "barrier" that only molecules with sufficient energy can pass. Chemists use mostly the term **energy**, but the words **irregularity**, **stress**, and **deformation**, are also used in discussing regarding transition states.

The same transformation of substitution of Nu for Y, or one circle for the other can run through a different transition state, Figure 13.2 . It is known as **SN1**, Substitution Nucleophilic Monomolecular.
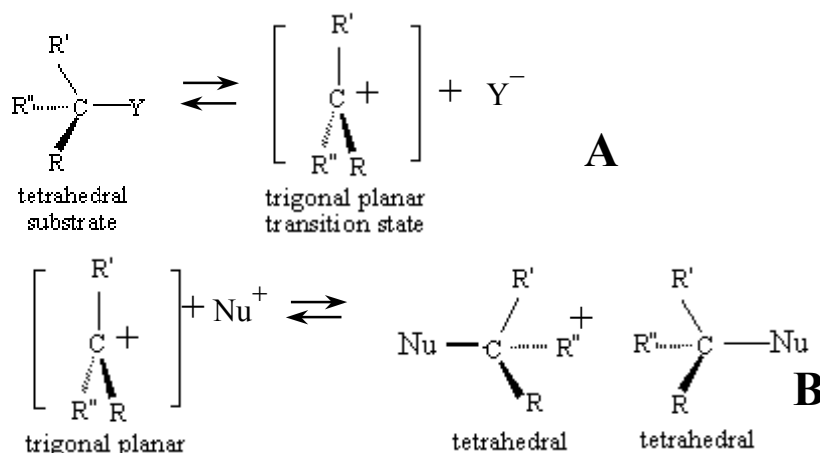

Figure 15.5 Substitution SN1

The initial state splits into Y^- and a crippled planar transition state with three bonds at the carbon atom (A). Next, the transition state can form a bond with either Y or Nu. Both can approach the transition state from either side of the plane. The 2D metaphor for this apparent mess is shown in Figure 12.3

We can start with either one of the four stable configurations outside the large square brackets. The four-finger palm can attach to either circle, which means that all four stable forms exist in an equilibrium.
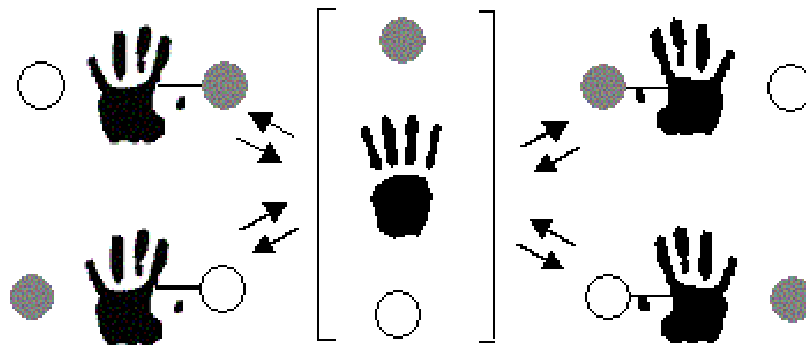
Figure 15.6 A metaphor of substitution SN1

Thus, the high energy transition states in **SN2** and **SN1** are irregular, unstable, and stressed because they are charged (in the eyes of a physicist) and have an abnormal number of bonds (in the eyes of a chemist).

Which mechanism takes place in reality depends on the part of reality which has been left out of the abstract picture: solvent, temperature, and the actual meaning of symbols Y, Nu, R, R', and R''. By studying the connection between the conditions of the transformation and its mechanism, chemistry acquired its modern theoretical sophistication even without observing the evasive transition state.

In short, the chemist who wants to predict or explain which alternative transformation will prevail in the short run, compares alternative transition states and gives the preference to the one with the less stressed (i.e., most probable) transition state. Generalizing this principle to the level of configurations in natural systems offers a new, **kinetic** approach to the dynamics of the complex systems built on the platform of life, including biological evolution itself, as well as border area between mind and society where language resides. The kinetic principle alone would do little good if not for another very general principal of natural complex systems, also inspired by chemistry: the change in any natural complex systems at any given time is mostly local. It means that most of complexity of the system is never involved in the change. This principle is well known to the historians of revolutions and, I believe, is applicable to language.

It is not accidentally that I selected human hands for a metaphor of substitution in chemistry. Molecules that, like hands, have no symmetry, possess chirality (handedness):

they are mirror reflection of each other. This property is of cardinal importance in biochemistry. The conclusions about the invisible transition states were drawn by chemists basing on the chirality of products. In **SN2**, a left-handed initial state reverses its chirality, while in **SN1**, the same state turns into a mixture of right and left final states.

A very similar method of oblique observations on utterances, leading to conclusions about unobservable thoughts, was first applied by no one but Zigmund Freud, a chemolinguist of a kind.

15.4   Program nean

```
% PROGRAM   nean
% input: NN (number of cycles), score (total score), S, words
% example S   =[1 1 5 2 1; 2 2 5 2 2; 3 1 6 1 2; 4 3 5 2 2; 5 2 3 1 1;
% 6 4 5 1 2;7 1 4 1 2; 8 2 4 1 2; 9 1 3 1 2; 10 3 4 1 2];
% example: words= ['Ug   '; 'Og   ';  'bone '; 'cave '; 'give '; 'grim
'];

LW=length(words(:,1)); LS=length(S(:,1)); LWS=LW-1;
E= 0;   A=[];
ns=0; %ns: number of selected strings
nw=0; %nw: number of different string in the output
tic
for n=1:NN;E2=0;
     p = randperm (LW); %random sequence of words
     DP=zeros(1,2);DS=zeros(1,2);
     E2=0;
     for i=1:LWS; DP=[p(i),p(i+1)]; %pair of neighboring words
          for j=1:LS  %number of doublets

        SD=S(j,:,:); DS= [SD(2),SD(3)]; %pair of generators in memory
        DSR=[SD(3),SD(2)];

          if ((DS == DP)|((S(j,5)==2)&(DSR==DP))), E2=E2+S(j,4);end ,
                     %compared with the pair of G in permutations
                     % score added to the total
           end
     end
             %if E2>E, E=E2; E
          if E2==score, p; A=cat(1,A,p);ns=ns+1; end
end,
n;A=unique(A,'rows');l=length(A(:,1));
W=[];
for w=1:l, W=A(w,:);WW=[];
```

```
    for v=1:LW,
        WW= cat(2,WW,words(W(v),:));
    end,
    disp (['      ',WW]);nw=nw+1;
end
t=toc; t=t/60;
disp ('     ');
disp ([ '  number of cycles: NN = ',int2str(NN) ,  ', run time: t='
,num2str(t,2),    ' min, ' ]);
disp([' score= ', num2str(score,2),', score matches ', int2str(ns), ',
number of different strings ', int2str(nw)]);
```

REFERENCES

See also http://spirospero.net/complexity.htm

Adami, Christoph. 1998. *Introduction to Artificial Life*. New York, NY: Springer.

Aitchison, Jean, 1881. *Language Change: Progress or Decay?* New York: Universe Books.

———. 1996. *The seeds of speech: Language origin and evolution*. Cambridge: Cambridge University Press.

Ashby, W. Ross. 1960. *Design for a Brain: The Origin of Adaptive Behavior*, 2nd Ed., New York: Wiley. Originally published in 1952.

———. 1964. *An Introduction to Cybernetics*, London: Chapman & Hall. Originally published in 1956.

A short gracious bio of Ashby: http://www.isss.org/lumashby.htm)

Baker, Mark C. 2001. *The Atoms of Language*. New York: Basic Books.

Baker, Mark and Travis, Lisa. 1995. WWW. *Mood as Verbal Definiteness in a "Tenseless" Language* equinox.rutgers.edu/people/faculty/ours/baker/mohawk-mood-prt.ps

Bhushan, Nalini and Rosenfeld, Stuart, Editors. 2000. *Of Minds and Molecules: New Philosophical Perspectives on Chemistry*. Oxford, NY: Oxford University Press.

Bickerton, Derek. 1981. *Roots of Language*. Ann Arbor: Karoma Publishers.

———. 1990. *Language & species*. Chicago : University of Chicago Press.

———. 1995. *Language and Human Behavior*. Seattle: University of Washington Press.

———. 2003. Symbol and Structure: A Comprehensive Framework for Language Evolution. See: Language Evolution, 2003, pp. 77-93.

Billard, Aude. 2002. Imitation: a means to enhance learning of a synthetic proto-language in an autonomous robot. In *Imitation in Animals and Artifacts* (Dautenhahn K. and Nehaniv, C.L., eds), pp.281–310, Cambridge, Mass: MIT Press.

http://asl.epfl.ch/aslInternalWeb/ASL/publications/ uploadedFiles/ABCyberneticsSystems2000.pdf

Bonabeau, E., Dorigo, M., and Théraulaz, G. 1999. *Swarm intelligence: from natural to artificial systems,* New York: Oxford University Press, 1999.

Bottéro, J. 1992. *Mesopotamia: Writing, Reasoning, and the Gods,* Chicago: The University of Chicago Press.

Bourbaki, Nicolas. 1968. *Elements of Mathematics: Theory of Sets*, Boston: Addison-Wesley, originally published by Hermann (Paris), 1968, p.259-382.

Calvin, William H. and Bickerton, Derek. 2000. *Lingua ex Machina: Reconciling Darwin and Chomsky with the human brain.* Cambridge, Mass.: MIT Press.

Cangelosi A. and Parisi D., Editors. 2002. *Simulating the Evolution of Language*. London: Springer.

Cangelosi, A., A. Greco and S. Harnad (2002). Symbol Grounding and the Symbolic Theft Hypothesis. In Cangelosi and Parisi (Eds. 2002), pp. 191-210.

Cavalli-Sforza, Luigi Luca. 2000. *Genes, People, and Languages*. New York: North Point Press.

Chomsky, Noam. 1997. Language and Mind: Current Thoughts on Ancient Problems (part I). Paper presented at Univ. de Brasilia, Nov. 25, 1996. Published in: Universidade de Brasília, *Pesquisa Lingüística* 3(4), 1997. http://www.linguistik.uni-bonn.de/Download/nc-pap.doc

———. 2002. *On Nature and Language*, Cambridge: Cambridge University Press.

Corballis, Michael C. 2002. *From hand to mouth: the origins of language*. Princeton: Princeton University Press.

Dawkins, Richard. 1989. *The Selfish Gene*. New Edition. Oxford, New York, etc.: Oxford University Press.

De Landa, Manuel. 1998. *A Thousand Years of Nonlinear History*, Cambridge, Mass.: The MIT Press.

Devlin, Keith. 2000. *The Math Gene: How Mathematical Thinking Evolved and Why Numbers Are Like Gossip*. New York: Basic Books.

Dunbar R.I.M. (1996/98) *Grooming, Gossip, and the Evolution of Language*. London: Faber & Faber/Cambridge, Mass.: Harvard University Press.

Ege, Seyhan. 1989. *Organic Chemistry*. Lexington, Mass.: D.C.Heat and Company.

Eigen, Manfred and Winkler, Ruthild. 1993. *Laws of the Game*. Princeton, New Jersey: Princeton University Press. Originally published in Germany in 1965.

Eigen, Manfred. 1971. Selforganization of Matter and the Evolution of Biological Macromolecules, *Die Naturwissenschaften*, **58,** 465-522 (1971).

———— 1977, 1978. The Hypercycle*, ibid*, 64, 541-565 (1977), 65, 7-1 (1978).

Eigen, Manfred and Schuster, Peter. 1979. *The Hypercycle - A Principle of Natural Self-Organization*, Berlin: Springer-Verlag.

Eldredge, N. 1985a. *Time Frames*. New York: Simon and Schuster.

_____. 1985b. Unfinished Synthesis. New York: Oxford University Press.

Eldredge, N. and Gould, S.J. 1972. Punctuated equilibria: an alternative to phyletic gradualism, in *Models in Paleobiology*. T.J.M. Schopf (ed.). San Francisco: Freeman, Cooper. pp. 82-115.

Eyring, H. and Polanyi, Z. 1931. *Phys. Chem*. B **1**2, 279 (1931).

Fodor, J.A. 1976. *The Language of thought*, New York: Thomas Y Crowell.

Goethe, J. W. v. 1988. *The Sorrows of Young Werther. Elective Affinities. Novella. Goethe's Collected Works,* Vol. 11. New York: Suhrkampf Publishers, Inc.

Greenberg, Joseph H. 1990. "The Patterning of Root Morphemes in Semitic" In: *On Language: Selected Writings of George H. Greenberg,*. Editors: Keith Denning and Suzanne Kemmer. Stanford: Stanford University Press, p. 365.

Grenander, Ulf. 1976. *Pattern Synthesis. Lectures in Pattern theory*, Volume 1. New York: Springer-Verlag.

————. 1978. *Pattern Analysis. Lectures in Pattern theory*, Vol. II. New York: Springer.

————. 1981. *Regular Structures.Lectures in Pattern theory*, Vol. III. New York: Springer.

————. 1993. *General Pattern Theory. A Mathematical Study of Regular Structures*, Oxford, New York: Oxford University Press.

————. 1995. *Elements of Pattern Theory*. Baltimore: Johns Hopkins University Press.

————. 2003. *Patterns of Thought*.   www.dam.brown.edu/ptg/REPORTS/mind.pdf

Halloran, John A. WWW. *The Proto-Sumerian Language Invention Process*. http://www.sumerian.org/prot-sum.htm.

Hammond, George S. 1955. A Correlation of Reaction Rates. 'Hammond Postulate.'

J. Am. Chem. Soc. , *77*, 334. See a historical sketch by Amanda Yarnell:
http://pubs.acs.org/cen/science/8120/print/8120sci2.html

Harnad, Stevan. 1990. The Symbol Grounding Problem. *Physica* D 42: 335-346.
http://eprints.ecs.soton.ac.uk/archive/00008175/01/sgproblem1.html

Hauser, Marc D. 1996. *The evolution of communication*. Cambridge, Mass. : The MIT
Press. c1996

Hoede, C. 2003. University of Twente. The Netherlands. Department of Applied
Mathematics. Memorandum No 1682. *Basic notions in mathematics: On the
"graph" in particular and on ontology in general*.
http://www.math.utwente.nl/publications/2003/1682.pdf

Hurford, James. R. 2003. The Language Mosaic and its Evolution. See *Language
Evolution*, 2003, pp. 38-57.

IUPAC. 1993. *IUPAC, Commission on Nomenclature of Organic Chemistry. A Guide to
IUPAC Nomenclature of Organic Compounds (Recommendations 1993)*. Oxford:
Blackwell Scientific publications. http://www.acdlabs.com/iupac/nomenclature/

Jackendoff, Ray. 1992. *Languages of the Mind: Essays on Mental Representation*.
Cambridge, Mass: The MIT Press.

———. 1987. *Consciousness and the computational mind*, Cambridge, Mass: The
MIT Press.

Jaffe. 2000. Jaffé, Charles, D. Farrelly, D. , and T. Uzer, 2000. Transition State Theory
without Time-Reversal Symmetry: Chaotic Ionization of the Hydrogen Atom.
Physical Review Letters, **84,** 610-613.

Johnson, Mark. 1987. *The body in the mind : the bodily basis of meaning, imagination,
and reason* . Chicago : University of Chicago Press.

Kauffman, S. (1993). *The Origins of Order: Self-Organization and Selection in evolution.*
 New York, Oxford: Oxford University press

———. (1995). *At Home in the Universe: The Search for Laws of Complexity*, New York:
Oxford University Press.

Kirby, Simon and Christiansen, Morten H. 2003. From Language Learning to Language
Evolution. See: Language Evolution (2003, p.274).

Komarova, Natalia L. and Nowak, Martin. A. 2003. Language, Learning, and Evolution. See: Language Evolution (2003), pp.316-337.

Krempel, Lothar and Plümper, Thomas. 1999. International Division of Labor and Global Economic Processes: An Analysis of the International Trade in Automobiles. *Journal   of World-Systems Research*, Vol V, 3, 487-498. http://csf.colorado.edu/jwsr

Krempel, Lothar. WWW. *A Gallery Of Social Structures.* http://www.mpi-fg-koeln.mpg.de/~lk/netvis.html

Kuhn, Thomas. 1962. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Lakoff, George and Johnson, Mark. 1980. *Metaphors we live by*. Chicago : University of Chicago Press.

Lakoff, George. 1987. *Women, fire, and dangerous things : what categories reveal about the mind.* Chicago : University of Chicago Press.

Language Evolution. 2003. Edited by Morten H. Christiansen and Simon Kirby. Oxford: Oxford University Press.

Language Origin. WWW. *Language Evolution and Computation Bibliography*, Web Portal maintained by Jun Wang at University of Illinois at Urbana-Champaign. http://www.isrl.uiuc.edu/~amag/langev/index.html Full reference: http://www.isrl.uiuc.edu/~amag/langev/full_reference.html

Laszlo, Pierre. 1997. *Belaboring the obvious: Chemistry as sister science to linguistics*. Conference *Language as an analogy in the natural sciences*, Munich, November 20-23, 1997. http://www.pierrelaszlo.net/science_writings/language.htm

Lieberman, Philip. 1998. *Eve spoke : human language and human evolution*. New York: W.W.Norton.

Lightfoot, David. 1999.The Development of Language: Acquisition, Change, and Evolution. Mass. and Oxford: Balckwell Publishers.

Liu, Xiaodong. 2002. *The Chemistry of Chinese Language*. Ph.D. thesis. University of Twente, The Netherlands. www.ub.utwente.nl/webdocs/tw/1/t0000022.pdf

Lucretius Carus, T. (1958). *De Rerum Natura. Of the Nature of Things* /

translated into English verse by William Ellery Leonard. New York : Heritage Club. http://www.gutenberg.net/etext/785

http://classics.mit.edu/Carus/nature_things.html

Mack, Gerhard. 2001. Universal Dynamics, a Unified Theory of Complex Systems. Emergence, Life and Death. *Communications in Mathematical Physics*, **219**, No.1, (141 – 178).

———. WWW. Web links to other works: http://lienhard.desy.de/ ; http://lienhard.desy.de/call.shtml?sy_1; http://lienhard.desy.de/call.shtml?sy_3.

Parwani, Rajesh R. WWW. 2002. Complexity. A course. http://staff.science.nus.edu.sg/~parwani/c1/book.html

Pauling, Linus (1954). *Modern structural chemistry*. Nobel Lecture, December 11, 1954. www.nobel.se/chemistry/laureates/ 1954/pauling-lecture.pdf

Peirce, C.S. (1992). *The essential Peirce*,Vol. 1 (1867-1893). N.Hauser & C. Kloesel (Eds.). Bloomington, IN: Indiana University Press.

Pinker, Steven. 1994. *The Language Instinct.* New York: William Morrow.

———. 2003. Language as an Adaptation to the Cognitive Niche. In: Language Evolution, 2003.

Pinker, S. and Bloom, P. 1990. Natural language and natural selection. *Behavioral and Brain Sciences* 13 (4): 707-784.

Prigogine, Ilya and Stengers, Isabelle. 1984. *Order out of Chaos* . New York: Bantam. Also: Nicolis, G. and Prigogine, I. 1989. *Exploring Complexity.* New York: W.H.Freeman. Stengers, I. and Prigogine, I. 1997. *The End of Certainty : Time, Chaos, and the New Laws of Nature*, New York: Free Press.

Quillian, M.R. 1968. Semantic memory. In: Minsky, M. (Ed.) *Semantic Information processing.* Cambridge, Mass: MIT Press 227-270

Rosen, Robert. 1991. *Life Itself* . New York: Columbia University Press.

——— . 2000. *Essays on Life Itself* . New York: Columbia University Press.

Raspe, Rudols Eric. WWW. *The Surprising Adventures of Baron Munchausen*. Published in 1895 (first appeared in1785). http://homepage.ntlworld.com/forgottenfutures/munch/munch.htm#ch6

Also at Project Gutenberg:

http://onlinebooks.library.upenn.edu/webbin/gutbook/lookup?num=3154

Ratnaparkhi, Adwait. 1998. *Maximum Entropy Models for Natural Language Ambiguity Resolution*. Dissertation. University of Pennsylvania.

http://www.ai.mit.edu/courses/6.891-nlp/READINGS/adwait.pdf

Richardson, Lewis Fry.1993. *Collected Papers of Lewis Fry Richardso*n. Edited by Oliver M. Ashford, *et al.* New York: Cambridge University Press.

Shannon, C.E. 1948. A mathematical theory of communication. *Bell Sys. Tech. Journal*, **27**, 1948.

Searls, David B. 1992. The linguistics of DNA. *American Scientist*, 80 (6), 579-591.

Senyak, Josh. WWW. *Shannonizer*, http://www.nightgarden.com/shannon.htm).

Sowa, John F. 2000. *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Pacific Grove, CA: Brooks Cole Publishing Co.

———. WWW. John F. Sowa's web site http://www.jfsowa.com

Spinoza, B. de. (1994). *A Spinoza Reader : the Ethics and other works* . Princeton, N.J. : Princeton University Press.

Stasko, John. T. WWW. Personal Page. Georgia Institute of Technology.

http://www.gvu.gatech.edu/gvu/people/faculty/john.stasko/

Studdert-Kennedy, Michael and Goldstein, Louis. 2003. "Launching Language: The Gestural Origin of Discrete Infinity." In *Evolution of Language*, pp. 235-254.

Tarnopolsky, Yuri. 2003. *Molecules and Thoughts: Pattern Complexity and Evolution in Chemical Systems and the Mind* . WWW:

www.dam.brown.edu/ptg/REPORTS/MINDSCALE.pdf

Or: http://spirospero.net/MINDSCALE.pdf

———. 2003. *Transition States in Patterns of History*. WWW.

http://spirospero.net/HistMath1.pdf

Terrace, H.S. 1970. *Nim*. NY: Knopf .

Thom, René. 1975. *Structural Stability and Morphogenesis: An Outline of a General Theory of Models*. Reading, Mass.: W. A. Benjamin, Inc.

Tikki, WWW. *Tikki Tikki Tembo*. Chinese folk tale.

A. http://www.pitt.edu/~dash/tikki.html , B. http://www.fairrosa.info/disc/tikki.html

Tomasello, Michael. 2003. On the Different Origins of Symbols and Grammar. See: Language Evolution (2003), pp. 94-110 .

Uriagereka, Juan. 1998. *Rhyme and Reason: An Introduction to Minimalist Syntax.* Cambridge, London: The MIT Press.

Varshavskaya, Adamina. 2002. Behavior-based early language development on a humanoid robot. In *Proc. of the 2nd Int. Conf. on Epigenetics Robotic*s, pages 149–158.

Viswanathan, G.M. *et al* . 2002. Lévy flight search patterns of wandering albatrosses. *Nature* 381, 413 - 415 (2002). Other authors: V. Afanasyev, S. V. Buldyrev, E. J. Murphy, P. A. Prince & H. E. Stanley

Vlasov, L. and Trifonov, D. 1970. *107 stories about chemistry*. Transl. from the Russian by D. Sobolev. Moscow: Mir http://www.todayinsci.com/stories/story020.htm

Waddington, C. 1957. *The Strategy of the Genes*. London: Allen and Unwin.

Wigner, Eugene. 1961. The probability of existence of a self-reproducing unit. In *The Logic of Personal Knowledge* (E. Shils, ed.) Glencoe, Ill: Free Press.

Willems, Marc. 1993. *Chemistry of Language; A Graph-theoretical Study of Linguistic Semantics*. Ph.D. thesis. University of Twente, The Netherlands.

Zhongwen. http://zhongwen.com/ . See also: http://www.chinaknowledge.de/Literature/script.htm#top

Zipf, George K. 1949. *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. Cambridge, Massachusetts: Addison-Wesley.

———— . 1965. *Psycho-Biology of Languages* . Cambridge, Mass: MIT Press. (First edition: Houghton-Mifflin, 1935). W.Li compiled a comprehensive bibliography on Zipf's Law at http://linkage.rockefeller.edu/wli/zipf/ .