The Three Little Pigs : Chemistry of language acquisition

Yuri Tarnopolsky
2005

Abstract

While problems of emergence are often treated in terms of **complexity** reaching a certain threshold, a different approach can be advocated in terms of **simplicity**. Inspired by fundamental principles of chemistry, it looks for very simple systems in tiny phase spaces, governed by simple local rules and capable of increasing their complexity by simple steps. It is hypothesized that the initial stage of language acquisition is a natural example of emergence through simplicity.
It might be difficult to reconcile evolving systems with the axiom of closure, which is the keystone of mathematics but leaves no place to evolutionary novelty. Chemistry, however, accommodates the concept of novelty quite well.
This e-paper continues, in a freewheeling fashion, the examination of language as a quasi-molecular system from the point of view of a chemist who happens to ask, "What if the words were atoms?" It further explores the parallel between cognitive and chemical systems. A unified conceptual groundwork for chemistry and linguistics, as well as cognition and all other discrete combinatorial systems, is borrowed from the atomism of Pattern Theory (Ulf Grenander).
As an illustration, the text of *The Three Little Pigs* is decomposed into triplets of adjacent words and some local principles of generator identification and categorization are examined. The principle of local equilibrium between the category and its entries is discussed against the background of basic chemical ideas.

CONTENTS

If words were atoms: an introduction

This e-paper continues the examination of language as a quasi-molecular system from the point of view of a chemist who, inspired by the book by Mark C. Baker *The Atoms of Language* [1] , quite seriously asks, "What if the words were atoms?"

The chemist happens to be myself. My motivation comes from the time when I, a student at a chemistry department in the mid-1950's, learned for the first time about Norbert Wiener and his previously forbidden in the Soviet Russia cybernetics. By the same time, facing a large body of chemical publications and starting to develop some passive skills in foreign languages, I felt the pull of the linguistic cosmos. Thirty years later I had learned, by mere accident, about Ulf Grenander and his Pattern Theory (PT), [2, 3] and the theory seemed to friendly embrace all available to me knowledge.

I want to look at cognition with the eyes of neither a mathematician, nor an engineer, nor a linguist, nor a cognitive scientist, but a chemist. My intuition tells me that chemistry may be relevant at least for one reason. Molecules and phrases, both observable, are configurations in PT.

What chemistry can contribute to the area is, first of all, the unique experience with discrete structural change over time, which hardly any other science possesses in comparably pure form. Another little used angle of vision, also inspired by chemistry, is the evolutionary one, but, again, not in the common sense. It is not that "everything evolves" but that everything **grows** on a historical scale from very simple structures by very simple steps up to an overwhelming complexity. Thirdly, the eye-catching but often misunderstood principle of catalysis has a very general extra-chemical meaning.

There is yet another subtle reason. We do not know whether it is essential for the brain and its cells to be a chemical system. If we knew, we could probably understand why human mind has been so stubbornly resisting any integral computer simulation for at least half a century of computer science.

Chemistry, as no other science, can efficiently master enormous complexity by simple means and a Spartan stock of ideas. Taking to account the **chemical** origin of life and tracing the origin of species, mind, and society back to the chemical cradle, we may expect to notice in the oblique light some new shadows invisible in the frontal glare of computer science.

I ended my previous e-paper [4] with a tentative Appendix as an illustration of some chemistry-inspired concepts regarding the first language acquisition by children and the Poverty of Stimulus argument. I decomposed a fragment of *The Three Little Pigs* into 1-neighborhoods of words, i.e., the word and its right and left neighbors, and tried to derive syntactic classification in a non-algorithmic manner through primitive local operations, ignoring the impressive achievements of Neural Networks, algorithmic Part-of-Speech tagging, corpus-based and context-based categorization, and other contemporary approaches to language processing. As before, I am interested here only in exploring the parallels between language and chemistry in the light of Pattern Theory, but always from the position and with habits of a chemist.

Further pursuing the program **If Words Were Atoms**, I am making here the next step within a larger program **The Chemical View of the World**, see [4, 5] . where some relevant literature was collected from distant domains of knowledge. It is the larger program that could be an excuse for numerous digressions from the immediate subject and references to distant times and places on the map of knowledge.

The main idea can be presented in the following way. The natural complex systems, especially, life, mind, language, society, and culture, all emerged at some elusive point. The dominating point of view is that order emerges in a dynamical system when its parameters exceed some threshold [6]. This is certainly true, as far as the origin of order in some physical and chemical systems is concerned, but there is no way to derive a particular kind of order, for example, origin and evolution of language, from the general systemic ideas about order. We can derive regularities from observing particular systems, but we cannot derive particular systems from the regularities, unless we have a conceptual bridge between particularities and regularities. This bridge naturally exists in

chemistry and follows from the idea of atomism. Pattern Theory is, in this sense, a meta-chemistry, i.e., a mathematical foundation for the study of atomistic structure.

If the prevalent direction in the study of emergence starts with **complexity**, the alternative idea is advocated here in terms of **simplicity**. The science of simplicity, possibly complementary to the science of complexity, starts with very simple systems in tiny phase spaces, governed by simple local rules and capable of increasing their complexity by simple steps. On its progress toward complexity, **simplicity** is not bound by the axiom of closure that makes mathematics and logic possible. Chemical systems, having served as the cradle of life and life's subsequent expansions into mind and society, are the natural source of such ideas. It is hypothesized that language acquisition is one of possible illustrations, too. Unlike the origin of life and society, not to mention the universe, it is perfectly observable in small children.

I suggest that the simple origin is a necessary condition of unfolding of any complex system and it should be included into the definition of the complex open system and taken to account in designing realistic simulations of life, mind, and society.

The paper consists of two parts: one is about principles and the other one with illustrations. As far as the style of this paper is concerned, if it appears tousled, it tells about the excitement of the adventure.

Let us think

PART 1　The mind and the flask

Name your friends

However far from chemistry, the part of computer science known as unsupervised learning is most closely related to the subject of this essay. From the distance I can limit myself to only a few allegorical remarks. The distance, however, is beneficial because it opens a panoramic view not encumbered by detail. Operating with symbols, both computer science and linguistics possess extended overlapping areas which are compartmentalized into something like the sealed off sections of a submarine: they do not communicate unless adjacent.

As an example, the prototype of Eleanor Rosch [7], template of Ulf Grenander [2], and prototypes of Shimon Edelman [8] all acknowledge the genes of Ludwig Wittgenstein in their chromosomes, but hardly ever gather for a family reunion.

The bits of data in symbolic computation are brought together by the processor for a short time and large arrays have to be explored bit by bit. The alternative idea of non-symbolic Neural Nets with parallel processing took shape through the concept of an advanced Perceptron in which an intermediate layer between inputs and outputs is fully interconnected, so that everybody is everybody's neighbor. The interconnectivity is a parallel materialization of the consecutive symbolic loop.

The old adage "Tell me who your friends are and I will tell you who you are" expresses the essence of the main paradigm of non-symbolic computation: the atom of cognition interacts only with its neighbors. Following the adage, I will list some of my own intellectual neighbors.

All possible variations on the non-symbolic connectionist theme constitute the large body of Neural Networks (NN) where work has a double motivation: to make an **artificial** device which performs human functions and to understand how **natural** intelligence performs its functions.

There is a great assortment of NN, immersed into the protoplasm of technicalities.. Not being an expert, I can only make very superficial and subjective remarks. There is abundant online literature on particular types of nets. Some reviews, see [9]

The Hopfield Networks, inspired by thermodynamics, are the closest possible approximation to chemical equilibrium. The crucial difference is that energy in chemistry is approximately additive over varying structures, while in NN, the structure is mostly constant and energy can be locally tuned up at the nodes regardless of conservation.

Self-organized maps (SOM) of Teuve Kohonen and his school, as it seems to me, model both the symbolic consecutive search and the interconnectivity of Perceptron by massive **global** computation over the data, resulting in revealing the much sparser topology of connectivity.

The work of Shimon Edelman [8], with its wide and inter-disciplinary perspective, witty illustrations, and attention to the concepts of novelty, open sets, and

metric relations in representation spaces stands apart in **unsupervised learning**. Besides, his recent project ADIOS (Automated DIstillation of Structure) strikes a chord in the heart of a chemist by the word **distillation**. The core of Edelman's approach, inherited from the earlier and most creative stages of cybernetics, seems to be the use of feature-hunters that look for particular **local** features in the input, a kind of a squadron of Dr. Watsons, but without Sherlock Holmes.

Having started within the framework of NN, Stephen Grossberg and Gail Carpenter [10] have compounded a large volume of work on **adaptive resonance theory** (ART). Although it is often difficult but scarcely explained, ART also stands apart because of its far-reaching realism, straightforward handling of **novelty** by adding new categories, conservative principle (normalization), and gravitation toward **behavioral success,** as Grossberg calls what I prefer to call **homeostasis**. ART, in a way, depends on the **history** of the system and uses elements of Sherlock Holmes' ability to connect the dots by differentiating the essential from the accidental. Other fine properties of ART are its **tunable radius** of attention and **sharp selection** from probability distributions ("winner takes all"). ART appears to be the richest, constructive, most realistic, and capable of evolution theory of this kind. Besides, it strikes another chemical chord with the word **resonance**.

The arising problem with NN seems to be that in the absence of a brain-like parallel hardware, a huge volume of computation—or hardware—is needed for simulation.

Another big intellectual field has been deeply ploughed by Jeffrey Elman [11A], whose major contribution, regarding language acquisition, from my point of view, is the clearly stated idea of the necessity of simple origins for natural complex systems.

The term **bootstrapping** [13A] is used in more than one meaning. For example, an area of experimental linguistics, unavoidably leading to the name of Peter Jusczyk [12], studies the language acquisition by small children from phonological environment, while syntactic bootstrapping (Lila Gleitman [13C,13D]) looks at the lexical environment.

As soon as I had opened Peter Jusczyk's "*The Discovery of Spoken Language*" [12A], I felt myself at home in the experimental science of spoken language with the "chemical" exuberance of objects, observability, controlled conditions, reproducibility and testability of results, a lot of small incremental works that gradually reveal big and reliable principles, inventive working techniques, which are as important as principles, and with possibility of consensus. I could also feel some background bitterness of the author that experimental psychology of language acquisition was regarded an off-stream area drowned in all the talk about language with ears wide shut. Peter W. Jusczyk's unordinary and productive life was, unfortunately, short [12B]. His delightful book is still available.

In computer science, bootstrapping means using a very small initial structure to extract a much larger one from extensive data. There are also works where bootstrapping straddles the border between computer science and linguistics. Bootstrapping as applied to a linguistic corpus works similarly to Kohonen nets and requires initial seed data [13E].

Bootstrapping is highly relevant for the current paper as a test of ideas, but it deserves a separate review. So is a more recent direction of **novelty detection,** which branches off the traditional and well-seasoned NN, rising fundamental questions about the nature of novelty. For my own view of the subject see [5].

I might have missed other distinctive directions.

In my eyes, the experimental study of language acquisition can have very important consequences for general system theory because unfolding of language is one of a few processes of evolutionary genesis we can see on earth. As much as I believe in extra-chemical message of chemistry, I believe in extra-linguistic message of linguistic. Both are chapters in the future Pattern Theory of Evolution.

The particular designs and implementations of cognitive computer science may become obsolete with time, but each of the schools mentioned above has added a constructive and important idea to the stock of the earlier contributions. It is the ideas and not technicalities that will go into the foundation of consensus when the dust over the

construction site settles. Linguistics is far from that stage: "Linguistics is a field that is known for controversy" [14].

While everybody in cognition, computer science, and probably even on Wall Street uses the word "pattern" as often as a teenager says "cool," the most general and far-reaching **Pattern Theory** (Ulf Grenander) seems all but unnoticed by linguists. The novelty it introduces in cognition is not only a universal mathematical language for patterns but also the potential ability to associate a structure with an individual "node" of connectionist concept in the same way a chemist associates the structure of phenyl salicylate with the word "aspirin" and calculates the energy of aspirin from its formula. PT introduces quantitative meta-chemistry into cognitive sciences, but unless you are a chemist, it is easy to overlook. My personal enthusiasm about PT is to a great extent fueled by the role of PT as a wide bridge between the contrasting mathematical and chemical modes of thinking.

I am not in a position to analyze or criticize the mainstream of computational linguistics. My impression is that every approach mentioned above intrinsically counts on a participation of not just a homunculus but a full blown educated adult human with the I.Q. of at least Sherlock Holmes' himself, either in the preparatory, or the operational, or the interpretive phase. For example, novelty is what is presented to the system by the **operator** for the first time, the self-organized maps require some initial **selection** of seed elements to be put on the map, the supervised systems are **trained** or need some help, bootstrapping needs some stock of "**seeds**," and, quite unbelievably, the corpus of *The Wall Street Journal* is considered natural language.

Finally, I would refer to the unfinished work of Gerhard Mack [15], theoretical physicist, and Clark Barrett [16A], anthropologist, on enzymatic computation and cognitive equilibrium, which appeal to the very soul of a chemist by turning to catalysis. Both use the term computation more like cognition and enzyme more like a metaphor, deviating from what chemistry understands by catalysis (see [5]). This, however, by no means nullifies their ideas inspired by chemistry. While Gerhard Mack is focused on the universal locality of natural phenomena, Clark Barrett comes with an opposite intent,

using the chemical paradigm to show how chemistry could overcome the rigid mechanistic locality of cognitive models. It is the relation between chemistry and cognition that I am trying to explore here from the chemist's perspective. Indeed, if all cognitive interactions are local, no creativity is possible. I myself tried to illustrate by modest computer experiments that such long jumps, paradoxically, can be a result of strictly local interactions (*Molecules and Thoughts*, [5]).

Drowning by numbers

Being an outsider in computer and cognitive sciences and linguistics, I am overwhelmed by the volume of literature in the area. I doubt a linguist would be equally overwhelmed by chemistry because chemistry, unlike cognition, enjoys at a high degree of consensus regarding theoretical principles and most experimental technicalities. One can find them all in a couple of textbooks.

The evolving role of Internet in science opens an unlicensed way to fish for the ideas and even technicalities in the enormous literature, while most books and older works still need to be rummaged for in the libraries.

I realize that I might have missed a large number of important publications, but it is absolutely beyond my capabilities to explore even a small part of relevant literature with its intricate geography. It includes two giant continents of computer modeling and experimental psychology/neurophysiology, separated by a narrow Isthmus of Panama with little traffic between them. There is also another large and lush continent of anthropology, hardly noticed by the other two, into which, by the way, Clark Barrett's works [16B] offer a fascinating peek. Then, of course comes the huge Eurasia of linguistics with its Himalayas and Gobi Desert difficult for those who speak only a couple of languages. And, of course, formal and experimental linguistics are separated by the Ural Mountains. Where is Italy and where is Siberia, decide for yourself.

In spite of its shifting sands, the Internet, by its very searchable nature, can provide a reliable access to the cream of cognitive ideas across all the oceans and mountains.

I give here only a minimal number of references, preferably to online publications, because, although the sites are mortal and can become senile, there is always somebody who wants to share knowledge and achievement in response to a query through **Google** or **CiteSeer**. Bibliography sites and Ph.D. dissertations, now often available on the web, serve as rich depositories of concentrated food for thought.

There is yet another factor in the favor of the Web. For example, a wonderful introductory review of Walter Klinger "Learning Grammar by Listening" [17] was published in *Academic Reports of The University Center for Intercultural Education*, The University of Shiga Prefecture, No. 6. Hikone, Japan.

The following source [18], which is among the most relevant for this paper, cannot be so easily found in the library, either:

> Timo Honkela, Ville Pulkki and Teuvo Kohonen, *Contextual Relations of Words in Grimm Tales Analyzed by Self-Organizing Map* (1995). Proceedings of International Conference on Artificial Neural Networks, ICANN-95, F.Fogelman-Soulie and P.Gallinari (eds.), EC2 et Cie, Paris, 1995, pp.3-7.

The paper cannot be downloaded, for some reason, even from Timo Honkela's site, in spite of the link, but it can be downloaded from **CiteSeer** (http://citeseer.ist.psu.edu/), which is becoming a prototype of the scientific library of the not so far future. After a meandering search, I found a direct source: http://websom.hut.fi/websom/doc/grimmsom.ps.gz

It is always possible to miss something important by relying too much on the Web, but the main focal points of a knowledge domain still can be seen. Recoiling from the polluted with body parts, garbage, and dead links Web, one can either celebrate or mourn the end of the entire era when science had an endless capacity for pure paper memory and the typewriter restrained the rat-like fertility of documents boosted by the potent hormones of computers. The Web is a model of human mind which forgets most of the mundane information. Moreover, computerized word processing mercifully makes most of what is published as unreadable as the US Tax Code or US Budget just because of the volume.

I diverge here from the main course not for the purpose of lamentation. The problem is that the Web itself is a unique formation and **not a member of a population** of Webs. This example will serve later as an illustration of the significance of **uniqueness** for cognition—something an average chemist never deals with, but any mortal sooner or later comes to face. Another instructive property of the Web is how it **classifies** its pages: the category can be formed by a single non-trivial keyword common to all of them, as the social support group is bound by the single problem all its members have. This example illustrates the main idea of this paper.

The language elephant

I see the contest of titans in linguistics as a picture of esthetic value.

Noam Chomsky was a student of Zellig Harris (1909-1992) who followed the intellectual trend of structuralism, started by Ferdinand de Saussure (1857-1913), the founder of modern linguistics, well before the use of statistical tools and Bayesian inference. Since Harris is little known outside linguistics, a good introduction [19] to his ideas could be useful to evaluate his place among the prophets.

Structuralism believes that the structure, for example, grammar, can be derived from the object of study with no other tools but the atomistic analysis of the object and the transformations of the object that preserve structure (note the circularity of the concept of structure). Thus, various English texts, be it Henry James or Ernest Hemingway, preserve the same grammar but are strikingly different and the difference can be measured. We can talk about grammar or any other structure after we have identified, dissected, and described the invariance.

Comparing structuralism with PT, one can notice an important difference. PT is not a closed theory. Ulf Grenander noted that the choice of generators for describing a structure

is an intuitive process, which could be more or less successful. The same can be said about the choice of template. The principle of realism in PT [2B] breaks the so typical for formal systems cycle.

Harris was as much structuralist as Joseph Greenberg (1915-2001) [5B]. The structuralist credo is: "Here is the text and it will tell us everything, even if we do not understand the language, just squeeze (I am still short of "beat the hell...") the numbers out of it." Greenberg, however, applied the extracted knowledge to the forensic problem outside the texts, namely, to the past evolution of language, which cannot be observed at all. In other words, he focused not on what was preserved but on what made languages **different** in space and time.

Chomsky was disenchanted with pure structuralism and he and Harris took diverging roads. Harris needed to see the observable data. Like Greenberg, Chomsky wanted to see the invisible: the individual evolution of language inside human mind, but not through crunching the numbers. He focused on what makes languages identical. The Greenberg's way would mean that as many individual languages as possible should be examined for the evidence of the acquisition process. Many critics noted, however, that formal linguistics was too much preoccupied with English as the reference point for all languages.

The story of Harris, Greenberg, and Chomsky reminds me the old Indian parable, often evoked in cognitive sciences, of the elephant examined by six blind men. Depending on what part they touched, they described the animal as snake, fan, column, wall, spear, and rope.

This is, actually, a productive approach, used by Shimon Edelman [8] and other computer scientists, also with long roots and collaterals, converging on the Pandemonium of Oliver Selfridge [21], in which, metaphorically speaking, a group of narrow experts trained to recognize only particular parts of animals, successfully comment on the nature of an animal they never saw.

All that a chemist can honestly say, observing the internal currents in linguistics, is that everybody is right. The task of Chomsky, however, has been the hardest. I can vent

my irritation [4] with the Gothic world developed by his followers (Shimon Edelman did that too : "ADIOS, poverty of the stimulus" [22]), but the dark upperworld of human mind pulls like a magnet even through the eerie portal Chomsky opened.

Language (and the world) remains as invariant in its various projections as the elephant in human hands, but isn't the truth **distributed**? If so many projection of language already exist, could yet another one be found or even could there be a synthesis of them all?

In this e-paper I am giving the elephant a second touch. But…

A former child's credo

An outsider, I must each time present some credentials for the foreign territory of language acquisition by children. Well, here it is: I am a former child.

I hope that the following **credo** will serve the purpose. A *credo*, which literally means **I believe,** is still a kind of compass for navigating toward the destination.

First, I reject the religious belief of many linguists in the "discrete infinity," [punch it on Google], i.e. , unlimited combinatorial freedom of language. Combinatorial freedom does not exist even in chemistry. Any such freedom is chaos, while life and all its manifestations are an ordered and, therefore, limited chaos. Zellig Harris ingeniously called the constraint on language "non-equiprobability of combination of parts," which, if you trust me, is also the core of chemistry, as well as of PT.

Second, I believe, as a father, grandfather, and a former child, that the problem of language acquisition is different from the problem of unsupervised learning in computer science in the following overlapping aspects:

The child has no corpus to work on.
The child has no computer neither at hand nor in the mind.

The child does not keep statistical records.

The source of learning is a flow of sounds and events, not of printed text, until
appropriate time.

Any realistic corpus for modeling language acquisition can be only phonological.

Acquisition is diachronic and partly supervised.

The child has no idea about word delimiter, spelling, and categories of grammar.

The child's only wish is to maintain a comfortable homeostasis. The messages
between the child and the environment have meaning and emotionality.

All that is true only before a potent machine of supervised and self-supervised
(intentional) learning starts humming around the age of five years.

To summarize, the only complete experimental data for studying language acquisition can be a six year long 24 hours a day video. This is a difficult task, but to watch this video would be the closest we can come to watching the evolution of life on earth. Neither one nor the other is possible.

Third, I want to purge the algorithmic intelligence from the first language acquisition. One immediate consequence is that the natural uneducated mind cannot perform any systematic search in the form of an algorithmic loop (for $i = 1 : \mathbf{n}$ do … end). The concept of Universal Grammar tacitly assumes that a homunculus inside the child's mind performs some intelligent work by selecting the appropriate grammar that fits the cues. I understand distributed intelligence as local one, see next point.

Finally, I cling to the strong and somewhat controversial statement about the locality of all real processes in the mind and universe: molecules and human brain work by **short-distance interaction in a topological space**, regardless of the Euclidean length of the axon. This is the norm. The controversy lies in the necessity of non-local interactions for creativity. The exceptions from the norm are called fluctuations in physics, mutations in biology, and leaps of imagination in art and invention.

The term **local** is used in different and not always clear meanings. Thus, in AI **localistic** means that each concept is represented by one element of the system, which, actually, means **localization**., i.e., the opposite of **distribution** called **delocalization** by the chemists. From a certain angle, the history of our civilization is the history of shrinking metric distance. The radius of immediate neighborhood is constantly expanding in Euclidean metrics.

The idea of locality has an unusual aspect: evolution is also local and the criterion of any naturally evolved system is the ability to evolve from something extremely simple. Simplicity is not a circular notion: the simple can be counted on the fingers of one hand, actually, even on a chicken foot. A radical evolution (revolution) is also local, but at a higher conceptual level, also in a small phase space.

The New and the Different

The category of novelty, so vital for law and art, plays a particular role in chemistry [5]. Novelty is also the essence of evolution. There is an intuitive connection between novelty and locality, illustrated by the great geographical discoveries driven by the desire to overcome the locality of the known word. Discovery is a systemic phenomenon.

It is the principle of locality [15] that complements the idea of classical atomism and unites all natural sciences about combinatorial objects. I would formulate it as:

Any act of change happens, most likely, but not exclusively, in a single topological neighborhood.

This "not exclusively" is the margin of novelty. Language acquisition happens only once in individual life, as language emergence happens only once in history.

We can acquire only what is new. We explore and map the topology of the object space according to the observable events. **The topological neighborhood is where the changes happen,** however, and the vicious circle closes. As usual, we can break the cycle by stepping out of it and taking an outsider's position. Fortunately, observing natural sciences, we have an unnatural science to lean on. In terms of computer science, which is outside natural sciences in the classical sense (i.e., whatever actually happens in the computer is completely transparent from the printout), mind is not a computer as we know it. It does not prevent us from designing a computer which is like the mind we know. But if we design such artificial natural mind, we will never know completely what is going on inside. Remarkably, this is what the Turing Test is about. The Great Test is only tasting better with years.

> But what is **natural** science? The question unthinkable for classical philosophy remains: is it possible to model a **natural** object with an **artificial** one? Consult Robert Rosen [22] , who tried to design a virtuous logical cycle, but do not expect a revelation. Both Rosen and Mack turned to category theory as a caravel out of this circular world.

If not for this credo, I could not claim any role for chemistry in developing the most general picture of the world, where language is just one lonely figure standing apart, albeit high up on the hill.

The aforementioned examples lead us to the additional curious problem of scientific epistemology, regardless of Gödel theorem: many powerful statements are circular. This may not bother a mathematician, for whom it means that the terms are simply axiomatic, but a natural scientist could feel uncomfortable. As I have repeatedly suggested [5], the predicament may follow from the axiom of closure on which all mathematical systems—and the Aristotelian logic—are based. We can talk about **autopoiesis** [23], but it means just another circular definition or synonym of bootstrapping. What is *auto*, anyway? The clinch could be broken by a kind of diachronic mathematics which could formalize the notion of novelty. When a "new" shape is presented to ADIOS (Automatic Distillation of Structure) [8], or ART [10], it is just a **different** combination of the old elements from a closed set. Nevertheless, ART comes to capturing the novelty as close as it could be without laying hands on it: the new in ART is

what is absent from memory. I am inclined to regard as **new** whatever requires an addition to the generator space in order to be **understood**. The New cannot be understood, but it can be incorporated as understandable Old.

At this stage I cannot offer any logically consistent formal system and I am not quite sure whether it is possible or desirable to formulate it without a kind of axiom of closure that would kill it. I believe that the mind is an intrinsically open and fuzzy system and it is more appropriate to describe it than formalize. A formalized mind cannot discover anything outside the formal system. Knowledge acquisition moves on in an **expanding space**.

Life sciences owe their striking success to the low level of formalization in chemistry and biology. This heretic idea does not imply anything anti-mathematical. It is a vague doubt in the very Aristotelian foundation of "exact" sciences and an instinctive gravitation to something more Heraclitean, as far as the mind looking at itself is concerned. More about it, see [5A]   In short, the difference between the two complementary methods is that the A-science is beautiful, logical, and predictive, but blind to the distinction between the new and the different, while the H-science has a sharp vision of the novelty at the expense of prognostic power, symmetry, and logical perfection. As for beauty, it is, of course, in the eyes of the beholder.

Go I don't know where and bring me I don't know what

The title of this section, again about novelty, comes from a Russian folk tale [24]. It looks like a definition of novelty: we cannot request anything new in advance because it has no address and no name. It comes uninvited and unannounced and could be at first mistaken for something known.

I have a feeling that all modern cognitive approaches are just projections of the same paradigm to which there is no mathematical alternative. The world casts its Platonic shades onto our sensory fields which are a closed set of cells possessing a clear topology in the sense that some of them are in the neighborhoods of others. Our brain is a large set of cells also connected in a definite but, most probably, partially variable way. Finally, the output projects onto a combinatory space of muscle contractions. The brain is not a pushbutton device and normally cannot be turned off. It is usually in a state of spontaneous activity, most typically, when we think, which we can do with inactive sensors and effectors.

The question is what we can learn about the world. Obviously, it is not the constantly changing background, like the shapes of all leaves on a tree or the clouds in the skies. We learn what has some consistency and continuity, especially if we suspect it can change our live. **We learn only what is new** because we do not need to learn what we have already learned. We can afford learning a lot of what we do not need because we are blessed with forgetting.

The sensors provide us with an eternal soap opera, which a critically minded observer can only interpret as that nothing new can really happen in the mind because whatever happens is just a point in the phase space of the mind. Everything is just a different configuration of the same generators. This theoretical view is without any relevance because the phase space of the mind is so large that during the individual life time only a small volume of it can be accessed. The function of science (and ideology) is to map it onto a much smaller space of abstract models. The animals and plants have been so successful at higher taxonomic units because they did not have a complex mind. Watching history and unprecedented among animals interaction between individuals, we are often fearful whether human mind can go too far in its diabolical complexity. Animals are as vulnerable to bad luck as the humans to errors of judgment.

I am not certain that my suggestion is new, but the picture changes when we look at a **population** of minds. We can see that the minds are exchanging messages, which is what goes from the outputs of one mind to the sensors of another. The messages exist in a different space, where we do not find either cells and connections or inputs and outputs. The **conceptual** space is in a constant flux, but it seems to be expanding over time, even

though individual minds can lose their content and disappear. This collective space includes all of the culture, together with art, ideology, science and technology, and, since cultures are different, it may not be compact everywhere. Some subcultures (like infra-Red and ultra-Blue in modern America) do not communicate.

A significant part of the collective space is stored in reference dictionaries and, in general, is observable, so that we can easily establish novelty. The new is what requires a new generator to be included in the generator space, for example, in the form of a dictionary entry. It can be done in the act of communication. Only what can be formulated in language, image, sound, touch, and other sensations can be understood by the receiver. Until then we never know if the sender of the message understands it himself. To see this process in action, we need to compare two dictionaries separated by a long enough time interval. Obviously, in the times of Shakespeare, the following entry would be incomprehensible:

Electron: A subatomic particle in the lepton family having a rest mass of 9.1006 x 10^{-28} gram and a unit negative electric charge of approximately 1.602 x 10^{-19} coulomb.
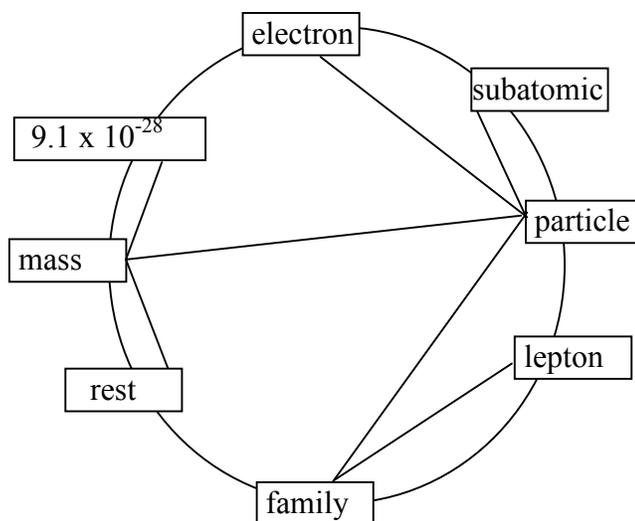

Figure 1. Semantic configuration for electron

This message is presented as a semantic configuration in **Figure 1**. Outside the textbooks, however, electron is just *electron* and is used as a single word. There was a rather short period after 1897, when it had been discovered by J.J. Thomson, during which electron lost its virgin novelty.

Suppose, a student finds the above definition of electron for the first time. How can it be understood and acquired if the new is what cannot be understood immediately?

To understand means to linearize the semantic configuration in whatever language, so that it could be communicated. This may look like a paradox because I attribute understanding to the teacher-sender, but not to the student-recipient who can just repeat what he heard. We understand when we send a message, not when we receive it. The recipient student must **prove his understanding by teaching** the teacher back.

A node in Kohonen net must be known and labeled in advance. But a different pattern in Grossberg's ART can be assigned a new node. Can ART discover electron? The actual history of discovery was much more intricate, but all we need to re-discover electron is to notice that the small electrical charges are multiples of the same minimal value. To make the next step, we need to apply the metaphor of counting coins, grains, and fingers, which are also multiples of the same minimal value, i.e., make a long and, strictly speaking, irrational leap of imagination. ART works on a raw input (retina, text), which does not explicitly carry any meaning, while the sophisticated mind operates in concept space. Honestly, I do not know the answer, but I suspect it can in principle, after some modifications and adjustments to the semantic space. But the same can be said about any contemporary cognitive approach: from whatever starting point we depart, if open to the entire body of expanding knowledge, we will sooner or later arrive to a valid and mind-like cognitive model.

If we all spoke the same language, we would not have linguistics. Languages arise from the one-to many mapping of the semantics (which itself can be muddled) onto linear speech, as **Figure 2** illustrates. It does not matter for semantics what the absolute positions of the nodes are because **semantics is out of this physical world**. What matters are their connections. But as soon as we try to linearize the meaning, a great variety of alternatives appear. The grammar, as I believe, works as a typical catalyst, minimizing

the thermodynamic barrier to decompressing and restoring the meaning from the utterance by the recipient.

Having started with the concept of **novelty**, I moved to **communicating information**, which has a non-zero content only if contains something **new**, and further to the necessity to **linearize** the content in order to **push it through** the single channel and to be **understood**.
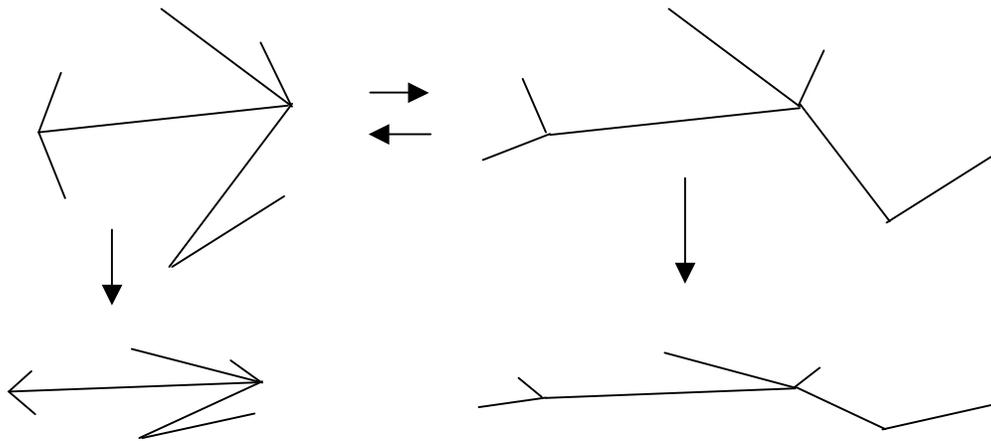

Figure 2. Linearization. The two topologically identical but different in the plane representations at the top generate different linear projections at the bottom

Speech is not the only possible way of communication. Pictograms and ideograms can perform the same function, but linearization problem will come up sooner or later when a new tidbit arrives: "The Normans are coming!", then "The Russians are coming!" then "The Mexicans are coming!" "The Chinese are coming!" etc. Finally, the Aliens are coming and we do not have a picture for it.

Connections and collisions

Today connectionism is synonymous with Neural Networks, but I look at it askance from my chemical perch. Most Neural Networks are processing devices. They have input, output, and direction of information flow, which could be backward as well as forward.

I am sentimentally attracted to the oldest chemical apparatus: reaction flask, initially a cooking pot. When the "input" reagents have been mixed, the reaction starts. Only in some particular cases there is an evolution of gas. The content of the flask is its output and it changes with time, coming to an equilibrium. Instead of giving the reaction some time to run, a continuous throughput apparatus can be used, but it just freezes the picture of the chemical process in time, without changing the results. The length of the tube corresponds to the reaction time.

Human mind works in both processing and spontaneous modes. The latter is what we call by different names, from philosophical reflection to creative art and problem solving. It has almost no external manifestation and the output may be delayed or absent. Spontaneous thinking has a stimulus, comparable with mixing the reagents, often subconscious, but no program. From what I know about computers, spontaneous thinking of a typical computer is always algorithmic. The machine just runs the program and spits out the result, the sooner the better. Neural Networks run the same way, only without a global program, instead of which there are local ones implemented in the design. Usually they are simulated by symbolic computation.

> I strongly believe that the wider our view of the subject (a.k.a. the language elephant) is, the deeper we can penetrate into it. This is why I allow myself the following digression, which is a trailer for a more extended treatment. Does the **computer** have consciousness? It **does not need** consciousness because its states are strictly linearized in time. Do **Neural Networks** have consciousness [25]? They **cannot have it** because they are not linearized by definition. What about **human mind**? It is a neural network, therefore, it **needs consciousness** for really complex tasks that NNs cannot even "think about." The mind needs consciousness to compete with computer. What about animal mind? Any

animal that displays **time-consuming** scenarios in its behavior, like predators do, and not just simple response to a stimulus, **needs consciousness**. Consciousness is needed for performing long sequences of actions because one state must be distinguished from another and there is no homunculus to do that. Automatic skills, however, can be quite unconscious. As a professor, I needed to think hard in order to find a way to convey to students complex ideas for the first time, but after many repetitions, I could be giving a routine lecture and catching myself on thinking about something else. What a Labyrinth waiting for an Ariadne's thread!

"It is suggested that the type of processing related to consciousness involves higher-order thoughts ("thoughts about thoughts"), and evolved to allow plans, formulated in a language, with many steps, to be corrected."(Edmund T. Rolls, [25B]).

Figure 3 illustrates the abstract parallel between chemistry and connectionism, as I see it, which, most probably, is not as NN sees it because there are no inputs, outputs, function, and nodes in it .
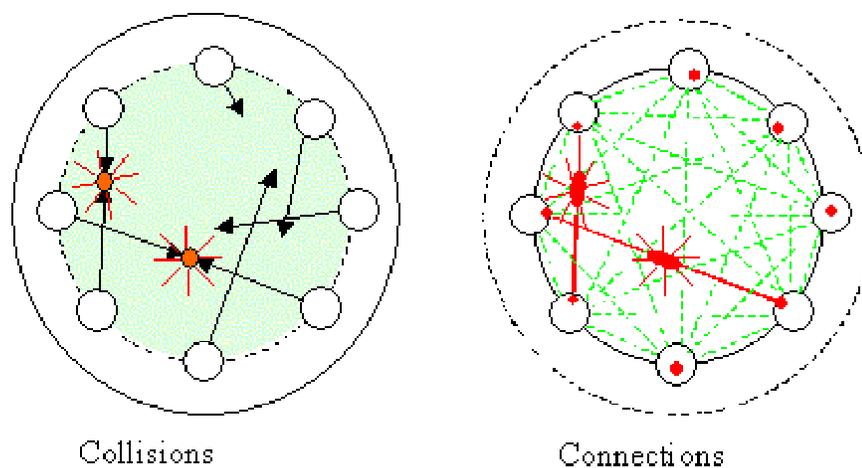

Collisions   Connections

Figure 3. Collisions and connections

In the chemical system, the molecules (small circles) are dashing around within a fixed volume, generating a certain average frequency of collisions per unit of time. In its cognitive analogue, working as a telephone switchboard, the fixed "molecules" connect at random, experiencing a certain average frequency of connections. In both models, the

event space is colored light green, which emphasizes the main (but by no means radical) difference: the space is continuous for collisions and discrete for connections.

The immediate connectionist question will be about the structure in the connections: where is it? The awfully heretic answer is that all small circles in the "connections" part of Figure 3 are fully connected, which effectively eliminates structure from consideration because all connectionist systems are identical. This is what we can say about any typical chemical system: all collisions are possible. Chemistry may look non-Harrisean to a linguist, but this is not so because collisions result in chemical bonds leading to structures with very different **probabilities**. A physical system such as ideal gas or liquid could be justly called non-Harrisean from the linguist's perch. What makes chemical collisions and cognitive connections structured and Harrisean is the non-equiprobability of elementary events.

> The non-equiprobability of atomic events is the core idea of Pattern Theory. PT should be used instead of my awkward language (langue-in-cheek?) intended only for teasing. Zellig Harris and Joseph Greenberg were the founders of linguochemistry. But what about Noam Chomsky? Far from linguochemistry, he matched the role of the physicist George Gamov who predicted the principles of genetic code in a very general but vague in details form. I believe that the final synthesis of linguistic theory is possible. But not peace: scientists today fight not so much for ideas as for grants.

How can we expect any function from a system without structure? A chemist would start telling about multiphase systems, like oil and water, in which some collisions are hindered because molecules have different affinities to different solvents. Then the chemist would give examples of creating an actual structure in a formally homogeneous solution, as we see in gels, such as the fruit desserts made with gelatin, and finally point at living cells, in which we have sophisticated structures made of nothing but water with some polymers, lipids, and sugars. Cellular membranes, organelles, and some cells that insulate neurons are examples.

With all my proclivity for digressions I cannot engage here into chemical details because we are interested not in chemistry but in meta-chemistry. What is important, practically all living cells are **small**, at least in two dimensions, and for a simple reason:

the small size ensures that in spite all constraints, all molecules in a living cell can collide within a short time. We will return to the importance of being small later.

The analogy with the flask can be misleading. In static fluid, molecules spread by diffusion at a limited rate. Therefore, even a gas in the flask does not guarantee a complete instant mixing of components. The mean free path (mean distance between collisions) for molecules of air has a magnitude of only 2×10^{-7} m. But the collision frequency is of the magnitude 2×10^{9} /sec. Two volumes of different gases brought into contact rather quickly "compute" their state of equilibrium. Nevertheless, a small molecule, like ammonia, travels across the room in several minutes. In liquids, however, diffusion is 10,000 times slower. The approximate equilibration of freshly combined but not mixed liquid components by diffusion can take months. This is why chemists do not rely on diffusion and stir their pots and pans as fast as possible. The typical cell has dimensions 2×10^{-6} m, which is only 1000 times more than the size of molecules (2×10^{-9}), but million times less than the size of the flask, so that diffusion alone may work reasonably fast.

The textbook explanation of the small size of the cell by its high surface-to-volume ratio ignores chemistry but not diffusion. The cells use various tricks to intensify molecular transport, some of them, like cytoplasmic streaming and axonal transport, remind of mechanical stirring. Besides, they do not need to reach equilibrium.

The cell is pretty close to a very small flask packed with even smaller test tubes, all with selectively permeable walls. The cell elongated in one direction, as neurons are, is a capillary. It takes a considerable time (could be days) for a molecule to travel along the macroscopic dimension by means of slow diffusion or even an enhanced transport. This is why neurons communicate not by diffusion, which may be good enough for protein synthesis and microscopic synapses, but by electricity. The human minds communicate ideas neither by chemistry nor by electricity, which the air does not conduct, but through sound waves, in modern times enhanced by electricity.

The connectionist modules must be small in order to move from configuration to configuration within reasonable time by **random** transformations, which would be

unreasonable in large realistic systems. The small system is the true cradle of life and cognition.

Most of the perceived world is metric and animals navigate it pretty well, acquiring knowledge but not sharing it. A new problem arises when non-metric knowledge, such as categorization, must be communicated, which is *sine qua non* of understanding. This function is performed by language. One may believe that language originated for the purpose of communicating thoughts, but others can see language as the substrate for the evolution of thinking. This is a separate and unrelated problem, as exciting as the chicken-egg problem, if you are in the mood and hate discarding the word "purpose," which is an empty egg shell, but still can be nicely painted.

Both connectionist (in my sense) and molecular systems display as a sequence of binary events, like matches in a football championship, but statistically distributed in time. In chemistry the temperature is the measure of the frequency of collisions. Similarly, the temperature of the mind can be regarded as the frequency and amplitude of mental events ("…he was feverishly thinking…" "he tried to bridle his racing thoughts…"). Statistical mechanics, however, too heavily influences and restrains the system theory. If we admit metaphor as legitimate tool of science, it will give us not just more freedom of imagination but also an opportunity to study the intimate physiology of metaphor *in vivo* in the Open Mind Theater. Metaphor is the best proof of understanding: you need a minimum of words to share it and will not be lost.

The heritage and the current span of connectionism are vast. For a chemist, the similarity of connectionism and chemistry is as striking as their differences are obvious. Drawing the parallel between the chemical flask and the mind, I must keep a safe distance from both. But the word "enzyme," used as metaphor, makes me hold on to both.

NOTE: After this e-paper was completed, I found some important sources for the "collision" computing, for example:

Banâtre, Jean-Pierre and Le Métayer, Daniel. 1996. Gamma and the chemical reaction model: ten years after. In *Coordination programming: mechanisms, models and semantics, IC Press*,. World Scientific Publishing. ftp://ftp.irisa.fr/local/lande/dlm-gamma10.ps.Z

Berry, Gérard and Boudol, Gérard, The Chemical Abstract Machine, *Theoretical Computer Science,* Vol. 96 (1992), 217-248. http://www.esterel-technologies.com/files/cham.zip Also: Boudol, Gérard. *Some Chemical Abstract Machines.* ftp://ftp-sop.inria.fr/mimosa/personnel/gbo/rex.ps

Adamatzky, Andrew. 2001. *Computing in nonlinear media and automata collectives*, Bristol, UK: IOP Publishing Ltd. Chapter 1:
 http://bookmarkphysics.iop.org/fullbooks/075030751x/adamatzkych01.pdf

Also: Adamatzky, Andrew, Editor. 2001. ***Collision-based computing*** . London: Springer-Verlag .
        See also: Yuri Tarnopolsky. Molecular computation: a chemist's view.
http://users.ids.net/~yuri/PTuter.pdf
.

Is the mind an enzyme?

The term *association* is as chemical as the *adjective* and *noun* of "**apple pie**" are linguistic. If one thinks about three little pigs (A), the wolf (B) promptly comes to mind because the entire story (C) is remembered.   A chemist could say that the remembered story is in **equilibrium** with all its components, which is pretty close to the chemical idea of equilibrium, as the following example illustrates.

The colorless gas N_2O_4 can be stored at low temperature in the freezer as a liquid. At room temperature the gas acquires brownish color because of the equilibrium:
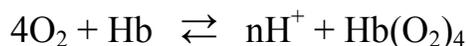
$$\bigcirc \quad N_2O_4 \rightleftarrows 2NO_2 \quad \bullet$$
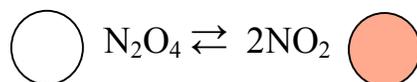
dinitrogen tetroxide $\rightleftarrows$ nitrogen dioxide

temperature $\xrightarrow{\hspace{2cm}}$

$\xleftarrow{\hspace{2cm}}$ pressure
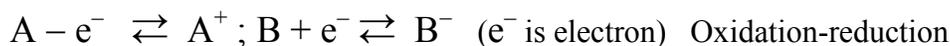
Whether we start from pure N_2O_4 or pure NO_2, both forms will be present and the mixture will move toward **equilibrium**. Its **position**, i.e., ratio of both forms, depends on the temperature and can be visually evaluated by the color. It also strongly depends on the pressure because when N_2O_4 turns into $2NO_2$, the volume of gas doubles. High presser shifts the equilibrium to the left, toward N_2O_4.

The direct association is not typical for the high school chemistry. Everybody knows that hemoglobin (Hb) absorbs oxygen, but in fact it is not association but a more typical oxidation-reduction:

$$4O_2 + Hb \rightleftarrows nH^+ + Hb(O_2)_4$$

The main patterns of chemical reactions are

A—B—C $\rightarrow$ **A—C + D**   Elimination

A—B + C—D $\rightarrow$ **A—C + B—C**   Exchange or substitution

A—B + C $\rightarrow$ **A—D—B**   Addition (inverse of elimination).

$A - e^- \rightleftarrows A^+$; $B + e^- \rightleftarrows B^-$ (e^- is electron)   Oxidation-reduction

Enzyme-substrate interaction belongs to the direct **association – dissociation** type of transformations, which are the bedrock of life chemistry: $A + B \rightleftarrows A—B$. They are not exactly the stuff of **common** chemistry, if such thing exists (it does not: chemistry is a submarine like any science, but with a captain). No surprise, **association** and alienation or **separation** are also terms of psychology and sociology. Assembly and aggregation are very general phenomena.

Let us consider for simplicity only direct association as the metaphor for universal interaction of generators. It is usually weak and does not result in a "normal" stable chemical bond. To form a strong ("covalent") bond $A—B$, chemists and living cell usually exploit exchange reactions of the type $A—x + B—y \rightleftarrows A—B + x—y$, where $x—y$ is a small molecule. In biochemistry it is water.

The term **recognition** is universally applied by chemists to the enzyme-substrate interaction. The "cognition" to precede re-cognition might have happened in distant evolutionary past or just on the fly. The interacting molecules must have a unique key-lock shape relation and be oriented in a unique way, as **Figure 4** illustrates. The bond between them cannot be pinpointed. It is **distributed**, notably, over a large number of weak atomic interactions which do not constitute the "normal" and "typical" covalent bond.

The enzyme looks like a chemical demon who not only recognizes its substrate but also performs an operation on it. It is very tempting to regard the enzyme as the simplest prototype of the mind, (a proto-homunculus, if you shun demons) from which the entire complexity of human life evolves, but this metaphor would be rather stretched. Let us keep it in mind, anyway.
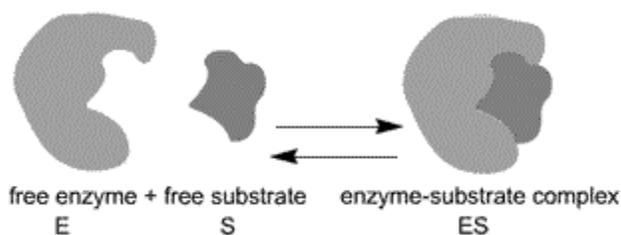
free enzyme + free substrate   enzyme-substrate complex
E   S   ES

Figure 4 . Enzyme-substrate interaction.

While the words *chemistry* and *catalysis* are components of commonplace metaphors ("catalysis for change," "the chemistry between them"), one aspect of enzymatic and, more generally, catalytic activity is all too often ignored outside chemistry. The chemical enzyme does not perform any work or operation (compare with Barrett [16A]). All it does is to increase the speed of both direct and reverse transformation, so that the system reaches equilibrium **much faster** than on its own. The predominant direction, whether forward or backward, is determined by the thermodynamics of the process, i.e., the position of chemical equilibrium, to which the system moves from whatever extreme, left or right.

Nevertheless, it may be stimulating to linger for a while at the mind as catalyst. Joseph Greenberg and colleagues started a computation of the Arabic verb roots at some moment and finished it at another [5B, 26]. With a computer, they would finish it much sooner. The hundreds of thousands of computation cycles on a parallel computer, needed to complete the self-organized map of the Grimms' Tales [18], would be impossible to do by hand. It is a hard nut even for symbolic computers. The computer, therefore, works as a catalyst, radically speeding up the arrival at the final stop point after which nothing can happen on its own. Similarly, a single wolf could not possibly kill a buffalo quick enough before having died of starvation, so that the minds of the wolves in the pack are working as catalysts speeding up the outcome. Further down, molecules of a freshly mixed reagents come to the final state of equilibrium either fast or slowly, depending on the presence of a catalyst. which itself can be used, like computers and the minds of scientists and wolves, many times over. This is the property we usually associate with machines. But the catalyst neither consumes energy nor performs work.

Is the human mind a catalyst? Obviously, only muscles perform work in mechanical sense and kidneys in thermodynamic one. If so, where does all the physical energy consumed by computers and the brain go? The quick (but incomplete) answer is that it goes for preventing both from coming to equilibrium and staying there indefinitely without any observable change **too soon**. Same applies to life. The function of life and the human mind is the race against time, always lost by an individual but sometimes won by the species.

The purpose of this digression was to emphasize the role of time and speed, always crucial for chemists, who call it **kinetics**, already somewhat relevant for computer designers and users, but not yet essential for cognitive scientists. The idea of irregular, unlawful, and even unobservable but speed-defining transition state seems to remain alien to cognitive sciences, but I might have missed signs to the contrary, except [16C, 16D].

On the mechanism of catalysis and transition state , see [5], as well as chemical textbooks and rich sources on the Web. It will quickly become clear to a reader what I meant by the unique chemical experience with time, which is called chemical kinetics.

As far as the unique insights of Gerhard Mack [15] and Clark Barrett [16A] are concerned, their significance is not diminished by deviations from chemistry. Following

the direction they indicated, somebody will sooner or later arrive, after me, at the simple but cardinal meta-chemical idea that **what happens** in the individual or evolutionary mind **is what can happen faster**.

No wonder we can think fast, finding an intuitive solution while computers keep crocheting their loops. If we want to think even faster, we boot up the computers. This is why our thinking and its verbal expression, as well as the chemical mechanisms of life, all seem so robust at normal conditions: they are faster than the errors and mutations crawling in. Be it otherwise …just read Hamlet. The trade-off is that our (**theirs**!) errors of judgment can be colossal.

Equilibrium and emergence of mind

The term "cognitive equilibrium" was introduced and used by Jean Piaget [27] in the same sense I use the term homeostasis and the evolutionary biologists use the term punctuated equilibrium [28].

Life and cognition are open systems which do not come to equilibrium. For them equilibrium is a misnomer. Piaget, however had no reason to be concerned with subtleties of thermodynamics. What cognitive equilibrium and homeostasis have in common is a sequence of stable states of low energy dissipation, **alternating** with states of high energy dissipation (stress, conflict, adaptation) during which the system searches for a new state of equilibrium. I would prefer homeostasis as the most general and sufficiently compact term for a large and fundamental class of phenomena.

An approximate mathematical image of homeostasis is the movement of the representative point through a landscape consisting of valleys, hills, and mountain ridges. In the real life homeostasis, for example, in the Lewis and Clark expedition, the landscape is unknown and must be either discovered or created. The homeostasis can be formulated as survival on the move, which is what life is generally about. Homeostasis requires a sufficiently high abstract temperature, i.e., degree of chaos, which would

knock out the representative point from any sleepy hollow and the polar explorer from his warm sleeping bag.

 Figure 5A shows how the landscape may look to a mathematician, but in evolving **open** systems it looks like **Figure 5B** in which only the close neighborhood is more or less visible. Moreover, even with hindsight, only the trajectories that already have been passed can be seen, while a lush valley or deadly abyss just two steps aside remains shrouded in fog. Evolution is not about closed mathematical structures.
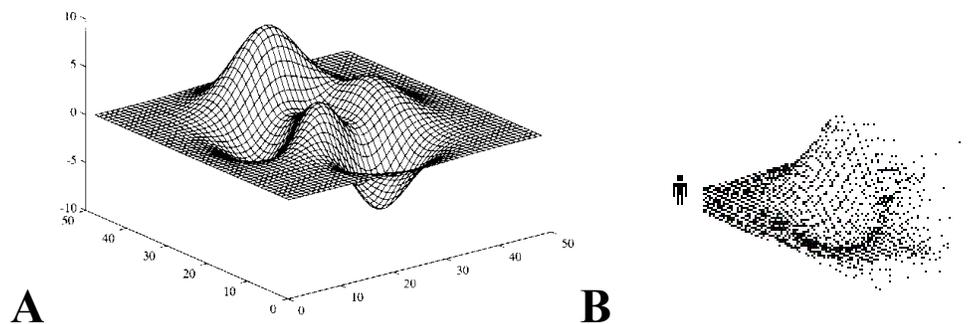

A   B

Figure 5. Evolutionary landscape: A. Closed system, B. Open system

 In the case of language acquisition, however, the landscape can be studied because the individual evolution of a child has been repeating billions of times.

> This process is epigenetic in some limited sense. There is very little molecular genetics in it but a lot of social genetics. I see the individual language acquisition as a homeostasis on a known landscape, but evolution of species, whether biological or social, is the creation of the landscape.

 In the mind flask, regardless of the physiological nature of the process, to which I have no clue, all we can metaphorically say is that there is a kind of "equilibrium" between the following components of unknown and irrelevant for us material nature:

The Three Little Pigs   3LP

Pigs   P

Individual pigs   **P1, P2, P3**

Wolf   W

$3LP \rightleftarrows P + W$

$P \rightleftarrows \textbf{P1} + \textbf{P2} + \textbf{P3}$

$3LP \rightleftarrows \textbf{P1} + \textbf{P2} + \textbf{P3} + W$

The above equations may look like an abomination to a chemist. They do not explain how a single W or $P1$ can retrieve the entire tale. This would violate the sacred for a natural scientist conservation of atomic particles. We are coming here to the most fundamental, although still hypothetical, difference between chemical and cognitive systems: **cognition, unlike chemistry, does not know multisets**. And yet there is a possible chemical interpretation.

Figure 6 shows the already familiar "chemical" equilibrium, in which there is no way $P1$ can **generate** P or $3LP$. Nevertheless, there is a chemical analog for the function of retrieval, but it is even less known outside chemistry than catalysis. Not surprisingly, it is called **extraction**, which is synonymous with **retrieval** and **recovery** and used for those purposes.
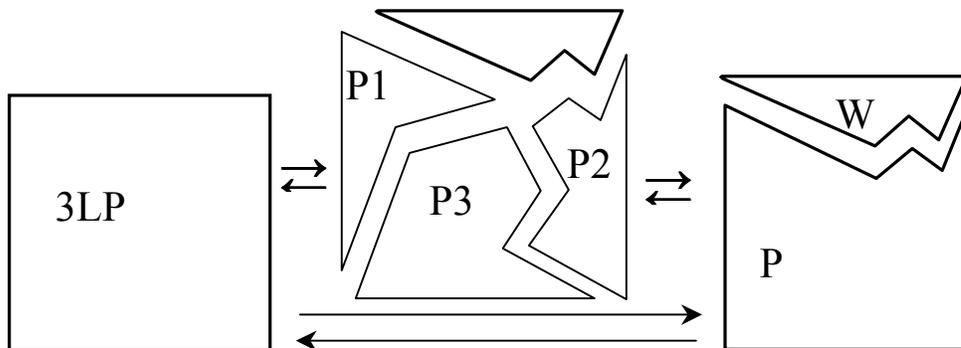

Figure 6. Equilibrium of the whole (3LP) and the components (P,P1,P2, P3, W)

The classical *homogenous* chemical equilibrium is only one kind of equilibrium, in which all components are in the same phase. In a heterogeneous medium, for example, immiscible oil and water, or in water solutions separated by a membrane, as it is the case in living cells, some the components do not have a complete freedom of movement between the phases and come to an equilibrium distribution between the two phases, like water and oil, which can be highly skewed.

In **Figure 7** the shaded lower part and the clear upper part, separated by an interface, represent the subconscious and conscious phases of the mind, with the equilibrium shifted toward the subconsciousness.

The lower part shows how component $\mathbf{W}$, which is in the focus of consciousness, recognizes and retrieves the rest of $\mathbf{3LP}$ from memory but leaves $\mathbf{Q}$ there. This is the case of a well-known and widely used in chemistry phase equilibrium: distribution of a component between two phases or **extraction** from one phase into another, for example, from water into oil or back.

Component W retrieves P into consciousness because it **recognizes** it in the same way enzyme recognizes its substrate. It is not the case with Q. This is the chemical metaphor, but what does stand for internal recognition in the real mind? Obviously, it is the connection, i.e., the bond couple of PT, for which the connection in a Hopfield type NN is the closest approximation (not an exemplary one because there is no normalization in Hopfield nets).
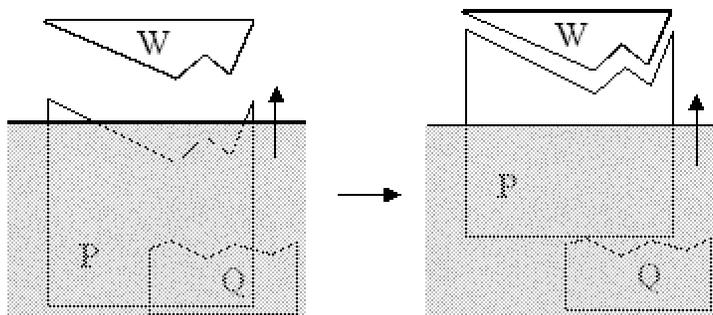

Figure 7. Extraction of components from memory into consciousness

We come here to the very starting point of the transition from chemical life to mind: the chemical components develop a long-term contact **over a macroscopic distance**, unlike collisions and enzyme-substrate interaction. This evolutionary moment is known in biology as emergence of nervous system. It is completely understandable for a chemist, but chemistry does not work through long capillaries (rather, it works, but very slowly) such as dendrites and axons. Here the Lewis and Clark of biological evolution have to change the horses for electrically powered vehicles. This is not yet truly electrical communication because the nervous impulse is still awfully slow as compared with the speed of light.

The macroscopic distance in space leads to macroscopic distance in time. For example, when ribosome starts working on a strand of mRNA, the macroscopic length of the template keeps the process going over a large number of consecutive steps.
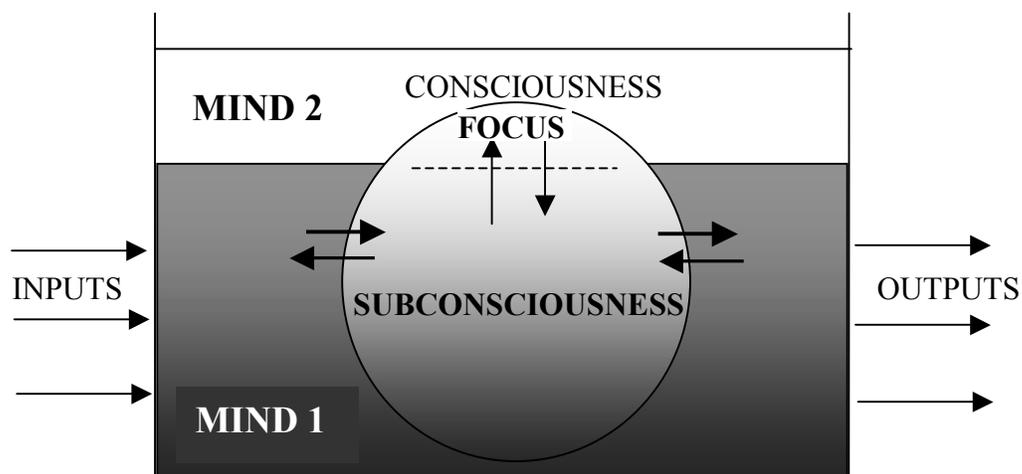

Figure 8. Mind as a two-phase system

The entire motion picture of the mind displays between consciousness (MIND 2) and subconsciousness (MIND 1), as **Figure 8** illustrates. The heterogeneous equilibrium and heterogeneous catalysis, which are two of chemical foundations of life, apart from its basic non-equilibrium thermodynamics, smoothly turn into the foundations of the mind, as soon as nervous cells turn into long capillaries. The practical visualization for this evolution of distance—the main content of our material civilization—is to take a piece of

chewing gum from your mouth (or use any other putty, slime, or goo) and start pulling it apart, dividing the piece into two parts with a thin fiber between them. Chemistry does not run well through capillaries, but electrochemistry is fine.

Christopher Davia's idea that consciousness is a soliton wave [16C] seems quite passable even as more than a metaphor. It is a wave of excitation through time.

The nature of chemical equilibrium is entirely physical and has nothing to do with the particular structure of molecules. The position of the chemical equilibrium can be calculated from the energies of participating molecules. If all energies are equal, the concentrations of the molecules will be equal, too. If not, the molecule, aggregate, or, better to say, **configuration with the lowest energy** will dominate, as in Neural Networks. It is necessary to know only the total number of atoms and relative energies of molecules, but not their absolute values. For a cognitive system, where every configuration is present in a single copy, the concentrations turn into probabilities. They are automatically normalized by the conservation of matter. In neural networks, the concentrations turn into weights.

More exactly, both chemical equilibrium and recognition by Neural Networks are the instances of optimization, namely, minimization. In terms of the landscape, it corresponds to a descent from a mountain or a slope to a valley. In computer science, unlike chemistry, the very fact of convergence of the computation used to be of a much lesser importance than how fast the minimization can be achieved. The very attractive principle "the winner takes all" in NN is a pipe dream for a chemist, who has to deal with mixtures of reaction products that populate all the nearby valleys, unless enzymes are involved.

The term *resonance* in cognition is excellent and the entire paradigm of adaptive resonance theory, regardless of its particular implementations, looks highly realistic for a chemist allergic to mechanical rigidity.

Chemical systems, however, although not perfectly predictable, are perfectly reproducible. This is possible because of the large frequency of collisions and large number of particles. Theoretically any particle can collide with any other, which in

cognitive systems translates into full connectivity, as for example, in self-organized maps. The full connectivity is typical for a telephone switchboard and email systems, but at a price: only a small part of all connections can be active at the same time and conference calls are exceptions.

The next question is how can the full connectivity of elements be achieved?

Small is big

The answer to the last question of the previous chapter is that only small configurations can realistically have a full connectivity. For large generator spaces the combinatorial explosion makes this impossible.

In the flask, an aggregate (called *complex*) always dissociates, however slightly, into components. If not, it is an atom and not a complex. Both the aggregate and the components are present **simultaneously** because they have **multiple** copies. The number of identical molecules per total number of atoms is called concentration. If the number of atoms in the system is constant, the concentrations are always normalized. This is the chemical reality. At the zero point of individual life, however, fertilization is an interaction involving **single** molecules which then begin to multiply.

We do not know what the cognitive reality is, but in any case we can represent it by any normalized parameter, whether we call it weight, probability, or frequency of firing.

In the pursuit of realism we cannot avoid the question of what **physically** makes the mind a conservative system of a kind. The answer is certainly not the number of atoms or neurons. It most probably comes from the biochemical nature of the brain. It is an open system that can work on a supply of energy coming through the arteries with blood flow. It is more realistic to assign certain margins to the conservation, depending on the intensity of neuronal activity, but there is always an upper limit. In short, all the neurons of the brain cannot work at once, which is exactly what we see in experiments with brain imaging. It is very natural to assume that this principle works also in any

separate segment of the brain because of the branching blood vessels. Any brain activity, therefore, is a case of competition for a limited resource for a simple physiological reason: you cannot have it all. The "winner takes all" is just a metaphor as far as the supply of energy is concerned. The winner can silence the loser.

Next, the chemical transformation happens stepwise, with a very limited number of atoms involved at any step, always at a close distance in space, although the overall distance along the chain of atoms can be large. A good popular example is the ribosomal synthesis of proteins on mRNA in which all the major events happen within the limited active area moving along the mRNA chain and spinning off the protein chain.

> This picture ostensibly contrasts with the mind, where we can jump from cabbage to kings and back in a moment, as if our mind was a computer, which it is not. In order to jump we need to know where to, but if the map is as big as our knowledge about the world, we can explore it only by crawling along some connecting graph or jumping at random within a close horizon, unless you are Columbus, Magellan, or a poet.

Let us use the terms **attention** and **consciousness** interchangeably, until we know what they actually mean, or call it **processing window,** or use **focus** (similar to **content** of PT) instead. As soon as we accept the premise that the focus is always a relatively small set, the difference between chemistry and cognition disappears, at least to a poet. But why is it so necessary for the active areas of molecule and the mind to be small? The effect of this principle is that in a small generator space set all possible connections can be searched (scanned, fired, repeated, etc.) within a **realistically** short time, as it is the case in the vigorously stirred chemical flask. The living cells are small in order to run the chemical reactions fast without stirring. The entirely mathematical idea of neighborhood "where everybody knows your name" attains a physical embodiment.

Both chemical synthesis and successful behavior are in fact possible because the chemical reaction and the act of cognition are both transformations of small configurations. They do not include the exhausting search in multidimensional spaces typical for symbolic computing, all the more, such abstract spaces are enormous in cognitive systems. Moreover, the small size of the focus makes unnecessary the massive computing of SOMs, which would require enormous energy consumption in the brain.

Here the little demons of Shimon Edelman come to help. Moreover, the small systems may be really the seeds for bootstrapping or, paying respect to the Greeks, autopoiesis [23].

From thought to language

Let us fantasize a bit about the smallest possible cognitive systems and their relation to language.

Figure 9 shows a quasi-chemical equilibrium with monovalent generators.

As everywhere in illustrations of this kind, small circles placed on a big circle symbolize a system of atomic objects in each other's neighborhoods. In the same sense all molecules in a volume of a fluid can potentially collide.

Here the "**no duplicates**" constraint causes no formal problem with translating configurations into molecules. This is a trivial case, however.
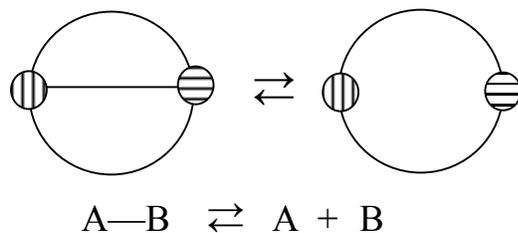
$$A\text{—}B \; \rightleftarrows \; A + B$$

Figure 9. Bonding in a cognitive or chemical system

The most important case of minimalist chemistry involves three **monovalent** generators.
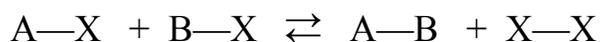
$$A\text{—}X \; + \; B\text{—}X \; \rightleftarrows \; A\text{—}B \; + \; X\text{—}X$$

This equation, so natural in chemistry, is impossible in cognition. The approximation $\quad A\text{—}X \; + \; B\text{—}X \; \rightleftarrows \; A\text{—}B \; + \; X \quad$ is no better because it still

contains two copies of X. We just cannot do it with the monovalent X. Having made the necessary fix, we come to the formally satisfactory:

$$A\text{—}X\text{—}B \;\rightleftarrows\; A\text{—}B \;+\; X \text{, or}$$

$$A + B + X \rightleftarrows A\text{—}X\text{—}B \rightleftarrows A\text{—}B \;+\; X$$
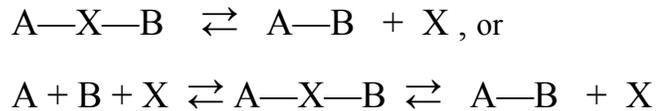
The last equation is the exact chemical symbolism for catalysis with bivalent catalyst X. This is what any enzyme or **heterogeneous** catalyst does. A and B must be attached to the **same** X and in close vicinity of each other. The catalytic effect is impossible with two separate molecules of the same enzyme involved. The human (though abused) metaphor for enzyme is the pair of human hands.

To explore chemical parallels even further, we need to look at the nature of the chemical and cognitive bonds. The bond space of the atom is defined by its quantum-mechanical properties. The question is whether the cognitive generator has something like the valences (*arity* in PT) of the chemical generator leads us into the most complicated, i.e., full of circular definitions, problem of the relation between language and thought and syntax and semantics, as well as relations within each, and I would rather skip it. But it is impossible to evade the distinction between language and semantics. Both are seen by the linguists as a kind of landscape stretching behind the horizon and best seen at some optimal distance, neither too close nor too far away.

I believe it is in the conformance with the idea of Ulf Grenander to suggest that in the small test tube of the mind where the current events happen there is a relatively small configuration (content) which is in a statistical equilibrium with all its subconfigurations by the very definition of what configuration is in PT.

I believe that the simplest non-trivial configuration consists of three generators.

This is intuitively very close to what the arrow in category theory means, in the sense that the process of evolution from simple to complex goes through a kind of triangulation, in which one generator, working as catalyst, enhances a bond between two other generators connected to it. Mathematics, however, is a representation to which time, physics,

chemistry, and physiology do not apply, while it is applicable to all four. The brain, on the contrary, is constrained by the quartet. Not accidentally, the linguists use the basic word order of subject, verb, and object for the primary classification of languages.
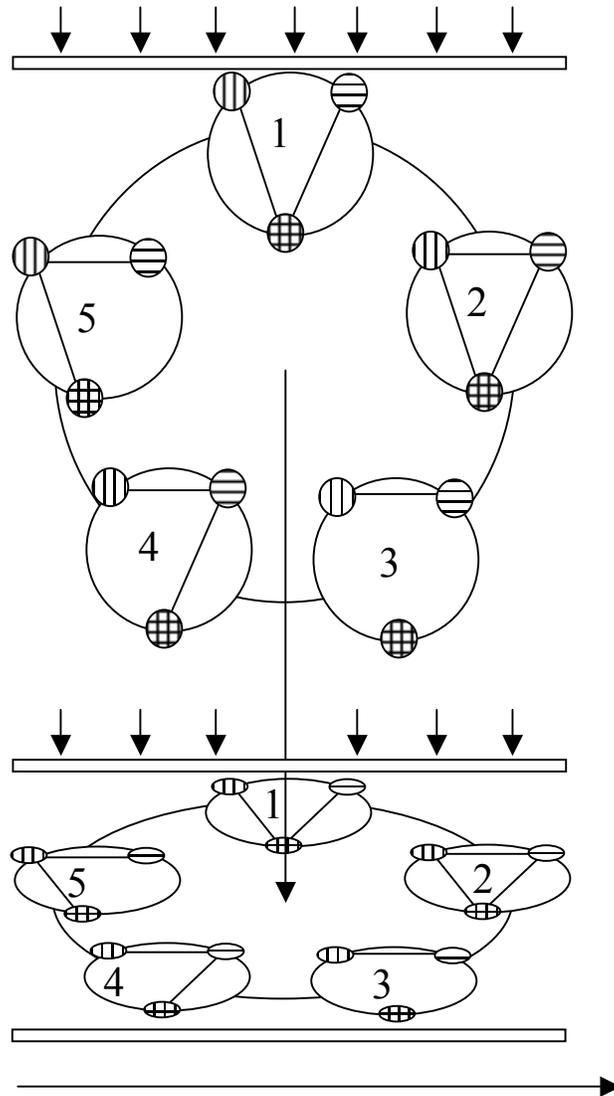
Figure 10. Linearization of the three-generator configurations

The terminal generators are signs for objects perceivable by sensors. The typical catalyst is the generator of a pattern, i.e., the non-terminal generator that stands for a class or category. For example, **cat** pulls **pet** into content and **pet**, in turn, pulls **dog** into the company. The members of the class fraternity recognize each other by the man-made badge of the class.

Coming back to the idea expressed earlier [4,5B], **Figure 10** shows how various graphs on three nodes behave during linearization. While all doublets can be projected onto a string without effort, only two triplets, No.1 and its flipped copy allow this without breaking one bond.

The lower part of **Figure 10** is the transition state stressed by compression and the uncertainty of it linearization outcome.

A natural hypothesis is that the innate mechanism of language acquisition works up to the stage of three-word utterances, which is (and was in evolution) the border area between Nean and language. After that the child learns the language as it learns to play games and do the shoelaces. The new stage, probably, starts with non-terminal words and morphemes that denote the invisible and imperceptible classes and categories. This is not up to a chemist, however, even if he is a former child, to speculate any further.

At this stage I cannot offer any logically consistent formal system and I am not quite sure it is possible without a kind of axiom of closure that would kill it. I believe that the mind is an intrinsically open and fuzzy system and it is more appropriate to describe it than formalize.

I believe that this is the very essence of the vaguely felt crisis in the branches of the science of the mind dominated by computers. The instinct of formalization hinders the cumulative continuity and consensus typical for natural sciences. The (not so) new wave in cognition (Arbib, Grossberg and Carpenter) feeds on the physiological reality more than on the mathematical formality, stepping outside the circle.

Notes on locality

The principle of locality has a limited application in systems reaching a certain level of complexity, which needs the principles of hierarchy and centralization as its complement. Neither mind, nor society, nor many forms of life could not achieve high complexity

without centralization and long-distance interactions exemplified in the central nervous system, as well as the centralized processor and memory of the computer. Locality, however, evolutionary precedes complexity.

The language of modern society is a far cry from the primitive Nean. If by locality we understand the neighborhood of a link in a linear chain, local principles alone are by no means sufficient for natural language processing. The evolutionary language acquisition by hominids, as well as the individual acquisition by infants, works locally only up to a point, after which conscious learning, self-control and motivation take over. The exact position of this point is a separate question. There are two languages, one very ancient and limited but sufficient for homeostasis in a primitive environment of either primeval tribal life or modern nursery, and the other the sophisticated language of modern educated people, well fit for the purpose of complex scientific reasoning, rhetoric, irony, art, and deceit. This can be compared with our two brains: the recent *neocortex* and the old *medulla oblongata*. It is the same as to say that the primeval tribe was the nursery of civilization.

> The watershed between the two languages is marked by the invention of writing that allows for keeping a large number of words and ideas in the focus by fixing them on the writing medium (or is it the third language?). Of course, this can be done, to a certain degree, just by keeping everything in a trained memory. I doubt that Aristotle could be possible without writing. I wonder whether the ancient codes and epic poems, not even segmented into words, have ever been analyzed from the point of view of the focus size.

A different version of the principle of locality—**whatever happens with a complex system concerns a small part of it**—is strong, purely hypothetical, but more universal. Applied to Ashby's homeostat, it means that if the number of interacting units is large enough, most of them will be perturbed only slightly. This seems unlikely if each unit is directly interconnected with all the rest. There is a practical constraint on any interconnected system, however: it cannot be too large because of the fast growing number of connections. If the telephone system consists of binary lines between all individual customers, one hundred customers would need 4450 lines. The principle of locality of change and the constraint on the interactions, in a way, enhance each other and

apply a selective pressure on the developing system in the direction of what we actually see in nature, including the mind and society: limited sphere of interaction, wave of change, active area, hierarchy, codification, and bureaucracy. The school of fish and the judicial court are two examples.

> Modern life poses very peculiar problems. Thus, the transition from the relatively simple "physical" monarchy to more complex "organic" industrial democracy seems to have an effect directly opposite of what is attributed to democracy. Making rational (i.e., based on the analysis of alternatives and their consequences) decisions in a representative democracy can be much more difficult than in monarchy or oligarchy. Small subsystems cannot maintain homeostasis because they have no access to each other and the only real power belongs today to the media.

This completes my attempt to put the glass flask and the mind flask side by side, as in the chamber of an alchemist.

Next, in **Part 2**, I would like to come back to the analysis of language acquisition from triplets, started earlier [4] and show how a few local operations can bootstrap the process.

Let's talk
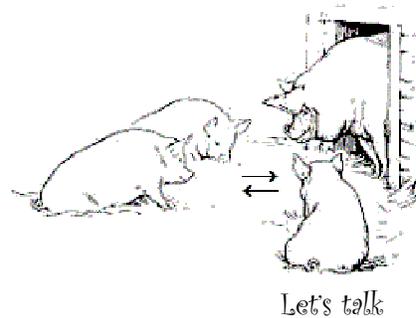
PART 2   The Three Little Pigs

Principles

The idea that understanding and generation of language operate with small chunks of text and are driven by the immediate environment of the word has been in the air for a long time, actually, since Claude Shannon's experiments with text generation and, even earlier, within the framework of structuralism. It is used in a very large body of work in computational linguistic.

When I wrote in [4] "the generators of language carry bits of grammar on their bonds like the bees carry pollen on their feet," I meant that in language acquisition a new word usually comes within a phrase, however short, from which the word identification and aspects of related grammar can be extracted. There is an experimental confirmation of the fact that it is not typical for a mothers to address the child with a single isolated word, see [12A, page 10].

I am not in the position to review the vast and technical but remotely relevant literature even selectively. I refer instead to [29], especially, its Part II. The contents of

the book and its very existence can give if not a bee's then a bird's eye view of the scale and variety of work in the area.

I would add just two representative references. The idea is that although the complete knowledge of language is needed for its computerized understanding, using only compact non-overlapping fragments (chunks, separated by chinks; term "chunking" is used in different ways, for example, as a mnemonic technique) is promising: "even if an exact solution is far beyond reach, a reasonable approximate solution is quite feasible" (Steven Abney, [30A]) The paper is interesting far beyond that remark, especially the chapter about language acquisition, but not by children. See also illustrative slides [30B].

I prefer to see chunks in language acquisition as "solidified" and overlapping (although they are defined in language processing as non-overlapping) neighborhoods of various but small enough radii. I also characterize the little child's command of language as rather approximate.

This extends even to adults, myself included. The Russian grammar, for example, could be as vexing as the English spelling. Even native speakers of Russian outside their milieu often demonstrate only an approximate command of the grammar.

What follows is that the narrow focus in language acquisition by hypothetical infant-robots who speak in doublets and triplets can be quite sufficient for their homeostasis. Both the first humans and the modern infants could speak dialects of Nean [5B] not because of some laws of nature requiring recapitulation of phylogenesis in ontogenesis but because of the universal requirement of simplicity for building complexity.

The notation used for **Part 2** includes symbols for generators, curly brackets $\{\ \}$, and double arrows $\Rightarrow$ and $\Leftrightarrow$, to which no mathematical or neurophysiological meaning should be attributed.

The symbols in curly brackets are **selected** from the entire generator space because they are, in some realistic sense, near each other. In Ulf Grenander's model of

the mind [3] this selection is called *content*: a small part of the very large mind, involved in a change, in other words, a dynamic focus of relevance.

In chemical imagery it is the small flask on the bench where only a few chemicals are interacting, while the lab has a large stock of various chemicals, some of which could be added during the process, some get into the flask by accident. Some products could escape as gas or precipitate.

I prefer to call the content between the brackets **focus**, by no means sharp, without linking it to the observable zones of activity in the brain. It is close to the reentrant dynamic core of Gerald Edelman [31]. I do not identify it with either short term memory or processing window, and Ulf Grenander's **content,** although it sounds best, is too exact for me. My **focus** is intentionally vague and filled with question marks. It should be better defined by its use.

There is a subtlety in chemical reality, ignored by chemists but essential for cognition. **All chemical transformations are irreversible**—a terribly heretic thing to say for a chemist, who believes—justly—in the basic reversibility of all elementary chemical acts. However, as soon as we notice that any chemical reaction must be irreversibly started by somebody at some moment in time and that nothing in the lab goes on its own, unless an earthquake, flood, or fire strikes, the analogy between chemistry, cognition, and evolution appears much closer. In fact, all three testify to the irreversibility of time in a way which is different from the approach of physics. This based on history principle of irreversibility is applicable to open systems that increase complexity and decrease entropy with time. While laws of physics are history-independent, the laws of evolution take note of unique events.

That the system cannot find its way back in a very large phase space, in spite of the micro-reversibility, is just one aspect of irreversibility. The other one is that the system has only a limited access to the larger system beyond its borders. Thus, two communicated people can exchange words and gestures but have a very limited ability to influence each other's mind and behavior, even at the gun point.

The basic principles, tentatively outlined in [4], can be substantially simplified. They remain, of course, entirely speculative, worse, dogmatic.

The unfolding of a complex open system from a small one is homeostasis: a diachronic sequence of alternating irreversible transition steps and subsequent states of a new equilibrium. Language acquisition is a chain of perturbations ($\Rightarrow$) interspaced with equilibrations ($\Leftrightarrow$).

The development of a small system through acquisition includes the following elementary mechanisms.

1. Irreversible acquisition steps that happen only once:

$$\{AX, BX\} \Rightarrow (X \in G); \qquad\qquad \text{generator acquisition (1)}$$

Meaning: If in the focus two generators A and B are bonded to X, then X is a generator. I use X not because it is something unknown but because it is **new**. As soon as it has been identified it becomes known, i.e., old for as long as it is remembered.

$$\{AX, BX\} \Rightarrow CX ; \quad (A, B \ \in C \in G); \qquad \text{class acquisition (2)}$$

Meaning: If in the focus two generators A and B can be bonded to X, then A and B belong to the same class C of all generators with affinity to X. The class acquisition can be also regarded as either categorization or pattern similarity acquisition, depending on the framework. I use *class* synonymously with *category*.

The left sides of (1) and (2) are identical, and both can be combined. Therefore, two **new** doublets, AX and BX with a common new block X , add a **new** generator X to the space and create a **new** class generator C which can be defined as all generators

connectible to X from the left (or right for $\{X—A, \ X—B \}$). A and/or B can be categories, so that classes can be expanded by new inputs.

When $X = \emptyset$, i.e., X is empty, and A and B are in the focus for the first time, a bond establishes between them in the sense of classical psychological association. They do not form a class, however, because a class is defined by its "classifier" capable of bonding with the entire class.

The two cases are, essentially, one because associations usually form against some common background which plays the role of X, not necessarily new. Nevertheless, the question whether associations in humans and animals require a non-empty X remains open. Will the Pavlov's dog be salivating when it is outside the lab while the bell sounds? The answer must be somewhere in the literature. This is a tempting problem.

2. Equilibriums:

$$\{A, B\} \Leftrightarrow \{AB\} \qquad \qquad \text{bonding equilibrium} \qquad (3)$$

$$\{A, B\} \Leftrightarrow C \ ; \ \ (A, B \ \in C \in G \ ); \ \text{class equilibrium} \qquad (4)$$

To combine both:

$$\{A, B\} \Leftrightarrow \{A, B, AB, C, X\}$$

or, further,

$$\{A\} \Leftrightarrow \{A, B, AB, C, X\}; \ \ \{B\} \Leftrightarrow \{A, B, AB, C, X\}; \ \text{etc.}$$

Meaning: A generator is in equilibrium with its neighborhood in horizontal and vertical (along the hierarchy) dimensions. This is yet another formulation of the locality principle. It follows from the principle that entire memory is in equilibrium, but the immediate neighborhood has the highest probability (equivalent of chemical concentration) to be retrieved into focus.

A reversible bond forms between A and B, similar to the association of two molecules of nitrogen dioxide NO_2 .

3. Irreversible spontaneous forgetting:

$$p\,(A \Rightarrow \emptyset) > 0\,,\tag{5}$$

where $\emptyset$ is the empty element and p is a nonzero probability.

Naturally, it is not so easy to lose a generator from the long term memory because it can always be re-equilibrated with configurations that contain it. If the generator disappears form memory, it can be later acquired anew, but always as the first time event.

Forgetting can be also deterministic and partially reversible, as for example, in case of repressed or corrected error.

For a simplistic view of memory, see **APPENDIX 4**.

4. The **identity constraint**:

$$\{A, A\} \Rightarrow \{A\}\tag{6}$$

Each generator exists and stored as a single copy.

This principle, actually, defines, so to speak, the systemic function of the mind, which Henri Poincaré [32] attributed to mathematics: to name many things with one name.

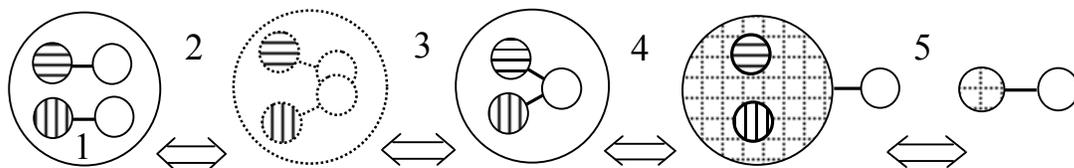What happens during the acquisition of language or knowledge in general is presented symbolically in **Figure 11** .

Figure 11. Processing of input in language/knowledge acquisition.
1: Input; 2, 3: Elimination of duplicates; 4, 5: Categorization

The four listed above principles make up a remarkably small and local set of rules. I would call them **simplicity principles** in knowledge acquisition.

It seems worth exploring where simplicity could lead us along the path of evolutionary expansion, which in our case is language acquisition. We overlook here the question about the initial pre-lingual set of generators and rules for knowledge acquisition which humans may share with animals. Another disregarded big question concerns, obviously, the neurophysiological mechanisms behind the rules.

Next we are coming to the main problem of both the mind and the flask. If any configuration is theoretically in an equilibrium with its components, why only a negligible part of all possible configurations appears in both focus and flask?

Regarding ideas, this question was first addressed by Henri Poincaré [32] who suggested that the ideas competed for the place in consciousness. The idea of competition in cognition is very common today. It was the chemist Manfred Eigen who first to give it a mathematical shape [33].

I try to never miss a chance to point to the roots of ideas. Scientists who discover common ancestors could be more inclined to respect family values.

The answer provided by chemistry is that only a few of configurations form **fast** enough because they have to overcome the barrier of a stressed transition state ("jumping to conclusions?"). Note that alternative chemical transformations also compete for the same initial product.

From the chemical perspective, the transition from one equilibrium to another goes through a transition barrier which is different for different pairs of stable states. Whether the transition barrier for cognition is measured simply by the distance between focus configurations is an attractively simple idea in itself, but it needs a separate discussion. In common language it means that if **cat** is in the focus, **dog** and **mouse** can be retrieved easier than **medicine** unless there is a catalyst, for example, **veterinarian**. It is the transition barrier that keeps us, literally, focused.

I believe that the study of the **kinetics** of transitions between the states, based on the study of **short-living stressed transition states**, is the key to understanding not only the mechanism of language acquisition but, more generally, the mechanism of the mind.

Earlier I have repeatedly described the theory of transition state [4,5], which is, I believe, the most important contribution of chemistry to the realistic theory of complex systems, and I will not repeat it here. This subject, introduced by **Figures 2** and **7** , should be considered separately.

My only purpose at this point is to illustrate how elements of grammar can be acquired not by analyzing a fragment of the entire synchronic text as a whole but by diachronically and incrementally updating and expanding the memory with each consecutive word in the input. Suppose the child robot is listening to a text for the first time. What can it derive from it? The words can be different but the principles will be the same.

This process is known as **incremental learning**, a topic of a substantial volume of literature, some well beyond language. Most interesting introductory discussion can be found in [11, 12A, 34, 35] and, for machine learning, in [36]. The spirit of simplicity is well articulated in this still underdeveloped area, initially overlooked in AI inebriated by computing power. In short, the system, natural or artificial, learns not from the entire set of data such as, for example, text or corpus, but starting with most simple inputs, comparing them with the available by that moment knowledge and updating the knowledge. Incremental learning is a separate subject, exciting and subtle, rising a set of fundamental questions, but I would like to comment here only on one aspect. A single step of incremental acquisition does not depend on any of the preceding steps, which are forgotten and leave no trace except the totality of knowledge. But the current state of knowledge may depend on the entire history of its acquisition. For different histories, the result can be different, and illustrations can be found in [12A], together with a summarizing discussion of the "less-is-more" hypothesis of Elissa Newport [12A, p.134, 35]. I find in this principle some support for my vision of two different languages, one of the Nean type [5B] and the other of the Aristotelian type, or of *The Wall Street Journal*, if you wish to blend into the crowd.

A sample of Aristotle:

For it is impossible that there should be demonstration of absolutely everything (there would be an infinite regress, so that there would still be no demonstration); but if there are things of which one should not demand demonstration, these persons could not say what principle they maintain to be more self-evident than the present one. (Aristotle. *Metaphysics*, Book 4; Translated by W.D.Ross)

A sample of Aristotle translated into Nean:

One cannot. One cannot demonstrate. Demonstrate everything. Absolutely everything. Infinite regress. Would be regress. No demonstration. One cannot. One cannot. If one. If one not. One not demand. Not demand demonstration. Demonstration. One not say. Not demand. Not say. Say what. Say what. What more self-evident. More than present. Present principle. No demonstration. No more self-evident.   No demonstration. No evident. No demonstration. No evident.

A better translation might be easier with the help of pictures on the sand.

To generalize, the only possible **explanation** of the particular state of a complex system is its history. This is why genealogy has been so important in tribal societies. In reconstructing evolution we try to complete the logical circle by going from the particular state back. As usual, to break the cycle, we have to step outside by turning to chemistry or meta-chemistry of the individual acts, **regardless of history**. We can study the elementary processes on observable examples picked up from the recorded history.

This is what makes the study of language acquisition a subject of general importance. What does actually happen day by day? If you want to understand a historical personality, you turn to his or her biography. The big personality just cannot hide it. Same applies to your humble partner, spouse, or friend, but this may be more difficult to uncover.

Illustrations

Armed with a small simplicity toolbox, let us look at language acquisition with the eyes of a naïve chemist. For an even higher degree of naiveté, see **APPENDIX 4**. Next degree of naiveté up, see **APPENDIX 5**. This section partially repeats and re-writes the **APPENDIX** to [4].

The following illustrations are by no means realistic because they do not deal with an audio input in an exchange with the environment, i.e., registering all utterances, actions, and facial expressions as a kind of symphonic score where a mother's smile is a syntactic marker, too. Besides, I take the **segmentation** into words for granted, for which there are sufficient reasons because infants seem to develop it very early [12A]. There are more reasons against it, however. It will be interesting to demonstrate that the word boundaries could be acquired basing on the general simplistic principles of generator acquisition but from the audio input.

The following text is a compact modified fragment from the tale of *The Three Little Pigs*. [37].

The input text P is a character array of 130 words given here in the form of MATLAB input:

```
P = char ('#', 'there', 'was', 'an', 'old', 'sow', 'with', 'three', 'little', 'pigs', 'and',
'as', 'she', 'had', 'not', 'enough', 'to', 'keep', 'them', 'she', 'sent', 'them', 'out', 'to',
'seek', 'their', 'fortune', '#', 'the', 'first', 'that', 'went', 'off', 'met', 'a', 'man', 'with', 'a',
'bundle', 'of', 'straw', 'and', 'said', 'to', 'him', '#', 'please', 'man', 'give', 'me', 'that',
'straw', 'to', 'build', 'a', 'house', '#', 'which', 'the', 'man', 'did', 'and', 'the', 'little',
'pig', 'built', 'a', 'house', '#', 'presently', 'came', 'along', 'a', 'wolf', 'and', 'knocked',
'at', 'the', 'door', 'and', 'said', '#', 'little', 'pig', 'let', 'me', 'come', 'in', '#', 'the', 'pig',
'answered', '#', 'no', '#', 'the', 'wolf', 'then', 'answered', 'to', 'that' , '#' , 'then', 'I', 'll',
'puff', 'and', 'I', 'll', 'blow', 'your', 'house', 'in', '#', 'so', 'he', 'puffed', 'and', 'he',
'blew', 'his', 'house', 'in', 'and', 'ate', 'up', 'the', 'little', 'pig', '#' );
```

Using a simple program, a **vocabulary** of 69 words, including # , which stands for the beginning or end of the sentence, was extracted from P and the text was presented as a sequence of overlapping triplets, see **Table 1**. The numbers, as for example, in **2-house** or **6-the** , refer to the number of occurrences in left or right position. The numbers in Column 2 record the order of the first appearance of the word and the overall occurrence in the text. Those numbers, as well as the sentence separators (#) are reported only for the sake of presentation and they are not supposed to be found in any form in the mind of the child robot, although the obvious effect of repetition can be the strengthening of memory and an increased probability of retrieval into focus.

The same was done also for the entire tale, see **APPENDIX 2**, **Table 2**, and a folk tale in Hungarian, **APPENDIX 3, Table 3**.

The complete **Table 1** is given in **APPENDIX 1**. Here we have only an initial fragment.

Table 1 (fragment) : **Vocabulary and neighborhoods of *The Three Little Pigs.***

1	2		Word	3
Left neighbors	No	Occur-rences	Word	Right neighbors
along, answered, fortune, him, 2-house, 2-in, no, pigs, 2-said, straw, them	1	16	#	I, a, little, no, please, 2-she, so, 6-the
#	2	1	there	was
there	3	1	was	an
was	4	1	an	old
an	5	1	old	sow
old	6	1	sow	with
man, sow	7	2	with	a, three

with	**8**	1	three	little
#, first, 2-the, three	**9**	5	little	4-pig, 4-pigs
little	**10**	1	pigs	#
……………	**…..**	…….	…….	…………………

There is nothing new in representing a string in triplets, i.e. 1-neighborhoods of the words. The immediate purpose of this artificial model is neither practical, nor theoretical, but simply an exercise in further establishing a parallel between chemistry and cognition.

Unlike SOM and other synchronic models, but conforming to one of the basic ideas of ART, the text should be considered as a diachronic sequence of first and subsequent appearances of the word, together with its left and right neighbors, in a very narrow focus of attention, but certainly not less than three words. The focus (moving window of perception and processing) can be much wider if a block of input comes within a limited time and is more or less remembered as a whole. This situation occurs when a child (and adult, as well) listens to a short story or tale for the first time.

When the focus is so narrow, the chemical interpretation of the model is the string of text drawn through a vessel as if it were a flow of reagents through a tubular reactor. What comes out of the reactor are forgotten or unnoticed ("unreacted") words. The template synthesis of DNA, RNA, and proteins offers another close analogy.

For our purpose we can neglect the forgetting and omitting, so that the tube turns into a flask fed with the string of text. What we cannot ignore is the cardinal difference from chemistry, which is the constraint of the single copy for each word inside the flask: $\{A, A\} \Rightarrow \{A\}$.

In the flask every incoming word is recognized as either old or new (Compare with the SCALE model in [3]).   Only the new words add to the content of the flask, while the old ones leave no material trace except strengthening the memory trace. The new word becomes instantly old after it is deposited in the memory.   The "chemical reaction" must run sufficiently fast.

If the focus of attention is wide enough, we can simplify the model and see it as a flask filled with the block of text and left for processing, which is a typical situation for listening to a very short story. Any discussion of details, however, should be better left to the linguists and cognitive scientists.

The short story situation, taken only as an illustration, is not suitable as a model for the earliest stages of language acquisition, in which short blocks of "Motherese" speech come in an irregular fashion, separated by intervals of various length and strongly tuned to environment.

Note that I assume that the child, at least from my own experience, never reads any scientific literature and knows nothing about Hebbian learning, Bayesian inference, linguistics, and many other things. In other words, the child has neither a homunculus in the brain nor any substantial knowledge database in storage.

Let us first take the first 26 words as the input. It is perceived as the string of sounds, in which apostrophe (') is used only for the convenience of reading :

THERE'WAS'AN'OLD'SOW'WITH'THREE'LITTLE'PIGS'AND'AS'SHE'HA
D'NOT'ENOUGH'TO'KEEP'THEM'SHE'SENT'THEM'OUT'TO'SEEK'THEI
R'FORTUNE

In reality there is much more to it, like prosody and emotional emphasis, which are disregarded not only by formal linguistics but also by most transcripts in child language corpora. To me it seems like studying the behavior of gas by measuring temperature and disregarding pressure. How crucial such aspects are, see [12A]. They could also be technically challenging to record and study.

Even the above short fragment is so long for an untrained mind that it has no chance to stay in focus as a whole. Suppose, for the sake of argument, it can. The generator identification and classification require at least two distinct fragments in the focus. There are no identified A and B in the above fragment, however, not to mention its excessive length. This simulated problem illustrates how **simplicity principles** work: they really have to start from scratch. To observe language acquisition or, for that matter, any evolution, we have to go to its very beginning, for which we never have a perfect

model, unless we have a time machine, but language acquisition and embryology are rare good surrogates.

Language acquisition, probably, requires a break-in stage where a limited pool of atomic words is identified by objects and events in the environment without any help from other words. As soon as this initial pool has accumulated, the words can be used for identifying new ones in fragments of increasing length. Sensory impressions, therefore, are the launching pad for bootstrapping.

Figure 12 illustrates the general scheme of simplicity by using visual inputs as generators **A** and **B** and audio input as **X**.
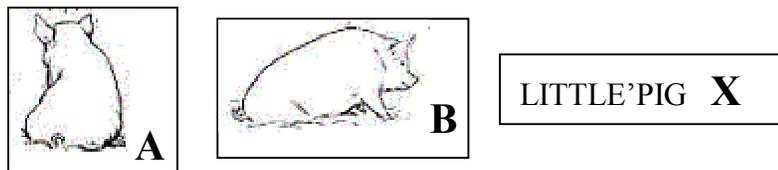
$$\{AX, BX\} \Rightarrow X \in G \text{ if X is new}$$
$$\{AX, BX\} \Leftrightarrow CX ; \ C : \text{"Class of all spatial positions}$$
of the real little pig"

Figure 12. Acquisition of a single speech generator LITTLE'PIG using images A and B as generators

The short fragment is represented below in such a way that repeated words are emphasized by color and font.

'there'was'an'old'sow'with'three'<u>little'pig</u>s'<u>AND</u>'as'**she**'had'not'enough'to'keep'**t hem**'**she**'sent'**them**'out'**to**'seek'their'fortune'#'**the**'first'that'went'off'met'**a**'man' with'**a**'bundle'of'straw'<u>AND</u>'said'**to**'him'please'man'give'me'that'straw'**to**'build' **a**'*house*'which'**the**'man'did'<u>AND</u>'**the**'<u>little'pig</u>'built'**a**'*house*'#'presently'came 'along'a'**wolf**'**and**'knocked'at'**the**'door'**and**'said'<u>little'pig</u>'let'me'come'in'#'**the**'pig' **answered**'no'**the**'**wolf**'then'**answered**'**to**'that'then'<u>I'LL</u>'puff'<u>AND</u>'<u>I'LL</u>'blow'your '*house*'in'so'HE'puffed'<u>AND</u>'HE'blew'his'*house*'in'<u>AND</u>'ate'up' **the**'<u>little'pig</u>';

We can see how **HE** and **she** classifies the right neighbor as a verb, **me** classifies the left word as a verb , **the** precedes a noun (**the**'first needs further specification) , etc.

The word **to** is ambiguous. It is used in the above short fragment as

out to seek , said to him , straw to build, answered to that , which is seen from the row of the table:

answered, enough, out, said, straw	5	**17**	**to**	build, him, keep, seek, that

It is used in the long fragment as

said to him, straw to build, furze to build, bricks to build, did to the other, want to go, hoped to get, had to climb, gone to pick, got to the fair, what to do, churn to hide, been to the fair, he wanted to get.

been, bricks, churn, did, furze, gone, got, had, hoped, said, straw, want, wanted, what	14	**30**	**to**	3-build, climb, do, 2-get, go, hide, him, pick, 3-the

The usage of **to** easily differentiates if the verbs and nouns/pronouns are represented as classes **V** (verb) and **N/proN** (noun/pronoun) or **V** and **non-V** .

V to non-V (or V to N/proN)

said to him, did to the other, got to the fair, been to the fair

N to V

straw to build, furze to build, bricks to build, churn to hide

There is a simpler principle: classes **V to-** X and X **-to V** , i.e., left or right positions of the verb regarding **to** and another word.

Note that classes correspond to patterns of PT where class is called a partition of the generator space and the similarity transformation is permutation.

The acquisition proceeds as an unsupervised updating by new inputs, i.e., incremental learning. The previous inputs do not need to be stored in memory. The current input is absorbed, digested, and forgotten. Such systems are formally regarded as **memoryless** , although the term has never been clearly defined, except in probabilities, and the term is misleading because the growing knowledge expands the content of the memory.

By definition, however, the **growing** acquisition system changes each time a **new** input is presented. Moreover, as I noted earlier, the state of the system depends on the previous path towards it. To compare with chemistry, the same reagents produce different result when added drop by drop in different order and quantities. The template synthesis of DNA and proteins is another, more common, illustration of a path-depended process. So are Lego constructs and all the other instances of structured discrete growth. For all such processes of growth (synthesis, as a chemist would say), the particular final result depends on the order of elementary acts and it can be achieved through different pathways, see **APPENDIX 4**. The memory content, so to speak, is the system itself, without any memory of the path.

The visual metaphors of the incremental learning and memory itself are in **APPENDIX 4**.

Let us examine some further illustrations of simplicity. We assume that if two words occur together, they have strong enough affinity to each other.

We encounter in the text some doublets and triplets of high occurrence in everyday speech, for example:

2	there	was

15	not	enough

54	let	me

39	give	me

she	14	had	not

she	20	sent	them

In such frequently used combinations the equilibrium is shifted toward the bonded state and the doublet or tripled functions as a single generator. This is where the transition from Nean toward the language of *The Wall Street Journal* begins: complex semantics starts driving the syntax and the speaker and writer operates with ideas instead of words.

The term **word**, however, has a fuzzy meaning in speech. We write the compound words **newborn** and **nevermore** without spaces but separate **not enough** and **let me**. There is, undoubtedly, linguistic literature on the subject, but it is rather obvious that other languages have different segmentation habits, especially the German language with its compounds, not to mention Chinese and Thai.

Equilibrium seems to me an elegant way to formulate the situation:

$$\{ \text{ \textbf{new} , \textbf{born} } \} \quad \Leftrightarrow \quad \{\textbf{newborn}\}$$

The position of the equilibrium can be shifted, which is, probably, the way language evolves. Thus, the equilibrium is shifted to the right in

$$\{ \text{ \textbf{strong} , \textbf{hold} } \} \quad \Leftrightarrow \quad \{\textbf{stronghold}\}.$$

I can even suggest, with the carelessness of an outsider, that the shift of equilibrium toward association is a general thermodynamically driven trend of language, so clearly seen in English. It is interesting whether it manifests in modern German. The origin of functional morphemes seems to fit this thermodynamically driven process. Today time is money, but long ago it used to be life, so better say it quick.

In chemical lingo, the doublets and triplets, if used frequently, can **crystallize** and form composite generators, provided the abstract **temperature** is low enough. In P, however, the statistics is meaningless because of the small size and the batch form.

I believe the position of equilibrium can be studied experimentally, but this should be better left to professionals.

The following is a series of examples of what can "chemically" happen with P as a substrate.

1. Let us take the row

7	25	the	door, first, 2-little, man, pig, wolf

THE creates the tentative class of all words right of THE. Such classification may further survive, as it in fact does, or fall apart. We need a name for the class, of course. As a Martian chemist who knows no English grammar, I would call it **THE-X**. Like a chemical formula, this is an iconic ideogram. I could also use THE—⊗ or ▲—○.

Class THE-X, X= {first, man, little, door, pig, wolf}

We know that **THE-X** includes both nouns and adjectives, but the child-robot does not know linguistic terminology.

2. Similarly:

5	31	a	bundle, 2-house, man, wolf

Class A-X: X= {man, bundle, house, wolf }

These two classes can be expressed in terms of the vocabulary entries.

Since the class is in equilibrium with its entries, MAN is in equilibrium with all its classes.

Class X-MAN : **X = {a, the}** and, therefore, **X= { A-X , THE-X }**.

Fed by many occurrences, this classification will, most probably, survive.

But then $\mathrm{A\text{-}X}$ and $\mathrm{THE\text{-}X}$ form also a class, for which we are out of names other than cumbersome $\mathrm{A\_THE\text{-}X}$. Of course, we now know the current name of the class: **article**, by the way, absent in many inflected languages.

Now we can switch from Martian to English.

3. Similarly, **pig** builds the class of verbs. The outsider **that** may remain in the class for a while before it loses to competition with other patterns.

4	45	pig	answered, built, let

The entire process of acquisition is a series of "triangulations" that gradually build the pattern of syntactic topology.

3. Similarly:

an	1	5	old	sow

and

2-the, three	4	9	little	3-pig, pigs

would allow for inferring the distinction between nouns and adjectives, not quite reliable
yet:

Class   X-Adjective-Y: X=Article, Y= Noun

In Hungarian, where articles exist and adjectives do not change, the mechanism of
acquisition will be the same as in English. In Russian, where articles are not known, the
adjectives are all clearly marked by endings or their absence: **old**-she **pig**-she ,
little-he piglet-he (**Стар**ая **свин**ья, **маленьк**ий **поросенок**).

A Martian chemist can wonder whether the total drive to aggregation, due to some
abstract Ice Age, eliminated the article from proto-Russian and adjective endings from
English and Hungarian. A Martian philosopher can mull an idea that culture is a great
refrigerator that prevents life from rotting quickly. I, on my part, consider culture a
thermostat-incubator for life: not too hot, not too cold for homeostasis, but warm enough
for change.

Is homeostasis a circular concept? No, because it can be objectively registered by outside
observers, such as Piaget, Ashby, and Gould.

There is not enough data to form the class of nouns, however, but we can easily
imagine that with enough verbs.

The above examples could make us feel what a little child feels when acquiring
the knowledge of the world: what we know and what seems elementary and obvious to us
must to be retrieved from the formless mass like the statue of David from the block of
marble. Unlike the sculptor, who cannot make a big mistake with the stone, the child's
mind works like a scientist—or living nature—creating, testing, and rejecting hypotheses.
There is a striking similarity in the processes of the individual acquisition of language by
children and the 150 year long acquisition of chemical knowledge by chemists,

generations of whom were busy putting two and two together. Both have to work with incomprehensible data, using the available knowledge and pulling it one step up.

Under circumstances, one can prolong the early childhood,. I remember how in Russia, at the age of 30, I repeatedly heard a song on the radio but could not understand what the words "**cash register of the mountains**" could mean, until one day it dawned on me: it was "**hill side.**" The two are phonetically identical (**kasa gor** and **kasagor**). The tune created the problem by putting two stresses on the single word. The song was about dangerous driving. I am sure that if I had a car, which I, as practically everybody in Russia, had not, I would catch the right meaning immediately.

Similarly, as a child, I heard my father speaking Yiddish, which I was never taught. I had always perceived "**A sochen vey**" ("such a misfortune") as a single word, until much later I started to study German and was able to parse the Yiddish expression.

Learning to read at the age of 6, I had a painful problem with the words "**on slez**" (he descended [from a tree]). In Russian, **slez** (слез) means **[he] descended**. **Slyoz** (слёз) is Genitive of **tears.** The problem was that the umlaut over е (ё) is traditionally not printed in Russian, unless absolutely necessary. As I interpret that case now, crying and weeping was something much more familiar to me than climbing trees, but the relevant cue in the form of two dots over е was missing.

The overall picture of language acquisition—and, therefore, of language genesis—becomes the field for competition of patterns, which are counterparts of biological species, and not individual sentences. When the starting pattern is as simple as doublet or triplet, further mutations can generate the largest variety of grammars, which explains why the languages are on the surface so different: they all came from the simple protolanguage which can be modified an a multitude of ways. The mutations of developed grammars are, of course, less radical. After the advent of book printing and general education they are very rare.

The German separable verb prefixes seem to contradict the principle of locality, but if we start with simple situations and short phrases, German with its **auf** and **an** at the end is no more strange than Japanese with its verb invariably at the end or the Arabic

words with the vowel inflections between the root consonants. Moreover, all that linguistic zoo is no more strange than the real zoo with elephant, snake, and parrot.

We can hope to reconstruct the process of linguistic genesis for two reasons: (1) we can understand the world of the first speakers where somebody does something to somebody or something, but not much more, and (2) we have only two choices for adding a new generator: either left or right of an old one.

I hope that the short fragment illustrates to an enthusiast the main principle: **if the words were atoms**, **there would be a chemistry of words**., leaving no place to any infinity. The "chemical" mechanism is local, it seems to lead to a distributed intelligence, and it works on many levels, from morphemes to phrases, creating, literally, a distributed grammar. I would really recommend to a linguist to browse through an introductory textbook of organic chemistry in search of inspiration.

The above and subsequent examples illustrate nothing but a hypothesis. Its further development, as well as comparison with other linguistic models of acquisition should be better left to those off-beat bees who might become attracted by the chemical smell of strange flowers. The entire direction of Darwinian linguistics, started by Manfred Eigen and continued by Martin Novak and others [33] becomes the only meadow where the nourishing flowers can grow.

Next let us increase the size of the text to the full bitter-sweet story. Its length is 775 words, maximal length 9 characters, but it contains only 167 different words. Its complete input (**P**) and output (Table 2) are given in **APPENDIX 2**.

Here are the beginning, a middle fragment, and the end of **Table 2**:

TABLE 2 (fragments)

LEFT		No	NAME	RIGHT
2-afraid, after3-noon, aga6-in, 2-along, angry, 2-answered, apple, 2-apples, bricks, came, chimney, 2-churn, 2-com6-ing, 2-d6-inner, do, 2-down, fair, far, field, fire, furze, garden, go, happened, hide, hill, him, 2-home, 3-house, 6-in, 2-it, late, 3-no, oclock, other, out, 2-pig, 2-pigs, 2-ready, replied, 15-said, 2-straw, 2-them, three, time, tomorrow, tree, turnips, up, well, w2-ith, wolf, 2-yes, 2-you	96	1	#	9-I, a, at, but, 11-he, 2-if, in, 2-man, no, 8-pig, please, s11-he, t11-he, t11-hen, very, 2-we, 2-what, w11-here, will
#, been, pig, 2-where	5	2	there	and, 3-is, was
and, he, there, 3-wolf	6	3	was	2-afraid, an, gone, late, very
down, was	2	4	an	apple, old
an	1	5	old	sow
old, the	2	6	sow	sent, with
hill, 3-house, 3-man, sow	8	7	with	#, 3-a, it, the, them, three
at, with	2	8	three	#, little

. . . .

LEFT		No	NAME	RIGHT
call, come, food, off, turnips	5	16	for	dinner, the, them, 2-you
for, sent, with	3	17	them	2-#, out
45-#, 5-?, and, are, 2-at, 3-before, blew, 2-blow, climb, day, 2-down, for, 2-gave, got, him, in, 4-in3-to, 2-saw, threw, 3-to, 3-told, 3-up, with	85	18	the	apple, 2-apples, bricks, chimney, 3-churn, door, 2-fair, field, fire, first, furze, garden, 2-hill, 3-house, 3-man, next, other, 28-pig, second, sow, straw, third, time, tree, turnips, 24-wolf
sow	1	19	sent	them
them	1	20	out	#
the	1	21	first	little
8-#, little, second, 28-the, third	39	22	pig	2-#, 2-I, answered, 2-are, 3-built, 3-got, in, jumped, 3-let, made, 3-met, 11-said, saw, somehow, there, threw, told, went
3-pig	3	23	met	3-a

. . . .

the	1	164	chimney	#
pig	1	165	made	up
a, the	2	166	fire	2-#
wolf	1	167	fell	into

The middle fragment illustrates the significance of frequent words: some carry a syntactic function (**the**), while others define a topic for reasoning (**pig**).

In order to calculate the occurrences, whether a human or a computer needs to pass the entire text, performing the following operations:

1. Accept the word
2. Check if it is new.
3. If new, create a counter with 1 in it.
4. If not, add 1 to the corresponding counter.

It seems that this algorithm can performs on-line incremental learning because there is no need to keep the text in memory. In fact, the structure of the text is lost and it is not known which word follows which. This is not important because it is the grammar that must be extracted. What is more important, the identical words can be separated by a long distance.

Let us take the row:

down, was	2	4	an	apple, old

"**An old**" first comes when "**an"** is the word No.3. Next time "**an"** is on the left of "**apple**" when "**an**" is word No. 143. A good memory is needed to remember the first occurrence. Infants hardly can have it.

In order to perform classification of the words, the second member of the doublet must be remembered. For example, in the following line (No.2)

#, been, pig, 2-where	5	2	there	and, 3-is, was

"**there was"** occurs right in the beginning of the text because "**was**" is the second word, but "**there is"** appears when "**is"** comes as word No.82.

We can see from the beginning of the **Table 1** that numerous repetitions allow for the inference of the pattern Article—Subject—Verb (**The pig said**), which is already a good deal of child's grammar, but, probably, not enough for mastering direct and indirect objects.

The main postulate of simplicity is that the generators for inference must be in a relatively narrow focus. This is possible when language acquisition starts with extreme simplicity and (I can't believe I am writing it!) **poverty of stimulus**. As soon as categories develop, only they must be kept in memory "as the pollen on the bee's feet" (self-quotation), together with the bees, of course, and a relatively poor assortment of syntactic patterns. This is what I believe is the earliest stage of language acquisition, corresponding to the earliest stages of language evolution.

APPENDIX 3 contains a preliminary simplistic analysis of a Hungarian folk tale [38] for those who would like to experiment with an unfamiliar language. A Chinese text would be a good example, too, but it is technically cumbersome for me.

The choice of whole written words as generators in Hungarian is certainly unproductive. Highly inflected and agglutinative languages should be, probably,

represented not in terms of words but in terms of syllables or smaller elements which can be still easily denoted by parts of written words. This seems to be the most natural way for most languages because phonology and morphology meet at the syllable. Mora is another option, but the discussion over it has not yet settled down.

A syllabic triplet representation of Hungarian is among my next plans.

APPENDICES 4 and **5** will speak for themselves in the language of 3 month old infants.

===

CONCLUSION

The phenomenon of growth is a particular case of evolutionary process in which a configuration is increasing its size in an increasing generator space. It can be described in terms of PT as expansion of the generator space. Mathematically, it is modeled by Bourbaki's scale of sets, see [5A] about that. The basic set of the scale is an alphabet, for examples, the set of elementary sensory inputs, whether original or processed, as in written language.

Three important concepts have been expressed systematically or casually in literature.

1. Pattern Theory (Ulf Grenander) as the most general theory of atomistic structures.

2. Evolution by triangulation (Gerhard Mack).

3. The principle of "less is more" in unsupervised incremental learning (Elissa Newport, Jeffrey Elman, Christophe Giraud-Carrier). Stephen Grossberg's ART is also related to this direction of thought. "Winner takes all" is a version of "less is more," naturally, because there is no gang to share the booty.

To these ideas, among which only the first one has been fully developed, I would add my still vague and crude idea of cognitive chemistry based on local non-algorithmic interactions. Probably, this is not new, either.

The comparison of the prediction with reality of language acquisition is a separate task. I am not interested here in solving any practical problem, like automatic translation—I am certain the problems will be solved anyway—but in understanding how human mind works. I suspect that it works, in many aspects, as the chemical system which long ago generated life, mind, and society. In the future, this knowledge may well be of great help in designing new gadgets to sell and new wireless chains to wear, but that kind of future has little appeal to me.

More specifically, I wanted to show that there could be some other approaches to language acquisition than the icy Cubist Universal Grammar at one end of the scale and equally icy and Kafkian Bayesian counting at the other end. There could be an entirely different dimension of non-algorithmic, distributed, random, thermodynamically constrained, kinetically driven, and mostly local computation—a real bootstrapping of the newborn mind. Most ideas are already in the air. And, yes, the bees are, too, doing their job of cross-pollination.

REFERENCES

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

1. Baker, Mark C. *The Atoms of Language*. New York: Basic Books, 2001.
   Mark Baker's publications: http://ling.rutgers.edu/people/faculty/baker.html

2A. Grenander, Ulf. *Elements of Pattern Theory*. Baltimore: Johns Hopkins University
   Press, 1995.
   Advanced works:

2B. ———. 1976. *Pattern Synthesis. Lectures in Pattern theory*, Volume 1. New York:
   Springer-Verlag, 1976.

2C. ———. 1978. *Pattern Analysis. Lectures in Pattern theory*, Vol. II. Springer.

2D. ———. 1981. *Regular Structures. Lectures in Pattern theory*, Vol. III. Springer.

2E. ———. *General Pattern Theory. A Mathematical Study of Regular Structures*, Oxford,
   New York: Oxford University Press, 1993.

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

4. Tarnopolsky, Yuri. *Pattern Theory and "Poverty of Stimulus" argument in linguistics.*
   http://spirospero.net/Poverty of stimulus.pdf

5A. ———. *Molecules and Thoughts: Pattern Complexity and Evolution in Chemical
   Systems and the Mind* , 2003.
   www.dam.brown.edu/ptg/REPORTS/MINDSCALE.pdf
   Or: http://spirospero.net/MINDSCALE.pdf

5B. ———. *Tikki Tikki Tembo: The Chemistry of Protolanguage*, 2004
   http://spirospero.net/Nean.pdf

5C. ———. *Transition States in Patterns of History*. 2003.
   http://spirospero.net/HistMath1.pdf

6. 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.

7. Rosch, E. 1977. *Human Categorization*. In N. Warren (ed.) *Studies in Cross-cultural Psychology*. London: Academic Press, vol. 1, pp. 1-49.

———. 1978. *Principles of categorization*. In E. Rosch and B. B. Lloyd (eds.), *Cognition and categorization*. Hillsdale, NJ: Erlbaum, pp. 27-48.

8. Edelman, Shimon. (1999). *Representation and recognition in vision*. Cambridge, MA.: MIT Press.

———. Numerous downloadable publications at :

http://kybele.psych.cornell.edu/~edelman/archive.html

Shimon Edelman's Experimental Epistemology Project

http://kybele.psych.cornell.edu/~edelman/

Among them, a most recent overview of Edelman's theory:

———. *Bridging language with the rest of cognition: computational, algorithmic and neurobiological issues and methods, (2005)*

http://kybele.psych.cornell.edu/~edelman/Archive/EMCL03-Edelman-chapter-final.pdf

9. There are a lot of materials on Neural Nets and connectionism in the shifting sands of WWW. Here are some:

Medler, David A. A brief History of Connectionism,

www.cs.rhul.ac.uk/NCS/vol1_3.pdf in: Neural Computing Surveys

www.cs.rhul.ac.uk/NCS/

Itti, Laurent, Lecture Notes at: http://ilab.usc.edu/classes/2002cs564/

For example, http://ilab.usc.edu/classes/2002cs564/ lecture_notes/08-Hopfield-Networks.pdf

An important meta-site:

CARLESCO: *The Complexity & Artificial Life Research Concept* for Self-Organizing Systems http://www.calresco.org/

FAQs, Introductions and Tutorial: http://www.calresco.org/tutorial.htm

Links to Journals & Resource Sites: http://www.calresco.org/links.htm#neur

Gurney, Kevin, Neural Nets by Kevin Gurney:

http://www.shef.ac.uk/psychology/gurney/notes/index.html

_____. (1997) An Introduction to Neural Networks. London: Routledge.

Bullinaria, John A. Introduction to Neural Networks - Course Material and

Useful Links,   http://www.cs.bham.ac.uk/~jxb/nn.html

10. Grossberg, Stephen, Publications :

http://cns-web.bu.edu/Profiles/Grossberg/onlinepub.html

11A. Elman, Jeffrey, Publications: http://crl.ucsd.edu/~elman/

11B. _____. 1993. Learning and development in neural networks: The importance of

starting small. *Cognition,* 48 (1993) 71-99.

http://crl.ucsd.edu/~elman/Papers/elman_cognition1993.pdf

12A.   Jusczyk, Peter W. (1997) *The Discovery of Spoken Language* , Cambridge, MA:

MIT Press.

12B. Hopkins Psychologist Peter W. Jusczyk, 1948-2001.

http://www.jhu.edu/news_info/news/home01/aug01/jusczyk.html

13A. Bootstrapping http://www.wikiverse.org/bootstrapping

13B. *Yuret, Deniz.* Bootstrapping Acquisition.

http://home.ku.edu.tr/~dyuret/pub/yuretphd/node4.html

in: Yuret, Deniz. 1998**.** *Discovery of Linguistic Relations Using Lexical Attraction.*

http://home.ku.edu.tr/~dyuret/pub/yuretphd/main.html

13C.   Gleitman, L. (1990). The structural sources of verb meanings. *Language*

Acquisition, 1, 3-55.

13D.   Gleitman, L. & Gleitman, H. (1992). A picture is worth a thousand words, but

that's the problem: The role of syntax in vocabulary acquisition. *Current*

Directions in Psychological Science, 1, 31-35.

13E. Thelen, Michael and Riloff, Ellen. *A Bootstrapping Method for Learning Semantic*

Lexicons using Extraction Pattern Contexts .

http://www.cs.utah.edu/~riloff/psfiles/emnlp02-thelen.pdf

Other publications of E. Riloff, see:

http://www.cs.utah.edu/~riloff/publications.html

14. Bates, Elizabeth and Goodman, Judith C. 1997. On the inseparability of grammar

and the lexicon : Evidence from acquisition, aphasia and real-rime processing

Language and Cognitive Processes, 1997, *12*(5/6)*,* 507-584.

http://crl.ucsd.edu/~bates/papers/pdf/bates-goodman-1997.pdf

15A.   Mack, Gerhard. 2001. Universal Dynamics, a Unified Theory of Complex Systems.

Emergence, Life and Death. *Communications in Mathematical Physics*, **219**, No.1, (141 – 178).

15B.   ———. 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.

16. Barrett, H. Clark. Enzymatic computation and cognitive modularity. Mind and Language. http://www.anthro.ucla.edu/faculty/barrett/barrett-enzymes.pdf
All publications: http://www.anthro.ucla.edu/faculty/barrett/research.htm

17. Klinger, Walter. Learning Grammar by Listening. Academic Reports of The University Center for Intercultural Education, The University of Shiga Prefecture, No. 6. Hikone, Japan.
http://www2.ice.usp.ac.jp/wklinger/QA/articles/kiyou2001/kiyou2001.html
Publications: http://www2.ice.usp.ac.jp/wklinger/research.htm

18. Honkela, T. , Ville Pulkki, and Kohonen, T. 1995. Contextual Relations of Words in Grimm Tales Analyzed by Self-Organizing Map. *Proceedings of International Conference on Artificial Neural Networks*, ICANN-95, F.Fogelman-Soulie and P.Gallinari (eds.), EC2 et Cie, Paris, 1995, pp.3-7.
http://websom.hut.fi/websom/doc/grimmsom.ps.gz
Timo Honkela's publications: http://www.cis.hut.fi/tho/publications/

19. John Goldsmith, John. Review of *The Legacy of Zellig Harris.*
http://humfs1.uchicago.edu:16080/ ~jagoldsm/Papers/ZelligHarris.pdf

20. Sumerian grammatical examples compared to Hungarian.
http://users.cwnet.com/millenia/Suwordorder.html
More about that can be revealed by search.

21. Marshall, Brian. *Selfridges's Original Pandemonium*.
http://www.agt.net/public/bmarshal/aipatterns/pan_orig.htm
Boeree, George. *Pandemonium*.
http://www.ship.edu/~cgboeree/pandemonium.html

22. Edelman, Shimon. CogSci03-poster. Slide 24.
http://kybele.psych.cornell.edu/~edelman/VL/CogSci03-poster.html
http://kybele.psych.cornell.edu/~edelman/VL/siframes.html

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

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

23. Autopoiesis on PRINCIPIA CYBERNETICA WEB

     http://pespmc1.vub.ac.be/ASC/AUTOPOIESIS.html

   Quick, Tom. Autopoiesis.   http://www.cs.ucl.ac.uk/staff/t.quick/autopoiesis.html

24. Russian Stories and Folklore

     http://www.story-lovers.com/listsrussianstories.html

25A. Consciousness for Neural Networks (1997). *Neural Networks*, Volume 10, Issue 7,

   1 Oct.   ftp://math.chtf.stuba.sk/pub/vlado/NN_books_texts/NeuralNetworks_Consciousness.pdf

25B. Rolls, Edmund T. 1997. Consciousness in Neural Networks? *ibid*., p. 1227–1240

26. 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.

27A. Darrah, Anita. Jean Piaget. http://shell.world-net.co.nz/~darrah/piaget.htm

27B. Rodriguez, Monica L. Rodriguez. Piaget's Cognitive-Developmental Theory.
   http://blue.csbs.albany.edu:8000/203/piaget.html

28. Eldredge, N. and Gould, S.J. (1972). Punctuated Equilibria: an alternative to
    phyletic gradualism. In: *Models in Paleobiology*, edited by Schopf, San Francisco:TJM Freeman, Cooper & Co, pp 82-115.
   http://www.blackwellpublishing.com/ridley/classictexts/eldredge.pdf
   Some S.J.Gould's publications: http://www.stephenjaygould.org/library.html
     At: The Unofficial Stephen Jay Gould Archive, http://www.stephenjaygould.org/

29. *The Oxford Handbook of Computational Linguistics* .2003. Edited by Ruslan
   Mitkov. New York: Oxford University Press

30A . Abney, Steven. 1996. Tagging and Partial Parsing**.** In: Ken Church, Steve Young,
    and Gerrit Bloothooft (eds.), *Corpus-Based Methods in Language and Speech.* Dordrecht: Kluwer Academic Publishers. www.vinartus.net/spa/95a.pdf
     Steven Abney's Publications: http://www.vinartus.net/spa/publications.html

30B. Neumann, Günter, Shallow Natural Language Parsing.
     http://www.dfki.de/~neumann/slides/GN-snlp.pdf

31. Tononi, Giulio and Edelman, Gerald . 1998. Consciousness and Complexity.
    Science, Vol. 282, 4 December 1998.

http://www.ini.unizh.ch/~kiper/Tononi_1998_consc_complex.pdf

32. Poincaré, Henri. 1946. *The Foundations of Science*, Lancaster, PA: The Science Press.

33. Nowak, Martin, 2002. From Quasispecies to Universal Grammar*, Z. Phys. Chem*. **216** (2002) 5–20. http://www.ped.fas.harvard.edu/pdf_files_old/ZPhysChem02.pdf Martin Novak's publications: http://www.ped.fas.harvard.edu/publications.html

34. Kirby, Simon and Hurford, James R. 1997. The evolution of incremental learning: language, development, and critical periods.

35A. Newport, Elissa. (1990). Maturational constraints on language learning. *Cognitive Science, 14,* 11-28. http://www.bcs.rochester.edu/people/newport/newport1990.pdf

35B. _____. Publications http://www.bcs.rochester.edu/people/newport/newport.html

36A. Giraud-Carrier, Christophe and Martinez, Tony. 1994. An Incremental Learning Model for Commonsense Reasoning . In *Proceedings of the 7th International Symposium on Artificial Intelligence (ISAI'94)*, 134-141. http://faculty.cs.byu.edu/%7Ecgc/Research/Publications/ISAI1994.pdf

36B. Giraud-Carrier, Christophe. A Note on the Utility of Incremental Learning http://faculty.cs.byu.edu/~cgc/ Research/Publications/AICOM2000.pdf

36C. _____. Publications. http://faculty.cs.byu.edu/%7Ecgc/Research/pubs.html

37. Jacobs, Joseph (1890). The Story of the Three Little Pigs. *English Fairy Tales*. London: David Nutt. http://www.surlalunefairytales.com/index.html

38. A só . www.magyarora.com/literature/Benedek_so.pdf

APPENDICES

Appendices 1 and **2** contain the short fragment and the complete Tale of *the Three little Pigs*, represented as triplets, as well as the corresponding input character arrays. **Appendix 3** contains triplet representations of fragments of the Hungarian Tale *Salt* (*A Só*) as a sample of a highly inflected language. A possible way to pattern analysis of such languages is to start with a syllabic input.

        Appendices 3, 4, and **5** mark points of departure for some future investigations.

APPENDIX 1

Table 1 : **Vocabulary and neighborhoods of *The Three Little Pigs.***

        The short fragment.

1	2	Word		3	
Left neighbors	**No**	**Occur-rences**	**Word**	**Right neighbors**	
answered, fortune, him, 2-house, 2-in, no, said, that	12	1	#	little, no, please, presently, so, 3-the, then, which	
#	1	2	there	was	
there	1	3	was	an	
was	1	4	an	old	
an	1	5	old	sow	
old	1	6	sow	with	
man, sow	2	7	with	a, three	
with	1	8	three	little	
#, 2-the, three	4	9	little	3-pig, pigs	
little	1	10	pigs	and	
did, door, in, pigs, puff, puffed, straw, wolf	8	11	and	I, as, ate, he, knocked, 2-said, the	
and	1	12	as	she	
as, them	2	13	she	had, sent	
she	1	14	had	not	

had	1	15	not	enough	
not	1	16	enough	to	
answered, enough, out, said, straw	5	17	to	build, him, keep, seek, that	
to	1	18	keep	them	
keep, sent	2	19	them	out, she	
she	1	20	sent	them	
them	1	21	out	to	
to	1	22	seek	their	
seek	1	23	their	fortune	
their	1	24	fortune	#	
3-#, and, at, up, which	7	25	the	door, first, 2-little, man, pig, wolf	
the	1	26	first	that	
first, me, to	3	27	that	#, straw, went	
that	1	28	went	off	
went	1	29	off	met	
off	1	30	met	a	
along, build, built, met, with	5	31	a	bundle, 2-house, man, wolf	
a, please, the	3	32	man	did, give, with	
a	1	33	bundle	of	
bundle	1	34	of	straw	
of, that	2	35	straw	and, to	
2-and	2	36	said	#, to	
to	1	37	him	#	
#	1	38	please	man	
man	1	39	give	me	
give, let	2	40	me	come, that	
to	1	41	build	a	
2-a, his, your	4	42	house	2-#, 2-in	
#	1	43	which	the	
man	1	44	did	and	
3-little, the	4	45	pig	#, answered, built, let	
pig	1	46	built	a	
#	1	47	presently	came	
presently	1	48	came	along	
came	1	49	along	a	
a, the	2	50	wolf	and, then	
and	1	51	knocked	at	
knocked	1	52	at	the	
the	1	53	door	and	
pig	1	54	let	me	
me	1	55	come	in	
come, 2-house	3	56	in	2-#, and	
pig, then	2	57	answered	#, to	

#	1	**58**	no	#
#, wolf	2	**59**	then	I, answered
and, then	2	**60**	I	2-ll
2-I	2	**61**	ll	blow, puff
ll	1	**62**	puff	and
ll	1	**63**	blow	your
blow	1	**64**	your	house
#	1	**65**	so	he
and, so	2	**66**	he	blew, puffed
he	1	**67**	puffed	and
he	1	**68**	blew	his
blew	1	**69**	his	house
and	1	**70**	ate	up
ate	1	**71**	up	the

APPENDIX 2

Table 2 : **Vocabulary and neighborhoods of *The Three Little Pigs.***

The complete text.

1	**2**	**Word**		**3**	
Left neighbors	**No**	**Occur-rences**	**Word**	**Right neighbors**	
2-afraid, afternoon, again, 2-along, angry, 2-answered, apple, 2-apples, bricks, came, chimney, 2-churn, 2-coming, 2-dinner, do, 2-down, fair, far, field, fire, furze, garden, go, happened, hide, hill, him, 2-home, 3-house, 6-in, 2-it, late, 3-no, oclock, other, out, 2-pig, 2-pigs, 2-ready, replied, 15-said, 2-straw, 2-them, three, time, tomorrow, tree, turnips, up, well, with, wolf, 2-yes, 2-you	96	1	#	9-I, a, at, but, 11-he, 2-if, in, 2-man, no, 8-pig, please, she, 45-the, 2-then, very, 2-we, 2-what, 2-where, will	
#, been, pig, 2-where	5	2	there	and, 3-is, was	
and, he, there, 3-wolf	6	3	was	2-afraid, an, gone, late, very	
down, was	2	4	an	apple, old	
an	1	5	old	sow	
old, the	2	6	sow	sent, with	
hill, 3-house, 3-man, sow	8	7	with	#, 3-a, it, the, them, three	

at, with	2	8	three	#, little
first, other, three	3	9	little	pig, 2-pigs
2-little	2	10	pigs	2-#
#	1	11	she	had
I, he, she	3	12	had	been, not, to
3-could, had, will	5	13	not	2-blow, deceive, enough, tell
not	1	14	enough	food
enough	1	15	food	for
call, come, food, off, turnips	5	16	for	dinner, the, them, 2-you
for, sent, with	3	17	them	2-#, out
45-#, 5-?, and, are, 2-at, 3-before, blew, 2-blow, climb, day, 2-down, for, 2-gave, got, him, in, 4-into, 2-saw, threw, 3-to, told, 3-up, with	85	18	the	apple, 2-apples, bricks, chimney, 3-churn, door, 2-fair, field, fire, first, furze, garden, 2-hill, 3-house, 3-man, next, other, 28-pig, second, sow, straw, third, time, tree, turnips, 24-wolf
sow	1	19	sent	them
them	1	20	out	#
the	1	21	first	little
8-#, little, second, 28-the, third	39	22	pig	2-#, 2-I, answered, 2-are, 3-built, 3-got, in, jumped, 3-let, made, 3-met, 11-said, saw, somehow, there, threw, told, went
3-pig	3	23	met	3-a
#, 2-bought, 3-build, built, got, 3-is, 3-met, up, 3-with	18	24 (4)	a	2-bundle, 2-churn, fair, fire, 4-house, load, 3-man, 3-nice, wolf
2-#, 3-a, please, 3-the	9	25	man	3-gave, 3-give, 3-with
2-a	2	26	bundle	2-of
2-bundle, field, load	4	27	of	bricks, furze, straw, turnips
of, that, the	3	28	straw	2-#, to
2-and, 3-he, 11-pig, 4-wolf	20	29	said	15-#, 2-no, to, 2-yes
been, bricks, churn, did, furze, gone, got, had, hoped, said, straw, want, wanted, what	14	30	to	3-build, climb, do, 2-get, go, hide, him, pick, 3-the
gave, to	2	31	him	#, the
3-man	3	32	give	3-me
deceive, 3-give, 3-let	7	33	me	I, 3-come, 2-that, those
found, 2-me, thought	4	34	that	furze, 2-he, straw
3-to	3	35	build	3-a
4-a, 3-his, 3-the, 3-your	13	36	house	3-#, 3-down, 4-in, 3-with
3-man	3	37	gave	him, 2-the
3-pig	3	38	built	a, 2-his
in, pick, with	3	39	it	2-#, up
a, 24-the	25	40	wolf	#, answered, 6-came, 2-coming, fell, felt, 2-huffed, knocked, puffed, replied, 4-said, 3-was, what

6-wolf	6	41	came	#, again, 2-along, as, up	
2-came	2	42	along	2-#	
wolf	1	43	knocked	at	
#, 2-?, knocked, 2-up, you	7	44	at	2-five, four, six, 2-the, three	
the	1	45	door	and	
again, bricks, churn, door, 2-down, fair, five, furze, 2-huff, 2-huffed, in, oclock, 2-puff, puffed, there, together, tree, up	22	46	and	4-I, ate, blew, bought, come, get, got, 4-he, puffed, ran, rolled, 2-said, the, was, went	
3-pig	3	47	let	3-me	
and, 3-me, will	5	48	come	back, for, 3-in	
#, 3-come, 4-house, pig	9	49	in	6-#, and, it, the	
pig, wolf	2	50	answered	2-#	
#, 2-said	3	51	no	3-#	
9-#, 4-and, me, morning, 2-pig, 2-then	19	52	I	am, got, had, have, 2-know, saw, 12-will	
#, 12-I, 2-we, 2-you	17	53	will	2-be, 3-blow, call, come, get, 2-go, 2-huff, not, 2-puff, throw, you	
2-will	2	54	puff	2-and	
2-not, 3-will	5	55	blow	2-the, 3-your	
3-blow	3	56	your	3-house	
and, wolf	2	57	puffed	and, but	
and, he	2	58	blew	his, the	
blew, 2-built	3	59	his	3-house	
and, he	2	60	ate	2-up	
2-ate, came, eat, 2-got, it, made	8	61	up	#, a, and, 2-at, 3-the	
the	1	62	second	pig	
of, that, the	3	63	furze	#, and, to	
2-#	2	64	then	2-I	
2-will	2	65	huff	2-and	
2-wolf	2	66	huffed	2-and	
11-#, 4-and, as, 2-but, 2-that	20	67	he	ate, blew, bought, 3-could, did, found, 2-got, had, hoped, ran, 3-said, saw, wanted, was, would	
get, 3-house, jumped, 2-rolled, you	8	68	down	2-#, an, 2-and, into, 2-the	
the	1	69	third	pig	
a	1	70	load	of	
of, the, those	3	71	bricks	#, and, to	
#	1	72	please	man	
me	1	73	those	bricks	
came	1	74	as	he	
he	1	75	did	to	
or, the	2	76	other	#, little	
#, angry, puffed	3	77	but	2-he, thought	
3-he	3	78	could	3-not	
he	1	79	found	that	
2-I	2	80	know	2-where	
2-#, 2-know	4	81	where	2-?, 2-there	
3-there	3	82	is	3-a	
3-a, apples	4	83	nice	?, apple, dinner, field	
nice, the	2	84	field	#, of	

of, some, the	3	85	turnips	#, before, for	
2-go, nice, 2-ready, 2-where	7	86	?	2-at, 5-the	
2-#	2	87	if	2-you	
are, do, 2-for, 2-if, saw, shall, throw, will	10	88	you	2-#, at, be, down, go, ready, want, 2-will	
2-will, you	3	89	be	3-ready	
am, 3-be, you	5	90	ready	2-#, 2-?, tomorrow	
oclock, ready	2	91	tomorrow	#, morning	
tomorrow	1	92	morning	I	
will	1	93	call	for	
2-#	2	94	we	2-will	
to, 2-will, you	4	95	go	#, 2-?, together	
go	1	96	together	and	
and, 2-to, will	4	97	get	back, down, 2-some	
2-get	2	98	some	apples, turnips	
for, nice	2	99	dinner	2-#	
#, felt, was	3	100	very	2-angry, well	
very	1	101	well	#	
2-#, tell, wolf	4	102	what	happened, 2-time, to	
the, 2-what	3	103	time	#, do, shall	
time, to	2	104	do	#, you	
you	1	105	want	to	
at	1	106	six	oclock	
five, four, six	3	107	oclock	#, and, tomorrow	
I, and, 2-he, 3-pig	7	108	got	a, 2-into, the, to, 2-up	
2-at	2	109	five	and, oclock	
back, off, turnips	3	110	before	3-the	
2-pig	2	111	are	the, you	
I	1	112	am	ready	
I	1	113	have	been	
had, have	2	114	been	there, to	
come, get	2	115	back	again, before	
back, came	2	116	again	#, and	
wolf	1	117	felt	very	
2-very	2	118	angry	#, but	
but	1	119	thought	that	
he	1	120	would	eat	
would	1	121	eat	up	
pig	1	122	somehow	or	
somehow	1	123	or	other	
an, nice, the	3	124	apple	#, far, tree	
apple, the	2	125	tree	#, and	
wolf	1	126	replied	#	
the	1	127	garden	#	
not	1	128	deceive	me	
some, 2-the	3	129	apples	2-#, nice	
at	1	130	four	oclock	
and, pig	2	131	went	2-off	
2-went	2	132	off	before, for	
he	1	133	hoped	to	
to	1	134	climb	the	
was	1	135	late	#	
I, he, pig	3	136	saw	2-the, you	
2-wolf	2	137	coming	2-#	

2-was	2	138	afraid	2-#	
2-said	2	139	yes	2-#	
will	1	140	throw	you	
pig	1	141	threw	the	
apple	1	142	far	#	
was	1	143	gone	to	
to	1	144	pick	it	
pig	1	145	jumped	down	
and, he	2	146	ran	2-home	
2-ran	2	147	home	2-#	
the	1	148	next	day	
next	1	149	day	the	
a, 2-the	3	150	fair	#, and, this	
fair	1	151	this	afternoon	
this	1	152	afternoon	#	
time	1	153	shall	you	
and, he	2	154	bought	2-a	
2-a, 3-the	5	155	churn	2-#, and, rolled, to	
not	1	156	tell	what	
down, fell, 2-got	4	157	into	4-the	
to	1	158	hide	#	
and, churn	2	159	rolled	2-down	
2-the	2	160	hill	#, with	
pig	1	161	told	the	
what	1	162	happened	#	
he	1	163	wanted	to	
the	1	164	chimney	#	
pig	1	165	made	up	
a, the	2	166	fire	2-#	
wolf	1	167	fell	into	

The full **3LP** story goes as follows:

P=char('#','there', 'was', 'an', 'old', 'sow', 'with', 'three', 'little', 'pigs', '#',...
 'she', 'had', 'not', 'enough', 'food', 'for', 'them', '# ', 'the', 'sow', 'sent', 'them',...
 'out', '#', 'the', 'first', 'little', 'pig', 'met', 'a', 'man', 'with', 'a', 'bundle', 'of',...
 'straw', '#', 'the', 'pig', 'said', 'to', 'him', '# ', 'man', 'give', 'me', 'that', 'straw',...
 'to', 'build', 'a', 'house', '#', 'the', 'man', 'gave', 'the', 'straw', '#', 'the', 'pig',...
 'built', 'a', 'house', 'with', 'it', '#', 'a', 'wolf', 'came', 'along', '#', 'the', 'wolf',...
 'knocked', 'at', 'the', 'door', 'and', 'said', '#', 'pig', 'let', 'me', 'come', 'in', '#',...
 'the', 'pig', 'answered', '#', 'no', '#', 'the', 'wolf', 'said', '#', 'I', 'will', 'puff', 'and'...
 , 'I', 'will', 'blow', 'your', 'house', 'in', '#', 'the', 'wolf', 'puffed', 'and', 'blew',...
 'his', 'house', 'in', 'and', 'ate', 'up', 'the', 'pig', '#', 'the', 'second', 'pig',...
 'met', 'a', 'man', 'with', 'a', 'bundle', 'of', 'furze', 'and', 'said', '#', 'man',...
 'give', 'me', 'that', 'furze', 'to', 'build', 'a', 'house', '#', 'the', 'man', 'gave',...
 'the', 'furze', '#', 'the', 'pig', 'built', 'his', 'house', '#', 'the', 'wolf', 'came',...
 'along', '#', 'the', 'wolf', 'said', '#', 'pig', 'let', 'me', 'come', 'in', '#', 'the', 'pig'...
 , 'said', 'no', '#', 'the', 'wolf', 'answered', '#', 'then', 'I', 'will', 'huff', 'and', 'I',...
 'will', 'blow', 'your', 'house', 'in', '#', 'the', 'wolf', 'huffed', 'and', 'he', 'blew',...
 'the', 'house', 'down', 'and', 'he', 'ate', 'up', 'the', 'pig', '#', 'the', 'third', 'pig',...

'met', 'a', 'man', 'with', 'a', 'load', 'of', 'bricks', '#', 'the', 'pig', 'said', '#',...
'please', 'man', 'give', 'me', 'those', 'bricks', 'to', 'build', 'a', 'house', 'with',...
'#', 'the', 'man', 'gave', 'him', 'the', 'bricks', 'and', 'the', 'pig', 'built', 'his',...
'house', 'with', 'them', '#', 'the', 'wolf', 'came', 'as', 'he', 'did', 'to', 'the',...
'other', 'little', 'pigs', '#', 'he', 'said', '#', 'pig', 'let', 'me', 'come', 'in', '#', 'the',...
'pig', 'said', 'no', '#', 'then', 'I', 'will', 'huff', 'and', 'I', 'will', 'puff', 'and', 'I', 'will',...
'blow', 'your', 'house', 'in', '#', 'the', 'wolf', 'huffed', 'and', 'puffed', 'but', 'he',...
'could', 'not', 'blow', 'the', 'house', 'down', '#', 'he', 'found', 'that', 'he', 'could',...
'not', 'blow', 'the', 'house', 'down', '#', 'the', 'wolf', 'said', '#', 'pig', 'I', 'know',...
'where', 'there', 'is', 'a', 'nice', 'field', 'of', 'turnips', '#', 'where', '?', 'the', 'pig', ...
'said', '#', 'in', 'the', 'field', '#', 'if', 'you', 'will', 'be', 'ready', 'tomorrow',...
'morning', 'I', 'will', 'call', 'for', 'you', '#', 'we', 'will', 'go', 'together', 'and', ...
'get', 'some', 'turnips', 'for', 'dinner', '#', 'the', 'pig', 'said', '#', 'very', 'well', '#',...
'I', 'will', 'be', 'ready', '#', 'what', 'time', 'do', 'you', 'want', 'to', 'go', '?', 'at', 'six',...
'oclock', '#', 'the', 'pig', 'got', 'up', 'at', 'five', 'and', 'he', 'got', 'the', 'turnips',...
'before', 'the', 'wolf', 'came', '#', 'the', 'wolf', 'said', '#', 'pig', 'are', 'you', 'ready',...
'?', 'the', 'pig', 'said', '#', 'I', 'am', 'ready', '#', 'I', 'have', 'been', 'there', 'and',...
'come', 'back', 'again', 'and', 'got', 'a', 'nice', 'dinner', '#', 'the', 'wolf', 'felt',...
'very', 'angry', 'but', 'thought', 'that', 'he', 'would', 'eat', 'up', 'the', 'pig',...
'somehow', 'or', 'other', '#', 'he', 'said', '#', 'pig', 'I', 'know', 'where', 'there', 'is',...
'a', 'nice', 'apple', 'tree', '#', 'the', 'pig', 'said', '#', 'where', '?', 'the', 'wolf',...
'replied', '#', 'at', 'the', 'garden', '#', 'if', 'you', 'will', 'not', 'deceive', 'me', 'I', ...
'will', 'come', 'for', 'you', 'at', 'five', 'oclock', 'tomorrow', '#', 'we', 'will', 'get',...
'some', 'apples', '#', 'the', 'pig', 'got', 'up', 'at', 'four', 'oclock', 'and', 'went',...
'off', 'for', 'the', 'apples', '#', 'he', 'hoped', 'to', 'get', 'back', 'before', 'the', ...
'wolf', '#', 'but', 'he', 'had', 'to', 'climb', 'the', 'tree', 'and', 'was', 'late', '#', 'he',...
'saw', 'the', 'wolf', 'coming', '#', 'he', 'was', 'afraid', '#', 'the', 'wolf', 'came', ...
'up', 'and', 'he', 'said', '#', 'pig', 'are', 'the', 'apples', 'nice', '?', 'the', 'pig',...
'said', 'yes', '#', 'I', 'will', 'throw', 'you', 'down', 'an', 'apple', '#', 'the', 'pig', ...
'threw', 'the', 'apple', 'far', '#', 'the', 'wolf', 'was', 'gone', 'to', 'pick', 'it', 'up', ...
'#', 'the', 'pig', 'jumped', 'down', 'and', 'ran', 'home', '#', 'the', 'next', 'day', ...
'the', 'wolf', 'came', 'again', '#', 'pig', 'there', 'is', 'a', 'fair', 'this', ...
'afternoon', '#', 'will', 'you', 'go', '?', 'the', 'pig', 'said', 'yes', '#', 'I', 'will', ...
'go', '#', 'what', 'time', 'shall', 'you', 'be', 'ready', '?', 'at', 'three', '#', 'the',...
'pig', 'went', 'off', 'before', 'the', 'time', '#', 'he', 'got', 'to', 'the', 'fair', '#',...
'he', 'bought', 'a', 'churn', '#', 'the', 'pig', 'saw', 'the', 'wolf', 'coming', '#',...
'he', 'could', 'not', 'tell', 'what', 'to', 'do', '#', 'the', 'pig', 'got', 'into', 'the',...
'churn', 'to', 'hide', '#', 'the', 'churn', 'rolled', 'down', 'the', 'hill', 'with', 'the'...
, 'pig', 'in', 'it', '#', 'the', 'wolf', 'was', 'afraid', '#', 'he', 'ran', 'home', '#', 'the'...
, 'pig', 'told', 'the', 'wolf', 'what', 'happened', '#', 'the', 'pig', 'said', '#', 'I',...
'had', 'been', 'to', 'the', 'fair', 'and', 'bought', 'a', 'churn', '#', 'I', 'saw', 'you', ...
'#', 'I', 'got', 'into', 'the', 'churn', 'and', 'rolled', 'down', 'the', 'hill', '#', 'the', ...
'wolf', 'was', 'very', 'angry', '#', 'he', 'wanted', 'to', 'get', 'down', 'into', 'the',...
'chimney', '#', 'the', 'pig', 'made', 'up', 'a', 'fire', '#', 'the', 'wolf', 'fell', 'into', ...
'the', 'fire', '#');

APPENDIX 3

Table 3. Vocabulary and word neighborhoods of Hungarian folk tale *A só* (*Salt*). Fragments

adni, arca, búzát, egyforma, egymagában, engem, is, király, királykisasszony, királykisasszonyt, kérdezlek, középsõt, legidõsebbiket, legkisebbikhez, leány, leánya, leányom, menjen, ország, szellõt, szereti, 2-szeretsz, 2-sót, te	28	1	#	az, azt, egyszer, elindult, ez, felelj, felelte, fordult, förmedt, hanem, hiába, hát, ki, 2-kérdezte, megtetszett, mint, mit, mondjad, mondta, ne, no, szépen, világgá, én, úgy
#, országa, piszkos, szép	4	2	volt	a, egyszer, különösen, mind
#, mondta, volt	3	3	egyszer	a, egy, mikor
beért, egyszer, már, még, élt	5	4	egy	darabig, esztendõ, nagy, órát, öreg
az, egy	2	5	öreg	2-király
2-a, 2-öreg	4	6	király	#, a, s, szerette
király, királyfi, palotájába, sírva	4	7	s	beért, három, két, meglátta
3-a, mert, nincs, s	6	8	három	egyforma, leányára, leányát, ország, országa, szép
három, ruhája	2	9	szép	leánya, volt
szép	1	10	leánya	#
#, ahogy, hogy, ki, különösen, még	6	11	az	arca, 2-emberek, országomból, udvaromból, öreg
király	1	12	szerette	volna
lett, szerette	2	13	volna	mind, nehéz
volna, volt	2	14	mind	2-a
adja, egyszer, elindult, emberek, felelte, fordult, galamb, hazavezette, király, királyfinak, 2-kérdezte, megfogta, meglátta, megtetszett, 2-mind, mint, mondta, nyárban, rá, szeretik, sírt, volt, úgy	25	15	a	galamb, 3-három, kezét, 2-kicsi, 2-király, királyfinak, 2-királykisasszony, királykisasszonyt, középsõt, legidõsebbiket, legkisebbikhez, legszebb, leány, leányainak, palotájába, ruhája, szellõt, 2-sót, tiszta
három	1	16	leányát	férjhez
leányát	1	17	férjhez	adni
férjhez	1	18	adni	#
#	1	19	ez	nem

……………………………………

2-#	2	55	2-a	kérdezte
a	1	56	#	legidõsebbiket
#, édesapám	2	57	a, forró	mint
a	1	58	a	galamb
a	1	59	búzát	tiszta
tiszta	1	60	#	búzát

………………………………

egyet, hetet, órát	3	140	sem	de, várt, és
sem	1	141	várt	de
még	1	142	egyet	sem
de	1	143	talán	még
egy	1	144	órát	sem
sem	1	145	és	megesküdtek
és	1	146	megesküdtek	#

Notes. 1. **Királykisasszony** (princess) means literally **Kinglittlewoman(lady).** Could it be that adjectives were initially just components of nouns?

2. A legidősebbiket : a leg-idő-sebb-ik-et   "the eldest one," in Object Case. This gives the taste of an agglutinative language.

3. Only the definite article a/az shows a sharp frequency jump because other auxiliary morphemes, even the possessive pronouns, are added to the root and serve as easy classifiers. This illustrates once again that the space between the words is to a significant degree arbitrary. A Chinese sentence can be seen as a single word.

4. It could be more productive to represent highly inflected languages not in terms of words but in terms of syllables, which would bridge phonology and morphology. This work is in process.

Fragments of the input:

P = char ('#', 'volt', 'egyszer', 'egy', 'öreg', 'király', 's', 'három', 'szép', ...

'leánya', '#', 'az', 'öreg', 'király', 'szerette', 'volna' , 'mind', 'a', ...
'három', 'leányát', 'férjhez', 'adni', '#', 'ez', 'nem', 'is', 'lett', ...
 'volna', 'nehéz', 'mert', 'három', 'országa', 'volt', 'mind', ...
'a', 'három', 'leányára', 'jutott', 'egy-egy', 'ország', '#', 'hanem', ...
'ahogyan', 'nincs', 'három', 'egyforma', 'alma', 'úgy', 'a',...

. . . .

 'darabig', 'egymagában', '#', 'egyszer', 'mikor', 'már', 'egy', ...
'esztendõ', 'is', 'eltelt', 'arra', 'járta', 'szomszéd', 'királyfi', 's', 'meglátta',...
 'a', 'királykisasszonyt', '#', 'megtetszett', 'a', 'királyfinak', 'a', ...
'királykisasszony', 'mert', 'akármilyen', 'piszkos', 'volt', 'a', ...
'ruhája', 'szép', 'volt', 'különösen', 'az', 'arca', '#', 'szépen', 'megfogta', ...
 'a', 'kezét', 'hazavezette', 'a', 'palotájába', 's', 'két', 'hetet', 'sem', 'várt', ...
 'de', 'még', 'egyet', 'sem', 'de', 'talán', 'még', 'egy', 'órát', 'sem', 'és', 'megesküdtek', '#');

Translation of the Hungarian tale (fragments):

Salt

Once upon a time there lived an old king with his three beautiful daughters. The old king wanted all three his daughters to get married. It would not be difficult because he had three lands, one for each daughter.

But as there are no three apples alike, so . . . .

. . . . completely on her own. Once , when already a year had passed, a neighboring prince came there and saw the little princess. The prince was struck by the princess because however shabby her dress was, she was beautiful, especially, her face. He gently took her hand and brought her to his palace, and not in two weeks, not even in one, maybe not even in an hour they got married.
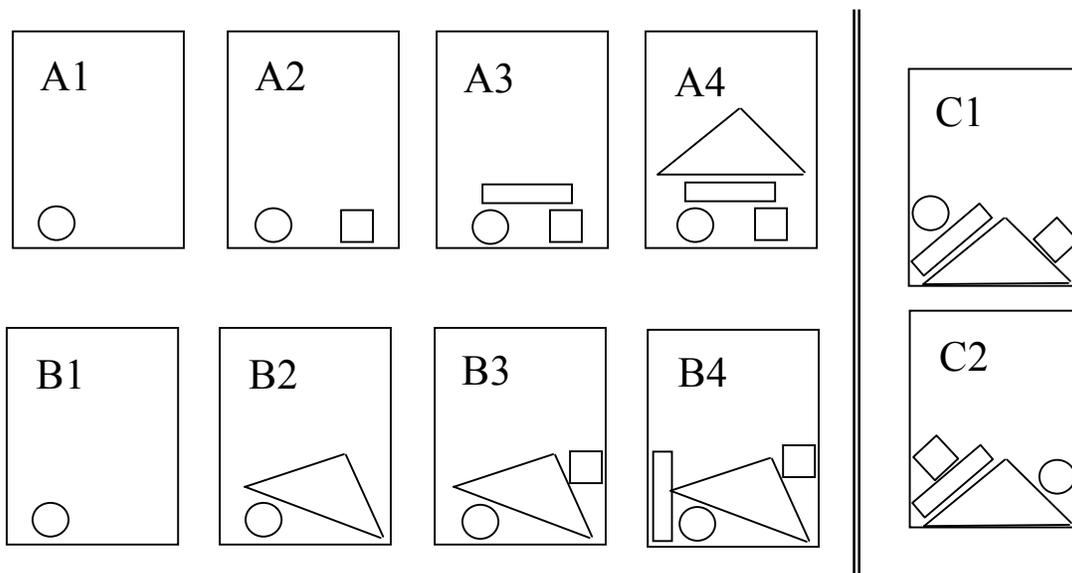
. . . .

APPENDIX 4 . Simplistic models

A. Simplistic model of memory

How can we count the frequency of inputs, for example, water droplets, without a computer and a data record?
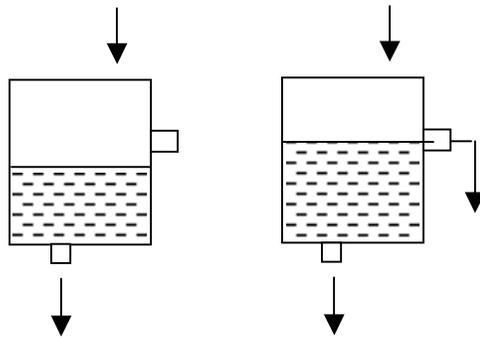
B. Simplistic model of acquisition/growth/evolution

Different histories from the same starting point **A1 = B1** mean different results (**A4** and **B4**) , although history is not remembered. It could be partially reconstructed from the properties of components and their interactions. History is stored in life and society.

If the temperature (imitated by shaking the box) is not too high and not too cold, one or very few stable states (**C1** and **C2**) are possible in equilibrium.

APPENDIX 5

Program OpenMind (OM)

I would like to formulate the following imaginary program of research.

Purpose:

    1. Consensus in understanding mind.

    2. Growing an artificial natural mind as the proof of consensus.

Components:

1. Mind as it is: psychology and neurophysiology.

2. Mind as it was: origin from chemistry and life.

3. Mind as what it creates: history, anthropology, linguistics, sociology, science, art,
    economics, colossal errors of judgment, etc.

4. Mind as it is created.

    4.1. Artificial mind (AI).

    4.2. Artificial natural mind.

5. Method:

    5.1. Completely open online source and exchange.

    5.2. Start with the simplest possible prototype.

    5.3. Develop measure of simplicity/complexity.

    5.3. Add components and increase complexity without any *a priori* general plan.

    5.4. Use the Turing Test, first adapted for an infant mind, then for adult one.

I believe that Ulf Grenander's GOLEM [3] can be a starting point in the project because of its meta-chemical properties and the perfect potential for starting small.