Information-Processing Model of Concept Formation – Is First Language Acquisition Universal?

Velina Slavova

New Bulgarian University, Department of Informatics, 21 Montevideo str., Sofia, Bulgaria
E-mail: vslavova@nbu.bg

Abstract: *The analysis of child's speech corpora shows that the process of acquisition of English and French displays identical development of children's expressions when the speech-utterances are presented as Fibonacci-weighted classes of concepts. A model of concept complexity and information processing based on principles of optimality is proposed to explain this statistical result.*

Keywords: *Concept formation, Cognitive model, Information, Entropy, Fibonacci sequence.*

1. Introduction

The core hypothesis explored in this work concerns the existence of biologically determined cognitive universals that underlie meaning-representations and language faculty, and which observe principles of optimality.

Based on data from child's language acquisition, this paper considers the process of conceptualization. More precisely, it focuses on information treatment and proposes an abstract model of hierarchical, step-like incorporation of information flows that trigger concept formation. The analysis performed to build the model is related to several domains, such as cognition, brain imaging, language, evolution, information theory and so forth. Questions regarding the semantic tiling of the cortex and as to the principles of obtaining meaningful internal representations are foci of research efforts that have led to several astonishing findings. However, many questions that arise when analyzing such findings are in want of definitive, or indeed any form of, answers, particularly at the level of algorithmic explanation (see [4, 5]). The effort in this paper is towards modeling the information-construction principles of concept formation, explored in terms of processing-complexity. The proposed model explains results revealed by statistical analysis of large corpora of child speech in two languages – English and French. The main contemporary results that prompt the model's underlying reasoning are briefly discussed in this introductory section.

Brain activity evoked by concepts when watching naturalistic movies, as well as when listening to passages of spoken narrative, has been found to share a common

basis across individuals (e.g., [20, 21, 27]). Recent studies suggest commonality of neural representations of sentences across individuals [52], while others show that the brain-representations of concepts ([11, 37]), and of sentences [55] in different languages, have a common basis. These recent neuroimaging studies suggest that concept- and language-related mental representations are based on distinguishable patterns of brain activation and have a common cortical activation-picture across individuals and across languages.

From the information processing standpoint, the questions that arise concern: 1) the information sources that provide inputs leading to such homogeneity of the language-related brain maps; 2) the principles that lead the information treatment to this similarity.

The first path of reasoning, followed herein is that conceptualized mental images are based on perception and actions of the individual in the environment. This assumption is derived from widely accepted theories in cognition (e.g., [3, 4]) and from brain studies (e.g., [15, 23]). The background adopted here follows the *grounded cognition* approach, where a *Concept* is a "*dynamically distributed network in the brain coupled with a category in the environment or experience, with this network guiding situated interactions with the category's instances.*" [2].

The second path of reasoning is related to the initial settings of the brain – in other words, its genetically underwritten capacity to conceptualize. Based on recent studies in sign language and home-sign languages (e.g., [17, 29]) showing that a child can create a combinatorial linguistic system without external linguistic input, the view adopted here is that there exist *inborn* conceptualization abilities.

The third path of reasoning concerns the role of information coming from the biological system itself and is related to the conceptualization of the *Self* as *Actor in the Environment*. The *Self-centred* model [45] posits that, due to the synchronized mechanisms of perception, proprioception, interoception, processing in the Mirror neuron system and Default Mode Network (DMN), a human has a biologically underwritten notion about the existence of his Self as Actor in the Environment, and that this is of fundamental importance for language faculty. Recent results [11] confirm the involvement of DMN in processing of language concepts. With relation to the "language of thought", contemporary results [30] show that brain activity in the self-relating domains of agency, self-recognition, emotions, resting state, etc., causes inner speech. Following the previous analysis [45], concept formation starts at perceptual and emotional levels, where meaning is accorded in active interaction with the world's instances as regards their importance for the Subject. This suggestion is supported by brain studies showing that the brain evaluates Self-related information based on the information's meaning/function for the Self [22]. It was recently suggested [33] that cortical networks exist, related to the representation of Self. Memories related to the Self are based on self-referential processes, rather than on familiarity of the information [26]. The consistency of the Self-centred model was shown with relation to language acquisition [46].

In addition to these arguments underlying the proposed model, the further reasoning follows one main suggestion promoted in previous papers [45, 46] – namely, that the process of meaning-congregation and the operation of syntactic

4

construction known in linguistics as "*Merge*" [8] are mediated by one and the same mechanism. In this paper, the binary Merge is seen as an information-constructing operation which leads to establishing of internal representations, called *meaning-units*. A meaning-unit corresponds to a distinct brain activation, obtained by integrating information from engaged brain subsystems, which is consequently memorized. The system of meaning-units is called in cognitive science a *semantic description of the world*.

One key idea promoted in the model here proposed is the relating of *interoceptive* and *exteroceptive* information in a common information-treatment scheme. Specialized studies [38] propose that such integration of interoceptive and exteroceptive signals occurs in contexts of Self-related processing. The connection between interoceptive and exteroceptive information is upheld by analysis of data from brain studies, concluding that "*there is no way for interoceptive stimuli other than to be processed in relation to exteroceptive stimuli and vice versa*" [32].

The study presented here was accomplished in the following chronological order: First, data from language-acquisition corpora were collected, treated, spread into meaning-classes and analysed. After that, the Fibonacci-tree model of syntactic operations, developed in earlier publications [44-47], was applied to the meaning-classes. The use of this model showed a remarkable fit to the statistical data. Effort then turned to elucidating the information-treatment model, proposed herein, with the objective of examining the general causes accountable for the Fibonacci-tree's statistical confirmation.

2. Data description and problem statement

If the brain's means of representing concepts has a common basis across individuals and across languages, this basis would be detectable in the earliest formative years of children's acquiring their particular first language. A number of comparative studies in language-acquisition concentrate on language-specific aspects (e.g., [48]) but do not investigate processes underlying a common mechanism of concept-formation. This paper's substance is directed toward finding some underlying information-processing schemes that can explain close correspondences observed in the acquisition of two languages – English and French.

Data from 42 corpora containing 1,515 free dialogues with child' speech in English and in French, annotated with Part Of Speech (POS), were extracted from CHILDES (CHIld Language Data Exchange System [56]) and used for statistical analyses [40]. The English data collection contains 620 dialogues (62 girls, 66 boys, and 7 children of unspecified gender); the French collection contains 895 dialogues (157 girls and 141 boys). Two large corpora containing speech of children aged between 9 and 62 months were obtained (125,873 speech-utterances for English and 153,824 for French language production), and analysed.

In order to compare the acquisition at the language level, the Ratio per Utterance (RU) for each POS was calculated using equation (1). The RU shows the extent of use of a given POS within an utterance "exemplary" for a dialogue:

$$RU\left(POS_{ij}\right) = \frac{N_{Pos_{ij}}}{N_j},$$

(1)

where: POS_i is one of the POS annotated in the corpora (25 identical for English and French); j is the number of dialogue; N_j is the number of child's utterances in the dialogue j; $N_{Pos_{ij}}$ is the number of the POS_i used by the child in the dialogue j.

The statistical analysis showed that the language-acquisition process displays surprising differences between the two languages when it is measured by means of use of Parts of Speech [41]. If the acquisition process is based on common semantic basis, it should display similarities at some semantic level. A set of *classes of meaning* was proposed and the POS of the two languages were spread into these classes. The proposed classes are: 1) *Entities*; 2) *Relationships*; 3) *Circumstances*; 4). *Quality and Attribution*; 5) *Quantity and Precision*, and 6) *Other*.

Fig. 1 shows the paths of use of the meaning-classes. The resemblance of language-acquisition processes, expressed in terms of use of meaning-classes is considerably bigger than when expressed with use of POS (for details, see [41]). It should be mentioned that the Sums of the expressed meaning-classes in the two languages, measured in words (Sum of POS), showing the emergent process of language production, are correlated at 0.87.
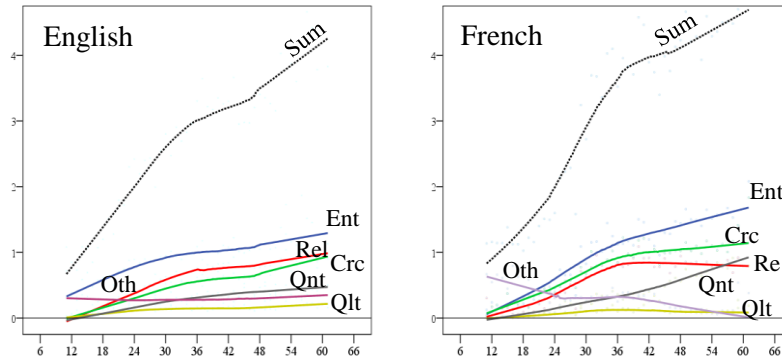

Fig. 1. Development with the age of use of the meaning classes. Ent – Entities, Rel – Relationships, Crc – Circumstances, Qlt – Quality, Qnt – Quantity, Oth – Other, Sum classes (Sum of POS RU) [41]

This statistical picture calls to mind several questions. One is why the classes, which start all being used from the onset of language production, are in different proportion. For example, Entities are acquired much more intensively in both languages – so, are Entities more easily conceptualized? The second question is related to the proportions of meaning-classes – in English the intensity of use is in the order Entities–Relationships–Circumstances (Fig. 1) but in French the order is Entities–Circumstances–Relationships. Since English and French are different languages and, knowing that children's expressions are influenced by the language environment, such a difference in their manner of uttering "events" (to use the linguistics term encompassing these three classes) is not surprising. Notwithstanding that, the Sums of the mentioned three Event-describing classes develop in a quasi-identical way (correlation up to 36 months – 0.905). That suggests that the mental process treats the two schemes with equal effort. One may suppose that there exists

some mechanism of a general nature regulating the advancement in the use of classes by the two language-groups of children. The hypothesis here is that this mechanism is related to the extent of mental-processing cost that the concept formation demands.

Expressing of concepts within speech utterances is an additive operation. That means that some "global-speech" additive term, showing the Overall Expressed Concepts (Expr), such as:

(2) $Expr = c_1*Ent + c_2*Rel + c_3*Crc + c_4*Qlt + c_5*Qnt + c_6*Oth,$

where: Ent, Rel, Crc, Qlt, Qnt, and Oth are the observed in the corpora RU of the classes and the coefficients c_i, $i = 1, …, 6$, denote the *complexity* of the classes should equilibrate the differences between the two language acquisition processes. One concrete problem is to find such c_i for which the term (2) leads to maximal similarity in the acquisition of the two languages.

3. Model assumptions and goal

In the model proposed here, the creation of meaning is portrayed as a process in which inborn neuronal mechanisms organize information so as to establish distinct *meaning-units*. The main framework of the model is that there exists a level of internal representation at which the information operable by the brain takes the form of meaning-units necessary to guide the actions of the Actor in the Environment. Meaning-units are *unlexicalized concepts,* to which word-form(s) are likely to become coupled as the consequence of interactions with the language-environment (Fig. 2).
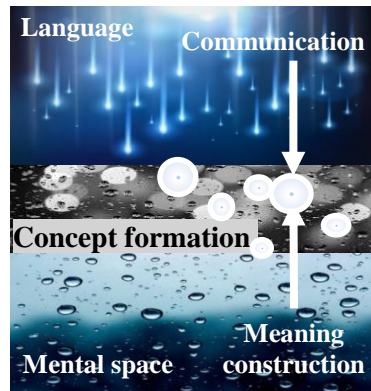

Fig. 2. Abstract Model. The process of concept-formation presented as interaction between internally generated meaning-units and the language environment

Blending of the internally created meaning-units with presented language ingredients results in the creation of a lexicalized concept – a meaning-unit labelled with a name. These names are subsequently applied as words in children's speech. It should be underlined that – as the data revealed – children often devise phonological labels of their own for the meaning-units they have formed, and use them with considerable persistence in communicating.

In the approach described here, it is important to distinguish the process of constructing meaning-units as a separate, internal and biologically supported initial stage of the extended process of concept-formation (Fig. 2). Language-environment, of course, influences concept-formation and the structure of the semantic space. This influence is however partial, as recent studies of language-related brain activity show. Even for speakers of French and Mandarin, the patterns of brain activation for representing lexicalized concepts indicate a common neural basis – the main differences found are related to social and cultural concepts [37].

It is proposed here that the construction of meaning-units is based on inborn information-treatment mechanisms which have the following features:

• They are based on two information flows consisting in coded neuronal activity (Fig. 3). The first (herein termed exogenous) is conducted from the outside world. The other (termed endogenous) is sourced inside the system and provides a prototypical signal created by the organism;

• The creation of concepts necessitates memory – genetically transferred brain mechanisms and stable paths as well as individual memory;

• Concepts are formed using different amounts of information-treatment operations. This separates them into classes with different processing complexity.
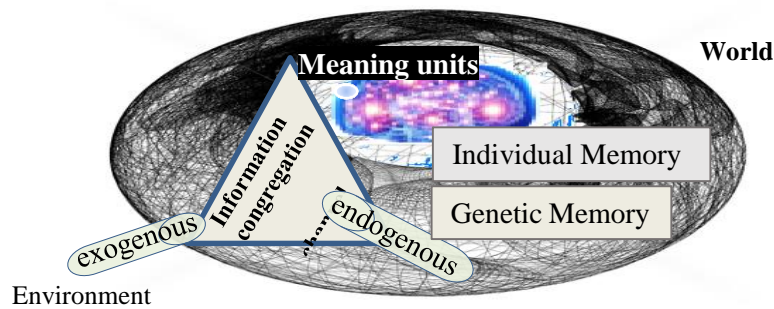

Fig. 3. General scheme of the model

As the reasoning that follows uses terms that have different definitions within different domains, how these terms should be understood in the context of the present work is clarified below.

Information. In compliance with the distinction between *information* and *data* accepted in the domain of data mining, where the process of data mining is defined as extraction of implicit, previously unknown and potentially useful information from data, the term *information*, used hereafter, is based on the following definition: "*Data are sensory stimuli that we perceive through our senses. Information is data that has been processed into a form that is meaningful to the recipient.*"[10].

In the approach described here, information is formed within the system by means of an information-congregation process (Fig. 3). The model is designed to describe the formation of meaning-units which are expressed by means of discrete language-units – hence, information is forthwith considered in a discrete context.

Information-units are presented here as the building material of meaning representations. They are seen as distinct activations represented and conveyed over

brain subsystems. *Meaning-units* are information-units that are related to semantic memory and can be used in autonomous regime.

Data and encoding. The nervous system is found to use numerous manners of coding (e.g., signal frequencies, sequences, oscillations, activation patterns, amplitudes, signal packages, wave densities of dendrites, etc.) to convey information. In the proposed model the physical nature of the transmission is not discussed; the diverse methods of coding are seen as features (parameters) of the input data and of the obtained and transmitted information-units.

The following principles of constructing of information-units are proposed:

1. The *Self-mapping* principle: Information is created inside the system by linking exogenous data from the environment with endogenous data sources;

2. The *multipath* principle: Meaning representations implicate information-units coming from neuronal activity dispersed over the brain;

3. The *parallel* principle: the information treatment is based on processes which run in parallel;

4. The *gradual integration* principle: Meaning obtaining in steps of gradual processing where each step uses information-units obtained in the previous-step;

5. The *grounding* principle: To form a meaning-unit, the treatment requires input from the environment, instantly or in the past;

6. The *temporal storage* principle: the processing of information-units on each step requires working memory;

7. The *recursive* principle: the results obtained on the steps are based on one and the same scheme of processing.

The goal of this work is to elaborate a general model of such treatment – its data sources, the manner of obtaining and classifying the gained information.

4. Model of information-gathering

In the proposed model, information-units are created by an operation which relates the endogenous and the exogenous paths, called *i-Merge*. The general logic of this operation is that it creates an information-unit by linking data coming from the environment with data generated by the system itself.

The i-Merge operation is presented as a general scheme; its description in neuroscientific terms, such as the manner of neural encoding involved and a listing of features contributing thereto (presented in Fig. 4a as encoded parameters), falls outside the scope of the present article. What is relevant here is that the nervous system *detects and responds to disparities*, and the general idea of i-Merge is that it operates in response to similarities/dissimilarities between the two inputs. The capacity of the nervous system to encode physical stimuli has biological constraints and the number of parameters used for neuronal coding is finite. Accordingly, the number of parameters entering i-Merge is finite and the treatment can produce a finite number of different information-units.

The information treatment necessitates a short-term memory buffer, known in cognitive science as Working Memory (WM). Exactly how this (and any other) form of memory works remains unclear, but numerous models and experiments show that

the storage capacity of WM is limited and suggest that WM buffers are specific for the semantic type of information being stored. To reflect this, each i-Merge uses WM for the processing and the produced unit is encoded with parameters that are not strictly the same as the input parameters (Fig. 4a).
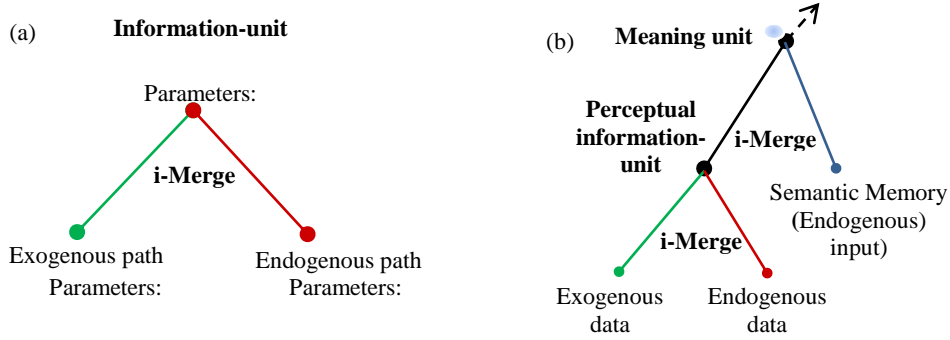
Fig. 4. General scheme of the i-Merge operation (a); Initial i-Merge (b)

Multimodal perception is involuntary, automatic and deeply incorporated in the sensory systems. In the model, the initial creation of an information-unit happens by i-Merging endogenous and exogenous data and indicates the gateway of multimodal perception (Fig. 4b). The created *perceptual information-unit* has its specific encoding-configuration and enters the higher layers of brain processing.

Perceptual information-units appear with different frequency depending on the experienced occurrence of exogenous data. Although it is not exactly understood how memory works, all agree that the repetition of an input creates a memory representation. The brain could not operate on information-units in its autonomous regime if the perceptual information were, moreover, not represented, memorized, and retrievable at some conscious level of functioning (see [50]). In order to become meaning-units, perceptual units have to be connected with higher levels of internal representation. The relation to *semantic memory* (Fig. 4b), proposed in the model, signifies a connection to levels of conscious activity. Initially, personally acquired semantic memory is meagre and its influence can be seen as noise which, inevitably, is compared with the perceptual information. The introduced uncertainty motivates extensive engagement with the environment – in fact, that is how the system works to guarantee ecological adaptation.

The information-units based directly on perception are the least abstract and their level of abstraction is determined here as A_1. In the proposed model, the higher levels of abstraction are based on information coming from the previous levels and on additional inputs from exogenous and endogenous sources.

The second level of abstraction, A_2, is based on level A_1 and requires novel exogenous and endogenous inputs (Fig. 5a). The novel data-sources are of different physical nature to the sources of level A_1, since, if they were of the same nature, they would be encoded using the entries of A_1 and i-Merged in A_1. Such a process of bottom-up i-Merging has to be applied successively so as to form a hierarchy of abstract levels. As i-Merge is defined as a two-input operation, the supposed hierarchy will be expressed by some binary tree.

10

It is known that the brain exhibits a hierarchical structure of meaning-related activation (e.g., [34]). There are no indications that, for example, the frontal areas are directly related to primary sensory areas. The modelled here information-units should not interact if they differ greatly in their levels of internal representation. The nodes of the tree should not relate subtrees with big differences in height – that is, the tree has to be somehow balanced in terms of height.

To suggest an accurate type of binary tree, the appropriate question to be asked is: "*has the underlying process some principles of optimality to observe?*". Next in this section, the tree is assembled by applying optimality principles.

The processing at level of abstraction A_1 relates some signals originating within the organism with signals describing environmental stimuli and creates internal representations that can be memorized. As noted earlier, these representations are limited in number as the encoding of the stimuli has biological constraints. Looking at the process in its evolutionary context, an organism, in order to further develop its informational description of the world, has to develop novel, different pathways to reflect the environment and other, novel internal prototypical signals based on its own functioning and on enquiring the existing meaning-units of level A_1.
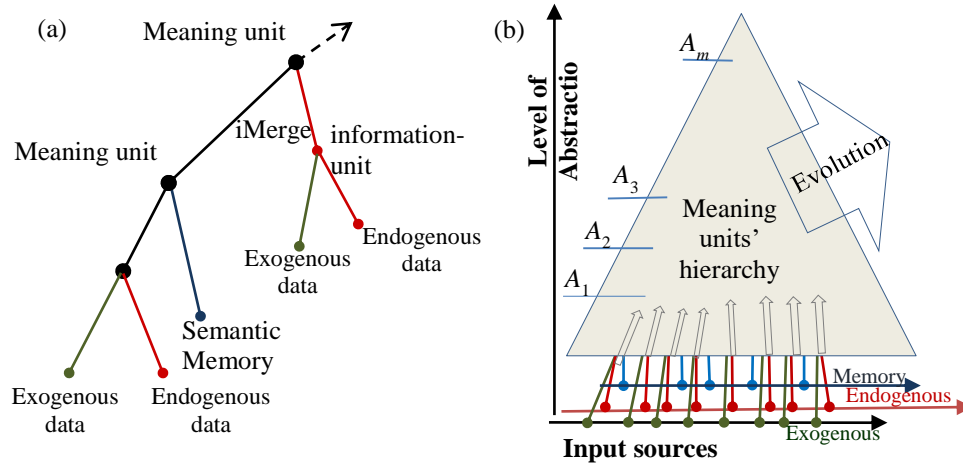

Fig. 5. Constructing the second level of abstraction (a); process of Information establishing, seen in the context of Evolution (b)

As shown in Fig. 5b, the overall process of creating increasingly abstract units necessitates memory and additional input gateways. Thus, three sources are engaged – the exogenous and endogenous inputs and the memory. This schematic description shows the nature of the constraints – the endogenous and exogenous inputs are biologically determined as number and capacities of transmission. Memory, too, is a fundamental resource limited by biological constraints.

Let us assume that biological evolution (Fig. 5b) is a process that observes principles of optimality because the changes implicated in organisms' speciation have become irreversible when they have led to a solution that is well-matched to the species' typical environmental conditions. In general, evolutionary development leads to organisms exhibiting increasingly complex behaviour. That signifies that their interaction with the environment is mediated by increasingly performant

information handling. Such increasingly effective interacting is an expression of the *number of information-units* that the information-treatment can create to represent the environment internally. This number is a factor which is maximised.

The factor that has to be minimized is the *metabolic cost*, reflecting the energy needed for signal transmission within the nervous system. The edges in the abstract scheme in Fig. 5a represent paths of information transmission. Biologically, these paths engage complex processes running within configurations of nerve-cells. In the model this cost is symbolized by the Cost of the branches.

Let us first look at the terminal branches (Fig. 5a). The left branch relates the exogenous input and depicts spontaneous sensory coding of the outside stimuli. This corresponds to bottom-up transmission. The right branches represent the inclusion of endogenous data related to genetic memory, interoception, proprioception, emotional state, semantic memory etc., and can be seen as involving the entire brain. The right terminal branches represent inclusion of a top-down information flow generated in consequence of the precise exogenous input. In terms of information system reasoning, the transmission through the left branch is analogous to insertion of new data entry while the top-down supply through the right branches is analogous to an operator, specifying "*Select* the necessary features *From* the brain *Where* their values correspond to this exogenous input (concrete stimulus)". In this very schematic representation, the exogenous (left) input is a parameter of the *where* condition transmitted on the endogenous (right) path. The process of top-down retrieval runs using several neuronal pathways and relies on WM – a metabolically expensive resource, found to engage complex structures distributed across the brain (see, e.g., [14]). In the generalized net model of WM [42], it was suggested that, due to the limited resource of WM, such a multipath retrieval process adjusts itself to the overall context and produces an ordered list of candidates where the head alone is returned as the top-down result. So, for this model representation, the direct transmission via the left branch has a cost of 1 and the query-and-retrieval via the right ones has a cost of 2.

Let us now analyse the nonterminal branches of the tree under construction. The internal nodes (the i-Merges) are relating information-units produced at the lower nodes and transmitted upward. The paths of i-Merging represent integration of information. The longer the paths are, the more information is integrated and the more brain paths are involved. When the two information-units feeding one internal i-Merge carry different amounts of integrated information, the optimal solution is to store in WM the information coming through the branch having a shorter subtree. If, for example, the right subtree is shorter, the right branch carries smaller amount of integrated information and, in order to minimise the WM charge, it will be the one which is stored in WM. Thus, the upper left branches' cost is 1 (for transmission) and the conduction by the right branches is charged with working memory storage-retrieval and has a cost of 2.

Suppose that a new-born infant is ready-equipped with such a mental device. For the new-born, the meaning-units that the device creates and classifies to levels A_1, ..., A_m consist of information about the world. The more input entropy the treatment is capable to translate into meaning-units, the more productive it is.

Following classical Shannon understanding, entropy is the capacity of a communication channel to transmit information. The transmission capacity here is determined by the entropy converted into meaning units. Let M be the highest level of abstraction that the biological device allows, and n be the number of inputs that it treats, where n corresponds to the number of entry points (endogenous, exogenous and memory) to the hierarchy (Fig. 5b.). This number denotes the biological constraint on the channel's capacity. Let us assign to the hierarchy's input a variable $F(f_1, f_2, f_3, ..., f_n)$ of n encoded data-entries, each presented by a feature vector f_i, where the probabilities of f_i, $(i=1, ..., n)$ are respectively $p_1, p_2, ..., p_n$, and $\sum p_i = 1$. The maximal entropy of the overall input is when $p_1 = p_2 = ... = p_n$.
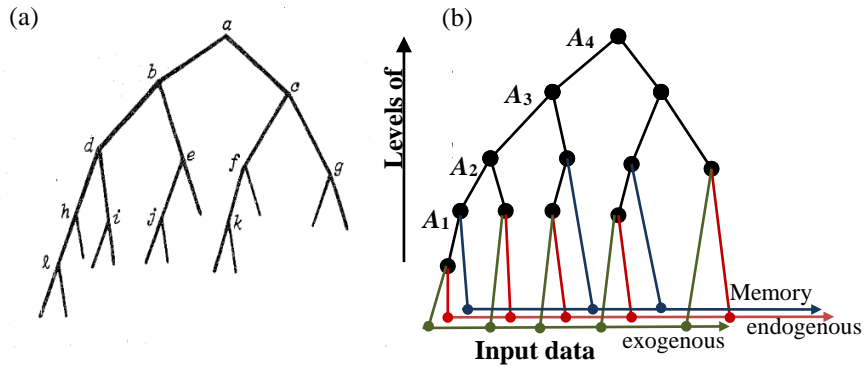

Fig. 6. The tree investigated by H o r i b e [18] (a); the model (b)

The model's settings correspond to the "*Maximal entropy–Minimal cost*" theorem of Horibe. H o r i b e [18, 19] demonstrated that in a binary tree with n terminal nodes weighted with probabilities $p_1, p_2, ..., p_n$, $\sum p_i = 1$, the Shannon entropy is maximal ($p_1 = p_2 = ... = p_n$), and the tree is optimal in the sense of minimal average cost if the tree is a tree of Fibonacci (obtained by φ-weight-balancing at each node, where φ is the Golden Ratio, $\varphi = (1 + \sqrt{5})/2$). The average cost of the tree is defined as $\sum p_i . a_i$, where a_i are the costs of the n terminal nodes. Each a_i is calculated as cost of the path from the leaf to the root, having assigned a cost 1 for the left and 2 for the right branches. Horibe demonstrated that a binary tree with any number of terminal nodes can be rebalanced in order to correspond to the theorem settings.

The "*Minimal cost–Maximal entropy*" tree of H o r i b e [18] is shown in Fig. 6a. The same tree, seen in the context of the model, is redrawn in Fig. 6b. The meaning-units are obtained by consecutive bottom-up i-Merges whose pathways converge to the leftmost branch, where the levels of abstraction are positioned.

The Self-centered character of the tree-model is expressed by the continuous inclusion in the treatment of endogenous information and of individual memory. As illustrated in Fig. 6b, the terminal branches that correspond to top-down inputs are involved as right branches with respect of the settings of Horibe's theorem (as discussed, their cost is 2). All higher i-Merges have shorter right subtrees and, as discussed, the right branches take the charge of WM with a cost of 2.

13

The importance accorded here to internally generated information is in line with contemporary neuroscientific understanding about the functioning of the brain as an autonomous system (e.g., [25]). As L l i n á s [24] concludes, "*Ultimately, we are faced with a system that addresses the external world, not as a slumbering machine to be awoken by the entry of sensory information, but rather as a continuously humming brain willing to internalize and incorporate into its intimate activity an image of the external world, but always in the context of its own existence and its own intrinsic electrical activity.*"

5. Complexity of concept formation

If the meaning-creating mechanism described in the previous section is available at birth, it has evolved so as to work in maximal entropy. Its manner of working can be simulated as a simple computerized model. For simulation purposes, an information-unit is presented as a set of parameters (frequencies, amplitudes, signal packages, etc.) in the form of a *feature-vector* (see Fig. 4a). In principle, i-Merge produces an information-unit (a feature-vector, in this context) on the basis of detecting disparities/similarities between its presented exogenous and endogenously stipulated inputs. Accordingly modelled, the presented feature-vector encodes on each position a particular feature which, if its value falls within a stipulated range, is ascribed a digitized value of 1, and if otherwise, of 0. Such binary representation allows simulating the successive upward steps of i-Merge processing. For implementing the simulation, it was considered that if the input feature-vectors are equal in position i, the binary vector produced by i-Merge will take a value 1 on i-th position. Thus, the resulting feature-vector reflects the similarity between the inputs by means of (reverse) Hamming distance, which represents a metric, so the outputted feature-vectors are ordered. Maximal entropy was simulated using uniformly distributed, randomly generated binary vectors, and the decision as to which of the levels is a given entry-set classified by the bottom-up i-Merging was based on Hamming distance. The simulation showed that in maximal entropy condition, the processing classifies to the levels A_1, …, A_m, with equal frequencies. Indeed, the statistical investigation of the child speech showed that all the meaning-classes start being used from the onset of language production (see [41]). However, their use does not develop at an equal pace over the course of time (see Fig. 1). Further in this section, this phenomenon is explained in terms of *cognitive complexity*, analogous to computational complexity.

For the new-born infant the mechanism described in the previous section in abstract terms is supposed to run automatically when the information flow coming from its novel environment starts being delivered to the treatment hierarchy. The metabolic cost expended by neuronal configurations involved (reflecting the energy needed) was taken into account when obtaining the treatment-structure (Fig. 7a). The main effort in the processing is to cause the information flow. This task is accomplished solely by bottom-level i-Merges. They pertain to the level of mere perceptual experiencing of the environment and express a boundary operation transforming data-input processed at sensory levels into information units amenable

14

to semantic classification at the higher levels of treatment. Let us call the produced units *elementary* units and reorganize the tree into a form comprising only paths that derive from the individual's environmentally triggered experiencing.

The reorganized tree is presented in Fig. 7b. It has the form of a balanced tree in the sense of Adelson-Velskii and Landis (AVL-tree). In computer science, the tree in Fig. 7b, known as a tree of Fibonacci, is qualified as the "worst case" of AVL-tree because it is the highest. As discussed, the number of information-units that can be created by each i-Merge is finite. At the same time, following the model, the bottom gateways represent costly products of biological evolution and their number is limited. So, the optimal solution for the treatment's efficiency is to obtain a maximal number of levels using a given number of input gateways. For the phenomenon modelled here the AVL tree in Fig. 7b represents the best solution – for a given number of inputs it creates the highest AVL-balanced hierarchy.
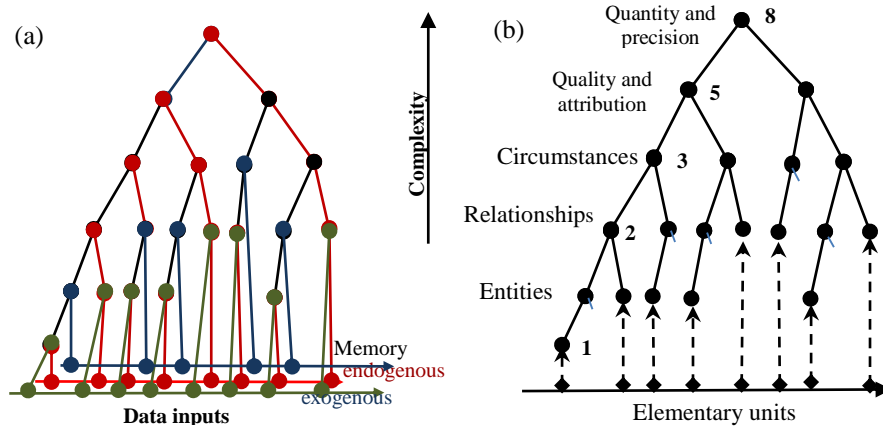

Fig. 7. The treatment tree (a); complexity of the classes (b)

In order to ascribe a measure of the processing load, the environmentally based and effort-requiring operation – the bottom i-Merge – is accorded a complexity load of 1. The complexity of each level is assigned a value equal to the number of bottom i-Merges (creating elementary units) of the corresponding sub-tree. Having this measure of complexity, the loads of the levels are 1, 2, 3, 5 and 8 (Fig. 7b).

The meaning-classes are placed in correspondence with the successive levels of the hierarchical treatment. Their complexity costs are: Entitles – 1, Relationships – 2, Circumstances – 3, Quality – 5, and Quantity – 8. This order respects the level of abstraction of each class. Entities are directly based on multimodal perception. Relationships are concepts about the actions and states of Entities, and obtain a concrete meaning when referred to Entities. Circumstances include spatial and temporal characteristics of Entities and/or Relationships. Quality and attribution require extracting of similar features from Entities, Relationships and Circumstances, and blending them in separate concepts. Quantity and precision need establishing of

concepts that estimate proportions, grouping of Entities or selecting one Entity from a conceptualized group.

6. Statistical investigation and results

As explained, the expressions used by children in free dialogues in two languages were stored in two extensive corpora and statistically treated. The calculated Ratios per Utterance (RU) of the POS over each dialogue reflect the use per utterance of each POS, and form a statistical space for each of the languages, further called "language spaces". The stored children's expressions are most often incomplete, phonetically erroneous and ungrammatical, but in the context of the dialogues they express very well the meaning that the children wished to communicate (see [41]). The assumption is that the expressions' observed content reflects the advancement in the process of concept-formation in children's mental representations.

The meaning-classes proposed can be seen as semantic prototypes of syntactic categories. The set of classes inferred from the Actor in the Environment model proved to be quite similar to the categories proposed by P i n k e r [36] in his "semantic bootstrapping hypothesis". Pinker proposed a model of language acquisition where he supposes the existence of a "semantic inductive basis" that helps children in the acquisition of language rules by means of "syntax-semantic pairing". The basis proposed contains categories such as "name of person or thing", "action or change of state", "attribute" and "spatial relation, path or direction". Pinker's work shows how these categories can lead to the acquisition of syntactic rules.

In order to estimate only syntax-procreative classes, the class "Other" (onomatopoeias, interjections and communicators) was removed from the language spaces. The tree-model allows ascribing weights to the overall expressed concepts in equation (2). The Concept Complexity (CC) based on the classes which meaning can be seen as semantic archetype for building syntactic structures, follows:

(3) $\qquad CC = 1*Ent + 2* Rel + 3*Crc + 5*Qlt + 8*Qnt.$

Equation (3), derived by the model, allows evaluating the development of the complexity of children's utterances with age. The CC was calculated using the statistical result for the use of meaning-classes in each dialogue. Fig. 8 shows the development of CC in language production from 11 up to 62 months.
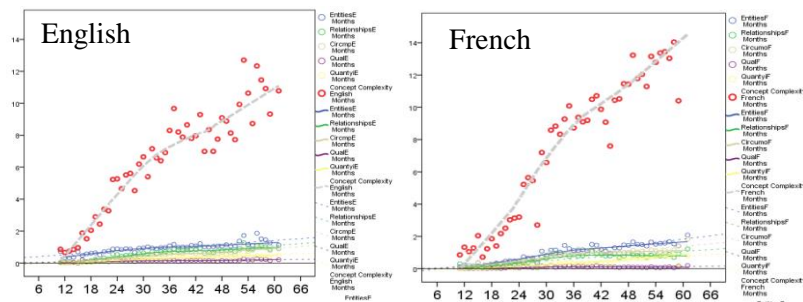

Fig. 8. Development of the CC over the time-course of language acquisition. The lower lines show the development of the classes of meaning

16

The between-languages correlation of the CC is 0.933. This is the biggest similarity between the language-acquisition processes found in the statistical study (Table 1).

Table 1. Between-languages correlations (Pearson, two-tailed, the *p*-values are 0.00)

language and model parameters	Between-languages correlations	
	Correlation (Average)	Maximal correlation
Parts of speech (25 identical POS in English and French)	0.46	0.84 (for Prepositions)
Classes of meaning (without the class "Other")	0.77	0.89 (for Circumstances)
Sum of POS (indicating the length of an utterance, in words)	0.87	
Concept Complexity	0.93	

The fact that the reasoning of this approach led to the appearing of the Golden ratio, a parameter often detected in natural processes and usually associated with behaviour called "efficient growth", is not extraordinary *per se*. It is however somewhat surprising that the statistical treatment of the use of POS spread into meaning classes fits this mathematical scenario. The latter fact can be interpreted as follows: Data from language-acquisition corpora suggests that concept-formation occurs in the same way in the examined languages because it is based on a general underlying principle which expresses a gain of maximal amount of classified information using minimal amount of energy.

The question is: to what extent is this statistical result reliable? The language parameters measured in children's speech are described by increasing functions and the correlations between them are high in every case. What needs to be ascertained is whether a similar result could be obtained by chance because of the nature of the acquisition process. A standard Monte-Carlo approach was applied for bootstrapping the correlation between the two language-spaces. The data, that is – the RU of the classes, was bootstrapped with random integers in the interval 1, ..., 8 used as coefficients c_i in (2) (as explained, the class "Other" was removed) and the between-languages correlation was calculated for each compound of five coefficients.
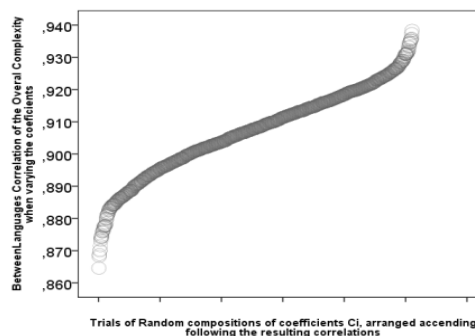

Fig. 9. Values of the between-languages correlation of the Overall Expressed Concepts when generating random compositions of coefficients in (2)

This manipulation applied to the statistical spaces captures the sensitivity of the between-languages correlation to the weights of the classes and, also, to the number and groupings of classes, since, when the coefficients of two or more classes are equal, they are fused. Thus, the languages' similarity is examined in quite a free condition – the classes can be randomly fused and the complexity of any class can be 2-8 times greater than the complexity of another class.

The plot of the obtained this way values for the between-languages correlation is given in Fig. 9. A correlation value close or bigger then the value 0.933 can be attained with a probability 0.011. This rejects the hypothesis that the between-languages correlation obtained by applying the model is gained by chance.

The statistical investigation of the corpora of child speech is in support of the model. The result suggests that the Concept Complexity develops in identical way in the two languages. The model proposes an explanation for this similarity. Additional investigation related to linguistics and semantics factors is needed to explain the differences in the distribution of the classes in the two languages.

This result supports the notion of the existence of a Universal Grammar based on inborn mechanisms for generating semantic representations of the world.

7. Discussion

The incontestable result of this study is its statistical finding: when parts-of-speech are spread into meaning-classes and these classes are considered with complexity loads following the Fibonacci sequence, language-acquisition by English and French infants can be discerned as one and the same process. The proposed model, based in reasoning related to optimization of entropy and energy-consumption and upon a schematic representation of information-treatment, can explain this statistical result. There are other recent studies that relate language phenomena with the Fibonacci sequence (e.g., [35]), investigating an isomorphism between the linguistic features of the Minimalist Program (the syntactic Merge-tree) and the formalism of quantum field theory. In addition, theoretical works (e.g., [1]) have appeared, discussing, in a combining framework, language, Fibonacci series, entropy and the second law of thermodynamics.

One question that arises is how, in more precise substantive terms, the Fibonacci sequence appears at this high level of the world's "phenomenological hierarchy". This work cannot give an answer to such a profound question. However, what is proposed in this Discussion section provides some paths of reasoning and possible directions for further investigation.

The model that has been described in this paper is grounded in perception, so one hypothesis is that the Fibonacci-like motif is related to some prominence in human perception. There are studies (e.g., [49]) and known facts indicating that the frequencies of the pitches (see **https://www.goldennumber.net/music/**) show Fibonacci ratios and the manner of constructing consonant chords deployed in the western music tradition (major and minor) equates with the Fibonacci-numbers. A recent analysis of EEG recordings of central auditory processing in new-borns (mean age 1.7 days) suggests [51] that the discrimination of both dissonant vs. consonant

and minor vs. major chords were processed differently from each other by the infant brain. It is still not clear whether this auditory discrimination relates to perinatal learning and cultural factors or to some universal perceptual frame. The polemic in the domain of music perception continues (see [7, 28]). Concerning the visual modality, studies in human vision have demonstrated the existence of a perceptual blindness for Fibonacci-like shapes. As suggested by E l l i o t t et al. [13], a bias, possibly due to reduced inter-neural synchronization when treating these specific geometric ratios, could be the reason for the perceptual effects of the golden section (often used in arts and architecture).

Such results from vision and audition give rise to the hypothesis that the Fibonacci-like pattern is involved one way or another in our biology and projected onto the higher levels of perceptual functioning. It is known that some major structural components of the cell cytoskeleton lattices follow Fibonacci geometry. It became clear also that the supra-molecular 3D structure of DNA is assembled in accordance with the Fibonacci sequence [36]. Such findings suggest a possible DNA-driven projection of the Fibonacci-like motif. The reasonable question at this point is: from where could such a projection come?

The fact that the Fibonacci pattern is often observed in living systems has focused the attention of many researchers from different domains (e.g., [53]). Surprising resemblances between plant cells and neurons have been discovered [2]. It has been demonstrated that there are many structural and functional similarities between the human microtubule cytoskeleton and the plant microtubule cytoskeleton [16]. Concerning energy, plant growth is such that plants always strive to maximize the quantity of their energy source. Plant growth and structure often follow Fibonacci ratios. From analysis of 557 different plant architectures, it was found [9] that plant growth is based on mechanisms which are typical for dendritic and axonal morphologies in the brain. The contemporary results show that dendritic encoding exists as an autonomous path of information transfer involved in specific plastic, computational, and cognitive processes (see [12, 47]). And, analysis of the dendrites of pyramidal cells in rat has revealed that the geometric mean lengths of exterior and interior links accord with the Fibonacci series [54].

Mathematical models of the dynamic processes of asymmetric division of cells and their renewal are beginning to explain how Fibonacci numbers appear in patterns of growth in nature [6].

With regard to entropy, through going deeper and deeper into the organization of matter, one can find, for example, that Golden Ratios imply maximum entropy in nuclear physics [31]. All of these findings suggest that there should exist principles of general character related to the Fibonacci sequence that are projected at the level of living matter and so influence the behaviour of organisms. The discovery of such principles is a matter of future efforts in several domains.

References

1. A u r u e s, T. The Fibonacci Sequence in Nature Implies Thermodynamic Maximum Entropy (New Developments of Generalized Entropies by Functional Analysis). 2013, 13.12.2017.
https://repository.kulib.kyoto-u.ac.jp/dspace/bitstream/2433/195148/1/1852-18.pdf.
2. B a l u š k a, F. Recent Surprising Similarities between Plant Cells and Neurons. – Plant Signalling & Behaviour, Vol. **5**, 2010, No 2, pp. 87-89.
3. B a r s a l o u, L. W. Situated Simulation in the Human Conceptual System. – Language and Cognitive Processes, Vol. **18**, 2003, No 516, pp. 513-562.
4. B a r s a l o u, L. W. Cognitively Plausible Theories of Concept Composition. – In: Y. Winter, J. A. Hampton, Eds. Compositionality and Concepts in Linguistics and Psychology. London, Springer Publishing, 2017, pp. 9-30.
5. B a r s a l o u, L. W. What Does Semantic Tiling of the Cortex Tell Us About Semantics? – In: Neuropsychologia, Vol. **105**, 2017, pp. 18-38.
6. B o m a n, B. M., et al. Why Do Fibonacci Numbers Appear in Patterns of Growth in Nature? – The Fibonacci Quarterly, Vol. **55**, 2017, No 5, pp. 30-41.
7. B o w l i n g, D. L., M. H o e s c h e l e, K. Z. G i l l, W. T. F i t c h. The Nature and Nurture of Musical Consonance. – Music Perception: An Interdisciplinary Journal, Vol. **35**, 2017, No 1, pp. 118-121.
8. C h o m s k y, N. The Minimalist Program. Cambridge, MA, MIT Press, 1995.
9. C o n n, A., et al. A Statistical Description of Plant Shoot Architecture. – Current Biology, Vol. **27**, 2017, No 14, pp. 2078-2088.
10. D a v i s, G. B., M. O l s o n. Management Information Systems: Conceptual Foundations, Methods and Development. New York, McGraw-Hill, 1985.
11. D e h g h a n i, M., et al. Decoding the Neural Representation of Story Meanings across Languages. 2017. Preprint DOI 10.17605/OSF.IO/QRPP3.
12. D o r o n, M. et al. Timed Synaptic Inhibition Shapes NMDA Spikes, Influencing Local Dendritic Processing and Global I/O Properties of Cortical Neurons. – Cell Reports, Vol. **21**, 2017, No 6, pp. 1550-1561.
13. E l l i o t t, M. A., J. K e l l y, J. F r i e d e l, J. B r o d s k y, P. M u l c a h y. The Golden Section as Optical Limitation. – PloS One, Vol. **10**, 2015, No 7, e0131045.
14. E r i k s s o n, J. E., et al. Neurocognitive Architecture of Working Memory. – Neuron, Vol. **88**, 2015, No 1, pp. 33-46.
15. D r e y e r, F. R., F. P u l v e r m ü l l e r. Abstract Semantics in the Motor System? – An Event-Related fMRI Study on Passive Reading of Semantic Word Categories Carrying Abstract Emotional and Mental Meaning. – Cortex, 2017.
https://doi.org/10.1016/j.cortex.2017.10.021
16. G a r d i n e r, J., J. M a r c. Arabidopsis Thaliana, a Plant Model Organism for the Neuronal Microtubule Cytoskeleton? – Journal of Experimental Botany, Vol. **62**, 2010, No 1, pp. 89-97.
17. G o l d i n-M e a d o w, S., C. Y a n g. Statistical Evidence That a Child Can Create a Combinatorial Linguistic System without External Linguistic Input: Implications for language Evolution. – Neuroscience & Biobehavioral Reviews. Corrected Proof 12.2017 (in Press).
http://www.sciencedirect.com/science/article/pii/S0149763416301105
18. H o r i b e, Y. An Entropy View of Fibonacci Trees. – The Fibonacci Quarterly, Vol. **20**, 1982, No 2, pp. 168-178.
19. H o r i b e, Y. Notes on Fibonacci Trees and Their Optimality. – The Fibonacci Quarterly, Vol. **21**, 1983, No 2, pp. 118-128.
20. H u t h, A. G., S. N i s h i m o t o, A. T. V u, J. L. G a l l a n t. A Continuous Semantic Space Describes the Representation of Thousands of Object and Action Categories Across the Human Brain. – Neuron, Vol. **76**, 2012, No 6, pp. 1210-1224.
21. H u t h, A. G., W. A. d e H e e r, T. vL. G r i f f i t h s, F. E. T h e u n i s s e n, J. L. G a l l a n t. Natural Speech Reveals the Semantic Maps that Tile Human Cerebral Cortex. – Nature, Vol. **532**, 2016, No 7600, pp. 453-458.

22. K i m, K., M. K. J o h n s o n. Activity in Ventromedial Prefrontal Cortex during Self-Related Processing: Positive Subjective Value or Personal Significance? – Social Cognitive and Affective Neuroscience, Vol. **10**, 2014, No 4, pp. 494-500.

23. L i u z z i, A. G., et al. Cross-Modal Representation of Spoken and Written Word Meaning in Left Pars Triangularis. – NeuroImage, Vol. **150**, 2017, pp. 292-307.

24. L l i n á s, R. R. Intrinsic Electrical Properties of Mammalian Neurons and CNS Function: A Historical Perspective. – Frontiers in Cellular Neuroscience, Vol. **8**, 2014.

25. L l i n á s, R. R., et al. Reconstruction of Human Brain Spontaneous Activity Based on Frequency-Pattern Analysis of Magnetoencephalography Data. – Frontiers in Neuroscience, Vol. **9**, 2015, No 373, pp. 1-8.

26. M a o, X., Y. W a n g, Y. W u, C. G u o. Self-Referential Information Alleviates Retrieval Inhibition of Directed Forgetting Effects – An ERP Evidence of Source Memory. – Frontiers in Behavioral Neuroscience, Vol. **11**, 2017, No 187.

27. J u s t, M. A, J. W a n g, V. C h e r k a s s k y. Neural Representations of the Concepts in Simple Sentences: Concept Activation Prediction and Context Effects. – NeuroImage, Vol. **157**, 2017, No 15, pp. 511-520.

28. M c D e r m o t t, J. H., A. F. S c h u l t z, E. A. U n d u r r a g a, R. A. G o d o y. Indifference to Dissonance in Native Amazonians Reveals Cultural Variation in Music Perception. – Nature, Vol. **535**, 2016, No 7613, pp. 547-550.

29. M o r a n, P., J. T o m m e r d a h l. A Case Study of Linguistic Isolation and Questions about Subsequent Language Support and Educational Provision in the United Kingdom. – In: The Linguistic Association of Canada and the United States, LACUS Forum 36: Mechanisms of Linguistic Behavior, 2011, pp. 229-240.

30. M o r i n, A., J. M i c h a u d. Self-Awareness and the Left Inferior Frontal Gyrus: Inner Speech Use during Self-Related Processing. – Brain Research Bulletin, Vol. **74**, 2007, No 6, pp. 387-396.

31. N a k a g a w a, Y. The Golden Ratios Which Imply Maximum Entropy in Nuclear Physics. Japan Fibonacci Association, 2017 (Accessed 15.12.2017).
http://www.zg.em-net.ne.jp/~aurues/triage/room1/MagicNumbers20170926.pdf

32. N o r t h o f f, G. From Emotions to Consciousness–A Neuro-Phenomenal and Neuro-Relational Approach. – Frontiers in Psychology, Vol. **3**, 2012.

33. N o r t h o f f, G., F. B e r m p o h l. Cortical Midline Structures and the Self. – Trends in Cognitive Sciences, Vol. **8**, 2004, No 3, pp. 102-107.

34. P e n d l, S. L., et al. Emergence of a Hierarchical Brain during Infancy Reflected by Stepwise Functional Connectivity. – Human Brain Mapping, Vol. **38**, 2017, No 5, pp. 2666-2682.

35. P i a t t e l l i-P a l m a r i n i, M., G. V i t i e l l o. Quantum Field Theory and the Linguistic Minimalist Program: A Remarkable Isomorphism. – Journal of Physics: Conference Series, Vol. **880**, 2017, No 1, pp. 12-16. IOP Publishing.

36. P i n k e r, S. The Bootstrapping Problem in Language Acquisition. – In: B. MacWhinney, Ed. Mechanisms of language Acquisition, Lawrence Erlbaum Associates. 1987, pp. 399-441.

37. P l o u x, S., et al. Structural Stability of Lexical Semantic Spaces: Nouns in Chinese and French. – arXiv preprint arXiv:1710.04173, 2017.

38. S e t h, A. K. Interoceptive Inference, Emotion, and the Embodied Self. – Trends in Cognitive Sciences, Vol. **17**, 2013, No 11, pp. 565-573.

39. S h a b a l k i n, I. P., E. Yu G r i g o r'e v a, M. V. G u d k o v a, P. I. S h a b a l k i n. Fibonacci Sequence and Supramolecular Structure of DNA. – Bulletin of Experimental Biology and Medicine, Vol. **161**, 2016, No 1, pp. 193-196.

40. S l a v o v a, V. Data Collection for Studying Language Acquisition. – In: Proc. of 12th International Conference on Computer Science and Education in C. Sc., 2016, pp. 105-113.

41. S l a v o v a, V. On Native Semantic Roles – Comparative Study Based on Data from Child Language-Acquisition of English and French. – International Journal of Cognitive Research in Science, Engineering and Education (IJCRSEE), Vol. **5**, 2017, No 2, pp. 1-18.

42. S l a v o v a, V., K. A t a n a s o v, A Generalized Net for Modeling Working Memory and Language Processing. – In: Proc. of the 7th International Conference Cognitive Modeling in linguistics, V. Text Processing and Cognitive Technologies, No 9, 2004, pp. 90-101.

43. S l a v o v a, V., A. S o s c h e n. A Fibonacci-Tree Model of Cognitive Processes Underlying Language Faculty. – In: Proc. of 3rd International Conference in Computer Science, NBU, University of Fulda, Boston University, 2007, pp. 196-205.
44. S l a v o v a, V., A. S o s c h e n. Experimental Support of Syntactic Computation Based on Semantic Merge of Concepts. – International Journal Information Theories & Applications, International Journal Information Technologies & Knowledge, Vol. **3**, 2009, pp. 5-23.
45. S l a v o v a, V., A. S o s c h e n. On Mental Representations: Language Structure and Meaning Revised. – International Journal of Information Theories & Applications, Vol. **2**, 2015, No 4, pp. 316-325.
46. S l a v o v a, V., A. S o s c h e n. Syntactic Operations – Modelling Language Faculty, International Journal of Information Theories & Applications, Vol. **2**, 2015, No 4, pp. 326-337.
47. S m i t h, S. L., I. S m i t h, T. B r a n c o, M. H ä u s s e r. Dendritic Spikes Enhance Stimulus Selectivity in Cortical Neurons in Vivo. – Nature, Vol. **503**, 2013, No 7474, pp. 115-120.
48. T a r d i f, T. Nouns Are Not Always Learned Before Verbs: Evidence from Mandarin Speakers Early Vocabularies. – Developmental Psychology, Vol. **32**, 1996, No 3, p. 492.
49. v a n G e n d, R. The Fibonacci Sequence and the Golden Ratio in Music. – Notes on Number Theory and Discrete Mathematics, Vol. **20**, 2014, No 1, pp. 72-77.
50. V e l m a n s, M. When Perception Becomes Conscious. – British Journal of Psychology, Vol. **90**, 1999, No 4, pp. 543-566.
51. V i r t a l a, P., M. H u o t i l a i n e n, E P a r t a n e n, V. F e l l m a n, M. T e r v a n i e m i. Newborn Infants' Auditory System is Sensitive to Western Music Chord Categories. – Frontiers in Psychology, Vol. **4**, 2013.
52. W a n g, J., V. C e r k a s s k y, M. A. J u s t. Predicting the Brain Activation Pattern Associated with the Propositional Content of a Sentence: Modeling Neural Representations of Events and States. – Human Brain Mapping, Vol. **38**, 2017, No 10, pp. 4865-4881.
53. W i l l e, J. J. Occurrence of Fibonacci Numbers in Development and Structure of Animal Forms: Phylogenetic Observations and Epigenetic Significance. – Natural Science, Vol. **4**, 2012, No 4, p. 216.
54. W o l d e n b e r g, M. J., M. P. O ' N e i l l, L. J. Q u a c k e n b u s h, R. J. P e n t n e y. Models for Growth, Decline and Regrowth of the Dendrites of Rat Purkinje Cells Induced from Magnitude and Link-Length Analysis. – Journal of Theoretical Biology, Vol. **162**, 1993, No 4, pp. 403-429.
55. Y a n g, Y., et al. Commonalities and Differences in the Neural Representations of English, Portuguese, and Mandarin Sentences: When Knowledge of the Brain-Language Mappings for Two Languages is Better than One. – Brain and Language, Vol. **175**, 2017, pp. 77-85.
56. CHILDES – The Child Language Component of the TalkBank – System for Sharing and Studying Conversational Interactions. Online Resource.
http://childes.psy.cmu.edu/