



---

## Joint Acquisition Of Language And Cognition

**Corresponding Author:**

Dr. Leonid Perlovsky,  
Dr., Harvard University - United States of America

**Submitting Author:**

Dr. Leonid Perlovsky,  
Doctor, Visiting Scholar, Principal Research Physicist and Technical Advisor at the Air Force Research Lab ,  
Harvard University, New York - United States of America

**Article ID:** WMC00994

**Article Type:** Research articles

**Submitted on:** 13-Oct-2010, 10:17:43 AM GMT **Published on:** 31-Oct-2010, 12:01:54 AM GMT

**Article URL:** [http://www.webmedcentral.com/article\\_view/994](http://www.webmedcentral.com/article_view/994)

**Subject Categories:** BRAIN

**Keywords:** language, cognition, thinking, concepts, emotions, knowledge instinct, dynamic logic, mind, hierarchy

**How to cite the article:** Perlovsky L . Joint Acquisition Of Language And Cognition . WebmedCentral BRAIN  
2010;1(10):WMC00994

# Joint Acquisition Of Language And Cognition

**Author(s):** Perlovsky L

## Abstract

What is the role of language and cognition in thinking? Is language just a communication device, or is it fundamental in developing thoughts? Chomsky suggested that language is separate from cognition. Cognitive linguistics emphasizes a single mechanism of both. Neither led to a computational theory. Here we develop a hypothesis that language and cognition are two separate but closely interconnected mechanisms; the role of each is identified. Language stores cultural wisdom; cognition develops mental representations modeling surrounding world and adapts cultural knowledge to concrete circumstances of life. Language is acquired from surrounding language 'ready-made' and therefore can be acquired early in life. This early acquisition of language by five years of age encompasses the entire hierarchy from sounds to words, to phrases, to highest concepts existing in culture. Cognition requires experience. The paper presents arguments why cognition can not be acquired directly from experience; language is a necessary intermediary, a "teacher." A mathematical model is developed that overcomes previous difficulties towards a computational theory. This model implies a specific neural mechanism consistent with Arbib's "language prewired brain;" it also models recent neuroimaging data about cognition, remaining unnoticed by other theories. The suggested theory explains a number of properties of language and cognition, which previously seemed mysterious.

## Main Text

### Introduction

Interaction between language and cognition remains an unsolved scientific problem. Chomsky's separation of these two abilities does not seem reasonable, and after 50 years of development it still has not resulted in a computational theory of language. Cognitive linguistics attempted to unify language and cognition; these attempts also have not been successful. Fundamental questions remain unanswered and seem mysterious. How every child learns a correspondence between words and objects among an almost uncountable number of wrong combinations? Combinations even among 100 words and objects exceed all the particle in the Universe, and no amount

of experience would suffice to learn these associations using existing methods. Why are there no animals with human thinking but without human language? Evolutionary linguistics emphasizes evolution as a mechanism of language acquisition, yet existing approaches also lead to combinatorial computational complexity. Here we propose a computational theory, corresponding to existing data on language and cognition, explaining each ability, and the fundamental role of both in thinking processes.

### Mathematical models in linguistics

Inborn brain-mind mechanisms were not appreciated in the first half of the last century. Logic dominated thinking of mathematicians and the intuitions of psychologists and linguists. Logical mechanisms are not much different for language or cognition; both were based on logical statements and rules. Deficiencies of logic established by the fundamental Gödelian results [[1]] did not move thinking about the mind away from logic.

Chomsky [[2]] initiated contemporary linguistic interests in the mind mechanisms of language in the 1950s. Among the first mysteries about language that science had to resolve he identified "poverty of stimulus." The fact that the tremendous amount of knowledge needed to speak and understand language is learned by every child around the world even in the absence of formal training. It seemed obvious to Chomsky that surrounding language cultures do not carry enough information for a child to learn language, unless specific language learning mechanisms are inborn. These mechanisms should be specific enough for learning complex language grammars and still flexible enough so that a child of any ethnicity from any part of the world would learn whichever language is spoken around. This inborn learning mechanism Chomsky called Universal Grammar and set out to discover its mechanisms. Chomsky emphasized the importance of syntax and thought that language learning is independent of cognition. The idea of inborn or innate language mechanisms is called nativism.

Followers of Chomsky initially used available mathematics of logical rules, similar to rule systems of artificial intelligence. A new mathematical paradigm in linguistics was proposed in [[3]], rules and parameters. This was similar to model-based systems emerging in mathematical studies of cognition. Universal properties of language grammars were supposed to be modeled by parametric rules or models, and specific

characteristics of grammar of a particular language were fixed by parameters, which every kid could learn when exposed to the surrounding language. Another fundamental change of Chomsky's ideas [[4]] was called the minimalist program. It aimed at simplifying the rule structure of the mind mechanism of language. Language was considered to be in closer interactions to other mind mechanisms, closer to the meaning, but stopped at an interface between language and meaning. Still nativism assumes that meanings appear independently from language. Logic is the mathematics of modeling language.

Not everybody agreed with the separation between language and cognition. In the 1970s cognitive linguistics emerged to unify language and cognition, and explain the creation of meanings. Cognitive linguistics rejected Chomsky's idea about a special module in the mind devoted to language. Language is no different from the rest of cognition. It is embodied and situated in the environment. Related research on construction grammar argues that language is not compositional, not all phrases are constructed from words using the same syntax rules and maintaining the same meanings; metaphors are good examples [[5],[6],[7]]. Neither nativism, nor cognitive linguistics lead to a computational linguistic theory explaining how cognition and language is acquired, and meanings are created. Again, logic is used for mathematical modeling.

Importance of evolving language and meanings are emphasized by evolutionary linguistics. Language mechanisms are shaped by transferring from generation to generation [[8],[9]]. [7,8,9]. This transferring process was demonstrated to be a "bottleneck," a process-mechanism that selected or "formed" compositional properties of language and connected language to meanings [[10],[11],[12]]. The evolutionary linguistic approach demonstrated mathematically that indeed this bottleneck leads to a compositional property of language. Under certain conditions a small number of sounds (phonemes, letters) are aggregated into a large number of words. Evolutionary linguistics by simulation of societies of communicating agents [[13]] demonstrated the emergence of a compositional language, still existing mathematical formalisms face combinatorial complexity.

Existing theories of language and cognition cannot address many aspects of these mechanisms, which remain mysterious. This paper addresses the role of language in cognition. The proposed model resolves some long-standing language-cognition issues: how the mind learns correct associations between words and objects among an astronomical number of

possible associations; why kids can talk about almost everything, but cannot act like adults, what exactly are the brain-mind differences? Why animals do not talk and think like people. How language and cognition participate in thinking. Recent brain imaging experiments indicate support for the proposed model. We discuss future theoretical and experimental research.

Mathematical model of cognition

Bottom-up and top-down signals

A simple experiment reveals important properties of perception and cognition, ignored by most theories [[14]]. Imagine an object in front of you with closed eyes. Imagination is vague, not as crisp and clear as with opened eyes. When eyes are opened, the object becomes crisp and clear. It seems to occur momentarily, but actually it takes about 1/5th of a second. This is a very long time for neural brain mechanisms – hundreds of thousands of neural interactions. Let us also note: with opened eyes we are not conscious about the initially vague imagination, we are not conscious of the entire 1/5th of a second, we are conscious only about the end of this process: a crisp, clear object in front of our eyes. This experiment has become easy to explain after many years of research that have uncovered what goes on in the brain during these 1/5th of a second.

Mechanisms of instincts, emotions, and mental representations are required to explain this experiment. Perception and understanding of the world is due to mechanism of mental representations, or concepts. Concept-representations are like mental models of objects and situations; this analogy is quite literal, e.g., during visual perception, a mental model of the object stored in memory projects an image (top-down signals) onto the visual cortex, which is matched there to an image projected from retina (bottom-up signal; for more details see [[15]]).

Mental representations is an evolutionary recent mechanism. It evolved for satisfaction of more ancient mechanisms of instincts. "Instinct" is not used currently in psychological literature; the reason is that the notion of instinct was mixed up with instinctual behaviors, which was not helpful. Here "instinct" is a simple inborn, non-adaptive mechanism described in [[16]]. Instinct is a mechanism similar to the internal "sensor," which measures vital body parameters, such as blood pressure, and indicate to the brain when these parameters are out of safe range. (More details could be found in [[17]] and references therein). An organism have dozens of such sensors, measuring sugar level in blood, body temperature, pressure at various parts, etc.

Instinctual-emotional theory of Grossberg-Levine [13]

suggests that communicating satisfaction or dissatisfaction of instinctual needs from instinctual parts of the brain to decision making parts of the brain is performed by emotional neural signals. Emotion refers to several neural mechanisms in the brain [[18]]; here it always refers to the mechanism connecting conceptual and instinctual brain regions. Perception and understanding of the mental models, corresponding to objects or situations that can potentially satisfy instinctual needs, receive preferential attention and processing resources in the mind.

Top-down neural signals projected from a mental model to the visual cortex makes visual neurons to be more receptive to matching bottom-up signals, or 'primes' neurons. This projection produces the imagination that we perceive with closed eyes, as in the close-open eye experiment. Conscious perception occurs, as mentioned, after top-down and bottom-up signals match. For a while the process of matching presented difficulties to mathematical modeling, as discussed below.

Combinatorial computational complexity and logic  
Computers cannot compete with animals [[19]]. Mathematical models of perception and cognition during the last 60 years faced the difficulty of combinatorial complexity (CC) [[20],[21],[22],[23]]. Learning requires training, algorithms have to be shown objects in their multiple variabilities; but also objects have to be shown surrounded by all objects that could appear around. These combinations lead to incomputable number of operations. Combinations of 100 objects are 100100, a number larger than all elementary particle interactions in the entire history of the Universe.

These CC difficulties have been related to Gödelian limitations of logic; they are manifestations of logic inconsistency in finite systems [[24],[25],[26]]. Approaches designed specifically to overcome logic limitations, such as fuzzy logic and neural networks, encountered logical steps in their operations: Training requires logical procedures (e.g. "this is a chair").

Dynamic logic (DL) was proposed to overcome limitations of logic [[27],16,15,[28],[29],[30]]. The mathematical description of DL is given later, here we describe it conceptually [[31]]. Classical logic is static (e.g. "this is a chair"), DL is a process 'from vague-to-crisp', from a vague representation, model, statement, decision, plan, to crisp ones. DL could be viewed as fuzzy logic that automatically sets a degree of fuzziness corresponding to the accuracy of the learned models.

DL models the open-close eye experiment: initial states of the models are vague. Recent brain imaging

experiments measured many details of this process. Bar et al [[32]] used functional Magnetic Resonance Imaging (fMRI) to obtain high-spatial resolution of processes in the brain, combined with magneto-encephalography (MEG), measurements of the magnetic field next to head, to provide a high temporal resolution of the brain activity. The experimenters were able to measure high resolution of cognitive processes in space and time. Bar et al concentrated on three brain areas: early visual cortex, object recognition area (fusiform gyrus), and object-information semantic processing area (OFC). They demonstrated that OFC is activated 130 ms after the visual cortex, but 50 ms before object recognition area. This suggests that OFC represents the cortical source of top-down facilitation in visual object recognition. This top-down facilitation was unconscious. In addition they demonstrated that the imagined perception generated by the top-down signals facilitated from OFC to the cortex is vague, similar to the close-open-eye experiment. Conscious perception of an object occurs when vague projections become crisp and match a crisp image from the retina; next, an object recognition area is activated.

Neural modeling field theory

The mind has an approximately hierarchical structure from sensory signals at the bottom to representations of the highest concepts at top [15,[33]]. Here we describe interaction between two adjacent layers in the hierarchy. We give a simplified description, as if eye retina. Matching mental models in memory to bottom-up signals coming from eyes is necessary for perception; otherwise an organism will not be able to perceive the surroundings and will not be able to survive. Therefore humans and high animals have an inborn drive to fit top-down and bottom-up signals, the instinct for knowledge [17,15].

The knowledge instinct is similar to other instincts in that the mind has a sensor-like mechanism, which measures a similarity between top-down and bottom-up signals, between concept-models and sensory signals, and maximizes this similarity. Brain areas participating in the knowledge instinct were discussed in [34]. That publication discussed similar mechanisms considered by biologists since the 1950s; without a mathematical formulation, however, its fundamental role in cognition was difficult to discern. All learning algorithms have some models of this instinct, maximizing correspondence between sensory input and an algorithm internal structure (knowledge in a wide sense). According to Grossberg and Levine instinct-emotion theory [13], satisfaction or dissatisfaction of every instinct is communicated to other brain areas by emotional neural signals.

Emotional signals associated with the knowledge instinct are felt as harmony or disharmony between our knowledge-models and the world [[35]]. At lower layers of the mind hierarchy, at the level of everyday object recognition, these emotions are usually below the level of consciousness; at higher layers of abstract and general concepts this feel of harmony or disharmony could be conscious; as discussed in [15],[36],[37],[38]] it is a foundation of our higher mental abilities; experimental demonstration of these emotions associated with knowledge is discussed in [[39]]. A mathematical theory combining the discussed mechanisms of cognition as interaction between top-down and bottom-up signals is summarized below following [15,[40]].

In a single layer of the mental hierarchy, neurons are enumerated by index  $n = 1, \dots, N$ . These neurons receive bottom-up input signals,  $X(n)$ , from lower layers in the mind hierarchy.  $X(n)$  is a field of bottom-up neuronal synapse activations, coming from neurons at a lower layer. Top-down, or priming signals to these neurons are sent by concept-models,  $M_m(S_m, n)$ ; we enumerate these models by index  $m = 1, \dots, M$ . Each model is characterized by its parameters,  $S_m$ . The models represent signals in the following sense. Say, signal  $X(n)$ , is coming from sensory neurons activated by object  $m$ , characterized by parameters  $S_m$ . These parameters may include position, orientation, or lighting of an object  $m$ . Model  $M_m(S_m, n)$  predicts a value  $X(n)$  of a signal at neuron  $n$ . For example, during visual perception, a neuron  $n$  in the visual cortex receives a signal  $X(n)$  from the retina and a priming signal  $M_m(S_m, n)$  from an object-concept-model  $m$ . A neuron  $n$  is activated if both a bottom-up signal from lower-layer-input and a top-down priming signal are strong. Various models compete for evidence in the bottom-up signals, while adapting their parameters for better match as described below. This is a simplified description of perception. Models  $M_m$  specify a field of primed neurons  $\{n\}$ , hence the name for this modeling architecture, modeling fields [18].

A mathematical model of the knowledge instinct is maximization of a similarity between top-down and bottom-up signals,

$$L = \prod \sum r(m) I(n|m). \quad (1)$$

$$n \in N \quad h \in H$$

Here  $I(n|m)$  is a conditional similarity between a bottom-up signal in pixel (sensor cell)  $n$  and top-down concept-representation  $m$ , given that signal  $n$  originated from concept-model  $m$ ; the functional shape of  $I(n|m)$  often can be taken as a Gaussian function of  $X(n)$  with the mean  $M_m(S_m, n)$ . Conditional similarities

are normalized on objects (or concepts)  $m$  being definitely present, and coefficients  $r(m)$  estimate a probability of objects actually being present. Similarity  $L$  accounts for all combinations of signals  $n$  coming from any model  $m$ , hence the huge number of items  $MN$  in eq. (1); this is a basic reason for combinatorial complexity of most algorithms. A system could form a new model; alternatively, old models are sometimes merged or eliminated. This requires a modification of the similarity measure (1); the reason is that more models always result in a better fit between the models and data. Therefore similarity (1) has to be multiplied by a "skeptical penalty function,"  $p(N, M)$  that grows with the number of parameters in models  $M$ , and the growth is steeper for smaller  $N$ .

The knowledge instinct maximizes similarity  $L$  over the model parameters  $S$ . DL is a mathematical technique maximizing similarity  $L$  without combinatorial complexity. Its salient property is matching vagueness or fuzziness of similarity measures to the uncertainty of the models. DL starts with any unknown values of parameters  $S$  and defines association variables  $f(m|n)$ ,  $f(m|n) = r(m)I(n|m) / \sum r(m')I(n|m')$ . (2)

$$m' \in M$$

DL determining the Neural Modeling Fields (NMF) dynamics is given by

$$dS_m/dt = \sum f(m|n) [\partial \ln I(n|m) / \partial M_m] \partial M_m / \partial S_m, \quad (3)$$

$$n \in N$$

When solving this equation iteratively,  $f(m|n)$  is recomputed according to (2) after each step using new parameter values, eq. (3). Parameter values are not known initially and uncertainty of conditional similarities (their variances) are high. So the fuzziness of the association variables is high. In the process of learning, models become more accurate, and association variables more crisp, the value of the similarity increases. The number of models is determined in the learning process. The system always keeps a store of dormant models, which are vague, have low  $r(m)$ , and do not participate in parameter fitting, except  $r(m)$ . When  $r(m)$  exceeds a threshold, a model is activated; correspondingly, an active model is deactivated when its  $r(m)$  falls below the threshold. In modeling interaction between bottom-up and top-down signals, the NMF-DL is similar to ART [[41]]; otherwise it is a very different architecture and algorithm. In particular, it uses parametric models, it fits multiple models in parallel, while associating bottom-up and top-down signals.

The process of DL always converges [18], it is proven by demonstrating that at each time step in eq. (3), the knowledge instinct (1) increases; thus DL and the

knowledge instinct are mathematically equivalent. Cultural effects of the knowledge instinct is discussed in [[42]].

#### Perception example

Below, DL is illustrated with an example described in more details in [[43],[44]], which demonstrates that DL can find patterns below the noise at about 100 times better than previous algorithms in terms of signal-to-noise ratio [15,[45]]. DL solves problems that were previously considered unsolvable and in many cases DL converges to the best possible solution of a problem [[46],[47],[48],[49],21,[50]].

Exact pattern shapes are not known and depend on unknown parameters, these parameters should be found by fitting the pattern model to the data. At the same time it is not clear which subset of the data points should be selected for fitting. A previous state-of-the-art algorithm for this type of problems, multiple hypotheses testing, tries various subsets [[51]]. In difficult cases, all combinations of subsets and models are exhaustively searched, leading to combinatorial complexity. In the current example we use simulated EEG signals of cognitively-related events; as usual, EEG signals are highly noisy, which makes difficult the problem of identifying patterns. The searched patterns are shown in Fig.1 at the bottom row. These events are "phase cones," circular events expanding or contracting in time (t, horizontal direction; in this case two expanding and one contracting event measured by an array of 64x64 sensors (each image chip). Direct search through all combinations of models and data leads to complexity of approximately  $MN = 1010,000$ , a prohibitive computational complexity.

The models and conditional similarities for this case are described in details in [29]: a uniform model for noise (not shown), expanding and contracting cones for the cognitive events. The number of computer operations in this example was about 1010. Thus, a problem that was not solvable due to CC becomes solvable using dynamic logic. DL in this example performs better than the human visual system. This is possible due to the fact that the human visual system is optimized for different types of images, not for circular shapes in noise.

Figure 1. Dynamic logic operation example, finding cognitively-related events in noise, in EEG signals. The searched patterns are shown in Fig.1 at the bottom row. These events are "phase cones," circular events expanding or contracting in time (horizontal direction t, each time step is 5 ms); in this case two expanding and one contracting event are simulated as measured by an array of 64x64 sensors. Direct search through all combinations of models and data leads to

complexity of approximately  $MN = 1010,000$ , a prohibitive computational complexity. The models and conditional similarities for this case are described in details in [29]: a uniform model for noise (not shown), expanding and contracting cones for the cognitive events. The first 5 rows illustrate dynamic logic convergence from a single vague blob at iteration 2 (row 1, top) to closely estimated cone events at iteration 200 (row 5); we did not attempt to reduce the number of iterations in this example; the number of computer operations was about 1010. Thus, a problem that was not solvable due to CC becomes solvable using dynamic logic.

#### Cognition example

Here we consider a next higher level in the hierarchy of cognition. At each level of the hierarchy bottom-up signals interact with top-down signals. For concreteness, we consider learning situations composed of objects. In real brain-mind, learning and recognition of situations proceeds in parallel with perception of objects. For simplifying presentation, we consider objects being already recognized. Situations are collections of objects. The fundamental difficulty of learning and recognizing situations is that when looking in any direction, a large number of objects is perceived. Some combinations of objects form "situations" important for learning and recognition, but most combinations of objects are just random sets, which human mind learns to ignore. The total number of combinations exceeds by far the number of objects in the Universe. This is the reason for this problem having not being solved over the decades [[52]].

This example is considered in details in [[53]]. Here we summarize the results. The data available for learning and recognition situations in this example are illustrated in Fig. 2. Horizontal axes corresponds to situations, the total number of situations are 16,000. Each situation is characterized by objects shown along the vertical axes. The total number of objects is 1000. Objects present in a situation are shown as white pixels, and absent objects are black. Fig. 2a illustrates data sorted by situations (horizontal axis). In every "important situation" there are several objects that are always present in this situation, hence the white lines in the left part of the figure. In half of situations there are no repeated objects; these random collections of objects are on the right of the Fig 2a. The same data are shown in Fig 2b with randomized order along horizontal line, as various situations actually appear in real life.

Fig. 2. Learning situations; white dots show present objects and black dots correspond to absent objects. Vertical axes show 1000 objects, horizontal axes show 10 situations each containing 10 relevant objects and

40 ransom one; in addition there 5000 “clutter” situations containing only random objects. Fig. 4a shows situations sorted along horizontal axis, hence there horizontal lines corresponding to relevant objects (right half contains only random noise). Fig 4b show the same situations in random order, which looks like random noise.

To solve this problem using a standard algorithm one can try to sort horizontal axis until white lines appear, similar to Fig 2a. This would take approximately 1040,000 operations, an unsolvable problem. Nevertheless, NMF-DL solves this problem in few iterations, as illustrated in Fig. 3.

Fig. 3. (a) shows DL initiation (random) and the first three iterations (Fig. 5a); the vertical axis shows objects and the horizontal axis shows models (from 1 to 20). The problem is approximately solved by the third iteration. This is illustrated in Fig 3b, where the error is shown on the vertical error. The correct situations are chosen by minimizing the error. The error does not go to 0 for numerical reasons as discussed in [31].

Fig. 3a illustrates DL iterations beginning with random association of objects and (arbitrary taken) 20 situations. Fig. 3b illustrate that errors quickly go to a small value. The error does not go to 0 for numerical reasons as discussed in [31]. In the above example relationships (such as on-the-left-of, or under) have not been explicitly considered. They can be easily included. Every relation and object can include a marker, pointing what relates to what. These markers are learned the same way as objects [31].

The procedure outlined in this section is general in that it is applicable to all higher layers in the mind hierarchy and to cognitive as well as language models. For example at higher layers, abstract concepts are subsets of lower level ones. The mathematical procedure outlined above is applicable without change.

#### Language learning

The procedure outlined in the previous section is applicable to learning language in the entire hierarchy from words up. Phrases are composed of words, and larger chunks of text from smaller chunks of texts can be learned similarly to learning above situations models composed of objects. Grammar rules, syntax, and morphology are learned using markers as discussed above. Lower layer models, may require continuous parametric models, like laryngeal models of phonemes [[54]]. These can be learned from language sounds using parametric models [[55],[56],[57],[58],[59],[60],[61],[65],[63],[64],[65],[66]] similar to a preceding section on perception.

The dual model of language and cognition

Do we use phrases to label situations that we already

have understood, or the other way around, do we just talk with words without understanding any cognitive meanings? It is obvious that different people have different cognitive and linguistic abilities and may tend to different poles in the cognitive-language continuum, while most people are somewhere in the middle in using cognition to help with language, and vice versa. What are the neural mechanisms that enable this flexibility? How do we learn which words and objects come together? If there is no specific language module, as assumed by cognitive linguists, why do kids learn a language by 5 or 7, but do not think like adults? And why there is no animals thinking like humans but without human language?

Little is known about neural mechanisms for integrating language and cognition. Here we propose a computational model that potentially can answer the above questions, and that is computationally tractable, it does not lead to combinatorial complexity. Also it implies relatively simple neural mechanisms, and explains why human language and human cognition are inextricably linked. It suggests that human language and cognition have evolved jointly.

#### Dual model

Whereas Chomskyan linguists could not explain how language and cognition interact, cognitive linguists could not explain why kids learn language by 5 but cannot think like adults; neither theory can overcome combinatorial complexity.

Consider first how is it possible to learn which words correspond to which objects? Contemporary psycholinguists follow ancient Locke idea, ‘associationsim’: associations between words and object are just remembered. But this is mathematically impossible. The number of combinations among 100 words and 100 objects is larger than all elementary particle interactions in the Universe. Combinations of 30,000 words and objects are practically infinite. No experience would be sufficient to learn associations. No mathematical theory of language offers any solution. NMF-DL solves this problem using the dual model [[67],[68],[69]]. Every mental representation consists of a pair of models, or two model aspects, cognitive and language. Mathematically, every concept-model  $M_m$  has two parts, linguistic  $ML_m$  and cognitive  $MC_m$ :

$$M_m = \{ ML_m, MC_m \}; \quad (4)$$

This dual-model equation suggests that the connection between language and cognitive models is inborn. In a newborn mind both types of models are vague placeholders for future cognitive and language contents. An image, say of a chair, and the sound “chair” do not exist in a newborn mind. But the neural

connections between the two types of models are inborn; therefore the brain does not have to learn associations between words and objects: which concrete word goes with which concrete object. Models acquire specific contents in the process of growing up and learning, linguistic and cognitive contents are always staying properly connected. Zillions of combinations need not be considered. Initial implementations of these ideas lead to encouraging results [[70],[71],[72],[73],[74],[75]].

#### Dual hierarchy

Consider language hierarchy higher up from words, Fig. 4. Phrases are made up from words similar to situations made up from objects. Because of linear structure, language actually is simpler than situations; rules of syntax can be learned similar to learning objects and relations using markers, as described in the previous section. The reason computers do not talk English used to be the fundamental problem of combinatorial complexity. Now, that the fundamental problem is solved, learning language will be solved in due course. Practically, significant effort will be required to build machines learning language. However, the principal difficulty has been solved in the previous section. Mathematical model of learning situations, considered in the previous section, is similar to learning how phrases are composed from words. Syntax can be learned similar to relations between objects [29,26,[76]].

Fig. 4. Parallel hierarchies of language and cognition consist of lower level concepts (like situations consist of objects). A set of objects (or lower level concepts) relevant to a situation (or higher level concept) should be learned among practically infinite number of possible random subsets (as discussed, larger than the Universe). No amount of experience would be sufficient for learning useful subsets from random ones. The previous section overcame combinatorial complexity of learning, given that the sufficient information is present. However, theories of mathematical linguistics offer no explanation where this information would come from.

The next step beyond current mathematical linguistics is modeling interaction between language and cognition. It is fundamental because cognition cannot be learned without language. Consider a widely-held belief that cognition can be learned from experience in the world. This belief is naïve and mathematically untenable. The reason is that abstract concepts-representations consist of a set of relevant bottom-up signals, which should be learned among practically infinite number of possible random subsets (as discussed larger than the Universe). No amount of experience would be sufficient for learning useful

subsets from random ones. The previous section overcame combinatorial complexity of learning, given that the sufficient information is present. However, mathematical linguistic theories offer no explanation where this information would come from.

NMF-DL with dual model and dual hierarchy suggests that information is coming from language. This is the reason why no animal without human-type language can achieve human-level cognition. This is the reason why humans learn language early in life, but learning cognition (making cognitive representations-models as crisp and conscious as language ones) takes a lifetime. Information for learning language is coming from the surrounding language at all levels of the hierarchy. Language model-representations exist in the surrounding language 'ready-made.' Learning language is thus grounded in the surrounding language.

For this reason language models become less vague and more specific by 5 years of age, much faster than the corresponding cognitive models for the reason that they are acquired ready-made from the surrounding language. This is especially true about the contents of abstract models, which cannot be directly perceived by the senses, such as "law," "abstractness," "rationality," etc. While language models are acquired ready-made from the surrounding language, cognitive models remain vague and gradually acquire more concrete contents throughout life guided by experience and language. According to the dual-model, this is an important aspect of the mechanism of what is colloquially called "acquiring experience."

Human learning of cognitive models continues through the lifetime and is guided by language models. If we imagine a familiar object with closed eyes, this imagination is not as clear and conscious as perception with opened eyes. With opened eyes it is virtually impossible to remember imaginations. Language plays a role of eyes for abstract thoughts. On one hand, abstract thoughts are only possible due to language, on the other, language "blinds" our mind to vagueness of abstract thoughts. Whenever one can talk about an abstract topic, he (or she) might think that the thought is clear and conscious in his (or her) mind. But the above discussion suggests that we are conscious about the language models of the dual hierarchy. The cognitive models in most cases may remain vague and unconscious. During conversation and thinking, the mind smoothly glides among language and cognitive models, using those that are crisper and more conscious – 'more available.' Scientists, engineers, and creative people in general are trained to differentiate between their own thoughts and what they read in a book or paper, but usually

people do not consciously notice if they use representations deeply thought through, acquired from personal experience, or what they have read or heard from teachers or peers. The higher up in the hierarchy the vaguer are the contents of abstract cognitive representations, while due to crispness of language models we may remain convinced that these are our own clear conscious thoughts.

Animal vocalizations are inseparable from instinctual needs and emotional functioning. The dual model has enabled separation of semantic and emotional contents, which made possible deliberate thinking. Yet operations of the dual model, connecting sounds and meanings, requires motivation. Motivation in language is carried by sounds [[77]]. Future research will have to address remaining emotionality of human languages, mechanisms involved, emotional differences among languages, and effects of language emotionalities on cultures.

Evolution of the language ability required rewiring of human brain. Animal brains cannot develop ability for deliberate discussions because conceptual representations, emotional evaluations, and behavior including vocalization are unified, undifferentiated states of the mind. Language required freeing vocalization from emotions, at least partially [74],[78]]. This process led to evolution of ability for music [75]; this is a separate research direction not addressed in this paper.

Another mystery of human-cognition, which is not addressed by current mathematical linguistics, is basic human irrationality. This has been widely discussed and experimentally demonstrated following discoveries of Tversky and Kahneman [[79]], leading to the 2002 Nobel Prize. According to NMF-DL, the "irrationality" originates from the discussed dichotomy between cognition and language. Language is crisp and conscious in the human brain, while cognition might be vague. Yet, collective wisdom accumulated in language may not be properly adapted to one's personal circumstances, and therefore be irrational in a concrete situation. In the 12th c. Maimonides wrote that Adam was expelled from paradise because he refused original thinking using his own cognitive models, but ate from the tree of knowledge and acquired collective wisdom of language [28].

The dual-model also suggests that the inborn neural connection between cognitive brain modules and language brain modules is sufficient to set humans on an evolutionary path separating us from the animal kingdom. Neural connections between these parts of cortex existed millions of years ago due to mirror neuron system, what Arbib called "language prewired brain" [[80]].

The combination of NMF-DL and the dual hierarchy introduces new mechanisms of language and its interaction with cognition. These mechanisms suggest solutions to a number of psycholinguistic mysteries, which have not been addressed by existing theories. These include fundamental cognitive interaction between cognition and language, similarities and differences between these two mechanisms; word-object associations; why children learn language early in life, but cognition is acquired much later; why animals without human language cannot think like humans. These mechanisms also connected language cognition dichotomy to 'irrationality' of the mind discovered by Tversky-Kahneman, and to the story of the Fall and Original sin.

The mathematical mechanisms of NMF-DL-dual model are relatively simple (eqs. (2) through (4), also see details in the given references). These mathematical mechanisms correspond to the known structure and experimental data about the brain-mind. In addition to conceptual mechanisms of cognition they also describe emotional mechanisms and their fundamental role in cognition and world understanding, including role of aesthetic emotions, beautiful, sublime, and musical emotions [[81],[82],[75].

#### Experimental data

An experimental indication in support of the dual model has appeared in [[83]]. That publication has demonstrated that the categorical perception of color in prelinguistic infants is based in the right brain hemisphere. When language is learned and access to lexical color codes becomes more automatic, categorical perception of color moves to the left hemisphere (between two and five years) and adult's categorical perception of color is only based in the left hemisphere.

This provides evidence for neural connections between perception and language, a foundation of the dual model. It supports another aspect of the dual model: the crisp and conscious language part of the model hides from our consciousness the vaguer cognitive part of the model. This is similar to what we observed in the close-open eye experiment: with opened eyes we are not conscious about vague imaginations.

#### Syntax and language faculty

Whereas Chomskyan linguists postulated syntax to be a separate inborn "box" in the mind, dual hierarchy supports the cognitive linguistic idea that syntax is a conceptual mechanism. Specifics of syntax according to this paper are encoded in the concept-model contents at layers of phrases and sentences. The hierarchy evolves based on the dual model, and the syntax is learned from surrounding language. Is syntax

determined by structures in the world or is it a cultural invention? In addition to the dual model, what other linguistic knowledge must be inborn?

“Language is, fundamentally, a system of sound-meaning connections” [[84]]. A language faculty is responsible for this connection; it generates internal representations, maps them into the sensory-motor interface, and into the conceptual-intentional interface. This connects sounds and meanings. The above reference emphasizes that the most important property of this mechanism is recursion. However, it has not specified mechanisms how recursion creates representations, or how it maps representations into the sensory-motor or conceptual-intentional interfaces. The current paper suggests that recursion is not necessarily a fundamental property of a language faculty. Recursion is accomplished by DL and the hierarchy: a higher layer generates the next lower layer models, which accomplishes recursive functions. The hierarchy, in turn, is a result of the dual model. It emerges in operations of the dual model and DL in a society of interacting agents with intergenerational communications.

According to this paper, a single neurally-simple mechanism is sufficient for evolution of human language and cognition. Still further experiments elucidating properties of the dual model are needed. The dual model also maps linguistic and cognitive representations. This paper also challenges an established view that specific vocalization is “arbitrary in terms of its association with a particular context.” In animals, voice directly affects ancient emotional centers. In humans these affects are obvious in songs, and still persist in language to a certain extent. The dual model frees language from emotional encumbrances and enables abstract cognitive development to some extent independent from primitive ancient emotions. Arbitrariness of vocalization (even to some extent) could only be a result of long evolution of vocalizations from primordial sounds [[85],35]. Yet the sound-intentional interface might play an important role in functioning of the dual model. Connecting two parts of the dual model requires motivation, which is facilitated by the emotionality of language [53].

#### Future research

The dual model implies a relatively minimal neural change from the animal to the human mind. It could emerge through combined cultural and genetic evolution and this cultural evolution might continue today. DL resolves a long-standing mystery of how human language, thinking, and culture could have evolved in a seemingly single big step, too large for an evolutionary mutation, too fast and involving too many

advances in language, thinking, and culture, happening almost momentarily around 50,000 years ago [[86],[87]]. DL along with the dual model explains how changes, which seem to involve improbable steps according to logical intuition, actually occur through continuous dynamics. The proposed theory provides a mathematical basis for the concurrent emergence of hierarchical human language and cognition.

Solutions to several principled mathematical problems have been suggested, involving combinatorial complexity. Initial neuro-imaging evidence supports the DL mechanism proposed in this paper, still much remains unknown. DL was experimentally demonstrated for the perception of a single object; these experiments should be extended to the perception of multiple objects in a complex context, as well as for higher level cognition. Evolution of languages can be studied using the developed theory and societies of intelligent agents [[88]].

Recursion as the fundamental mechanism setting human language apart from animal abilities has been challenged in this paper. It is proposed instead that recursion is accomplished by the hierarchy. The fundamental mechanisms enabling the hierarchy, recursion, and connection of language and cognition are the dual model and DL. The paper also challenges the idea of arbitrariness of vocalization. It is suggested that a significant degree of arbitrariness in current languages is a distal result of millennia of language evolution in the presence of the dual model. Instead of assuming arbitrariness as fundamental, future research should concentrate on its emergence from the primordial fusion of sound and emotion.

Future research should address evolutionary separation of cognition from direct emotional-motivational control and immediate behavioral connections. Remaining emotionalities of different languages and their effects on cultural evolution shall be addressed.

Mathematical simulations of the proposed mechanisms should be extended to the engineering developments of Internet search engines with elements of language understanding. The next step would be developing interactive environments, where computers will interact among themselves and with people, gradually evolving human language and cognitive abilities.

#### Acknowledgment

I am thankful to M. Alexander, M. Bar, R. Brockett, M. Cabanac, R. Deming, F. Lin, J. Gleason, R. Kozma, D. Levine, A. Ovsich, and B. Weijers, and to AFOSR PMs Drs. J. Sjogren, and D. Cochran for supporting part of this research.

## References

---

- [1]. Gödel K. Collected works. Ed. S. Feferman, J. W. Dawson, Jr, S. C. Kleene. New York: Oxford Univ. Press. 1931/1994.
- [2]. Chomsky N. Aspects of the theory of syntax. Cambridge: MIT Press, 1965
- [3]. Chomsky N. Principles and parameters in syntactic theory. In N. Hornstein and D. Lightfoot, eds, Explanation in linguistics. the logical problem of language acquisition. London: Longman, 1981.
- [4]. Chomsky, N. The minimalist program. Cambridge: MIT Press, 1995.
- [5]. Croft W, Alan C. Cognitive linguistics. Cambridge: Cambridge University Press, 2004.
- [6]. Evans V, Green M. Cognitive linguistics: an introduction. Edinburgh: Edinburgh University Press, 2006.
- [7]. Ungerer F, Schmid H-J. An introduction to cognitive linguistics. New York: Pearson, 2006.
- [8]. Hurford J. The evolution of human communication and language. In P. D'Etorre & D. Hughes, Eds. Sociobiology of communication: an interdisciplinary perspective. New York: Oxford University Press, 2008:249-264.
- [9]. Christiansen MH, Kirby S. Language evolution. New York: Oxford Univ. Press, 2003.
- [10]. Cangelosi A, Parisi D. Eds. Simulating the Evolution of Language. London: Springer, 2002
- [11]. Cangelosi A, Bugmann G, Borisyuk R. Eds. Modeling Language, Cognition and Action: Proceedings of the 9th Neural Computation and Psychology Workshop. Singapore: World Scientific, 2005.
- [12]. Cangelosi A, Tikhanoff V, Fontanari JF, Hourdakis E. Integrating language and cognition: A cognitive robotics approach. IEEE Computational Intelligence Magazine, 2007; 2(3):65-70
- [13]. Brighton H, Smith K, Kirby, S. Language as an evolutionary system. Phys. Life Rev, 2005; 2(3):177-226.
- [14]. Perlovsky LI. 'Vague-to-Crisp' Neural Mechanism of Perception. IEEE Trans. Neural Networks, 20(8), 2009:1363-1367.
- [15]. Grossberg, S. Neural networks and natural intelligence. Cambridge: MIT Press, 1988.
- [16]. Grossberg S, Levine DS. Neural dynamics of attentionally modulated Pavlovian conditioning: blocking, inter-stimulus interval, and secondary reinforcement. Psychobiology, 15(3), 1987:195-240.
- [17]. Gnadt W, Grossberg S. SOVEREIGN: An autonomous neural system for incrementally learning planned action sequences to navigate towards a rewarded goal. Neural Networks, 21, 2008:699-758.
- [18]. Juslin PN, Västfjäll D. Emotional responses to music: The Need to consider underlying mechanisms. Behavioral and Brain Sciences, 2008; 31:559-575.
- [19]. Perlovsky LI. The Mind is not a Kludge, Skeptic, 2010; 15(3):51-55.
- [20]. Perlovsky LI. Conundrum of Combinatorial Complexity. IEEE Trans. PAMI, 1998; 20(6):666-670.
- [21]. Perlovsky LI. Toward physics of the mind: concepts, emotions, consciousness, and symbols. Physics of Life Reviews, 2006; 3:23-55.
- [22]. Perlovsky LI, Webb V.H, Bradley, S.A, Hansen, C.A (1998). Improved ROTH Detection and Tracking using MLANS. Chapter in Probabilistic Multi-Hypothesis Tracking, Ed. R.L. Streit, NUWC Press, Newport, RI, pp.245-254.
- [23]. Mayorga R, Perlovsky LI. Eds. (2008). Sapient Systems. Springer, London, UK.
- [24]. Perlovsky LI. (2000). Neural Networks and Intellect: using model based concepts. New York: Oxford University Press.
- [25]. Perlovsky LI, (1996). Gödel Theorem and Semiotics. Proceedings of the Conference on Intelligent Systems and Semiotics '96. Gaithersburg, MD, v.2, pp. 14-18.
- [26]. Perlovsky LI, Mayorga R. (2008). Preface. In Sapient Systems, Eds. Mayorga, R, Perlovsky LI, , Springer, London.
- [27]. Perlovsky LI, McManus MM. (1991). Maximum Likelihood Neural Networks for Sensor Fusion and Adaptive Classification. Neural Networks 4 (1), pp. 89-102.
- [28]. Perlovsky LI. Multiple Sensor Fusion and Neural Networks. DARPA Neural Network Study, Lexington, MA: MIT/Lincoln Laboratory, 1987.
- [29]. Perlovsky LI, Kozma, R. (2007). Editorial - Neurodynamics of Cognition and Consciousness, In Neurodynamics of Cognition and Consciousness, Perlovsky, L, Kozma, R. (eds), Springer Verlag, Heidelberg, Germany.
- [30]. Perlovsky LI, (2007). The Mind vs. Logic: Aristotle and Zadeh. Society for Mathematics of Uncertainty, Critical Review, 1(1), pp. 30-33.

## Illustrations

### Illustration 1

.

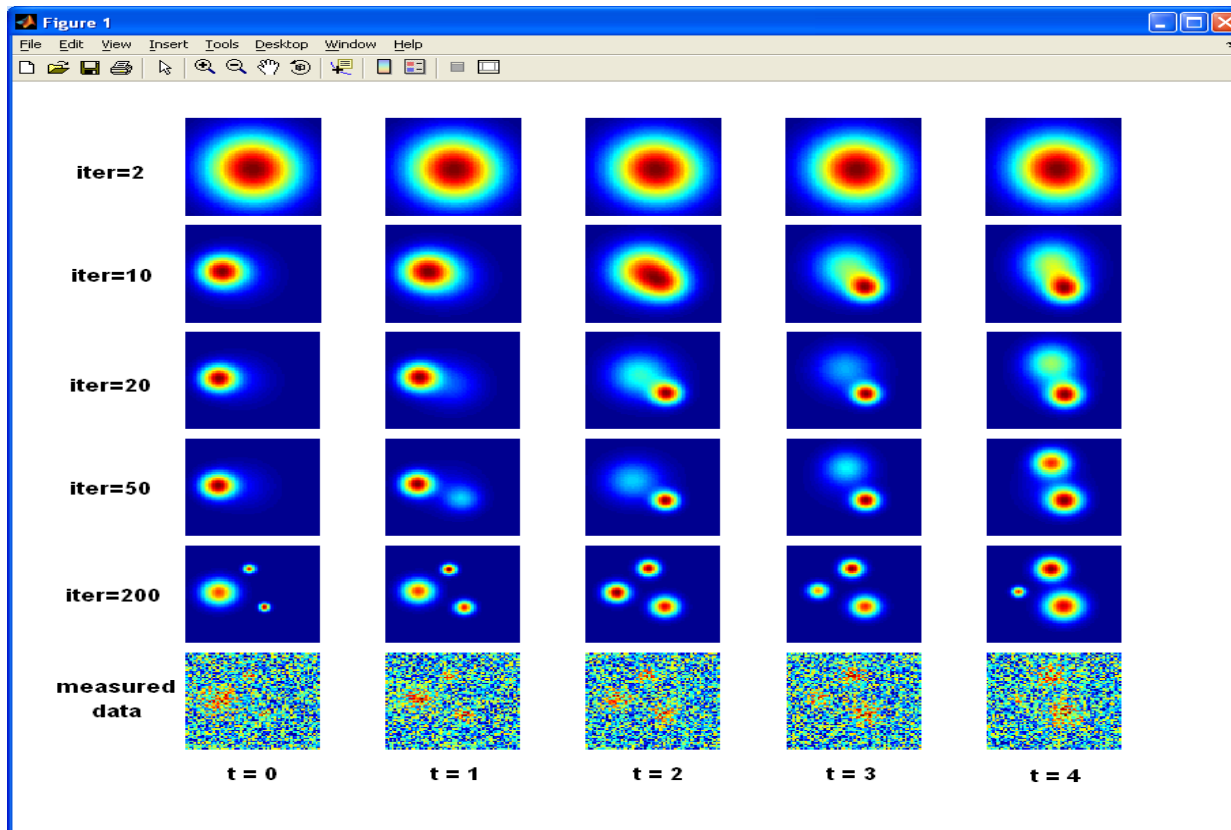
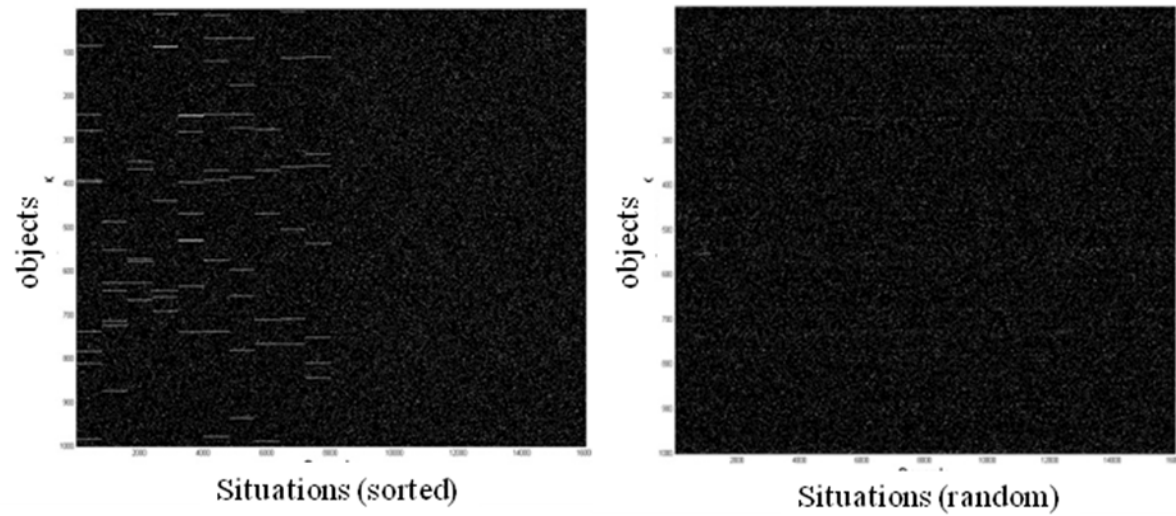


Figure 1. Dynamic logic operation example, finding cognitively-related events in noise, in EEG signals. The searched patterns are shown in Fig.1 at the bottom row. These events are “phase cones,” circular events expanding or contracting in time (horizontal direction  $t$ , each time step is 5 ms); in this case two expanding and one contracting event are simulated as measured by an array of 64x64 sensors. Direct search through all combinations of models and data leads to complexity of approximately  $M^N = 10^{10,000}$ , a prohibitive computational complexity. The models and conditional similarities for this case are described in details in [29]: a uniform model for noise (not shown), expanding and contracting cones for the cognitive events. The first 5 rows illustrate dynamic logic convergence from a single vague blob at iteration 2 (row 1, top) to closely estimated cone events at iteration 200 (row 5); we did not attempt to reduce the number of iterations in this example; the number of computer operations was about  $10^{10}$ . Thus, a problem that was not solvable due to CC becomes solvable using dynamic logic.

## Illustration 2



### Illustration 3

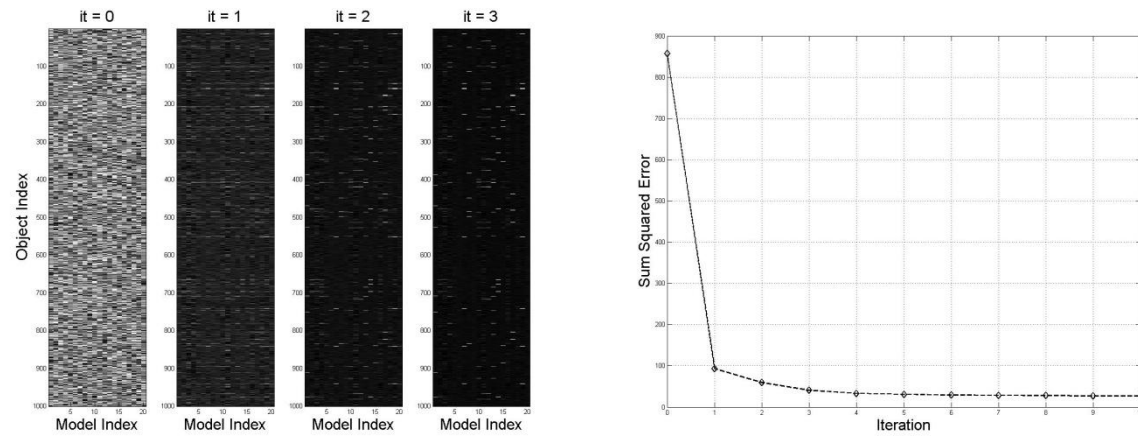
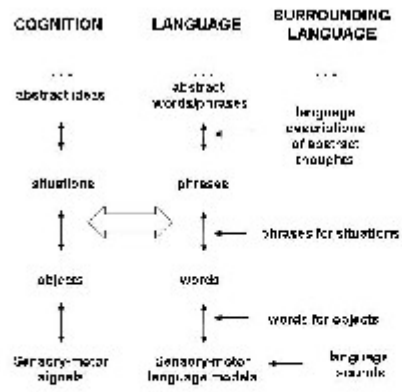


Fig. 3. (a) shows DL initiation (random) and the first three iterations (Fig. 5a); the vertical axis shows objects and the horizontal axis shows models (from 1 to 20). The problem is approximately solved by

## Illustration 4

Parallel hierarchies of language and cognition



## Disclaimer

This article has been downloaded from WebmedCentral. With our unique author driven post publication peer review, contents posted on this web portal do not undergo any prepublication peer or editorial review. It is completely the responsibility of the authors to ensure not only scientific and ethical standards of the manuscript but also its grammatical accuracy. Authors must ensure that they obtain all the necessary permissions before submitting any information that requires obtaining a consent or approval from a third party. Authors should also ensure not to submit any information which they do not have the copyright of or of which they have transferred the copyrights to a third party.

Contents on WebmedCentral are purely for biomedical researchers and scientists. They are not meant to cater to the needs of an individual patient. The web portal or any content(s) therein is neither designed to support, nor replace, the relationship that exists between a patient/site visitor and his/her physician. Your use of the WebmedCentral site and its contents is entirely at your own risk. We do not take any responsibility for any harm that you may suffer or inflict on a third person by following the contents of this website.