CREATIVITY STIMULATION IN CHAT CONVERSATIONS THROUGH MORPHOLOGICAL ANALYSIS

Daniela STAMATI[1], Mihai DASCĂLU[2], Ştefan TRĂUŞAN-MATU[3]

Computer Supported Collaborative Learning (CSCL) scenarios have emerged as viable alternatives to classic education in which technology plays the role of a facilitator. The choice of integrating creativity stimulation algorithms into chat conversations is derived from the ease of using such environments for brainstorming. Areas of interest are determined by applying General Morphological Analysis patterns built upon specific semantic models. Users are presented related concepts with those discussed which will ultimately lead to the adoption of new conversational patterns, thus increasing the rage of ideas and thinking scenarios. The results were ecouraging given the good feedback provided by 39 students.

Keywords: CSCL, creativity stimulation, chat analysis, morphological analysis

1. Introduction

Knowledge in general can be perceived as a network of concepts and their corresponding links. When users come across new information, they automatically connect or bridge it with previously known concepts [1]. Intrinsically, the complexity of the network is influenced by the amount of information retained by users, making possible the inference of new knowledge. Due to a potential lack of a path between ideas existing in our conscious thinking and latent ideas, people might fail to connect the pieces of a problem together. Therefore, the mere retention of information does not guarantee a successful outcome in terms of idea generation [2].

The main focus of our solution concentrates on helping students build a situation model which is a coherent mental representation of the studied information. [3]. On the micro level, students build a surface representation by linking key elements of the studied text together. On a macro level, the conceptual model is detached from the text, connecting the pieces of information with previously known concepts. It is immensely important to reach a coherent

[1] Eng., Dept.of Computer Science, University POLITEHNICA of Bucharest, Romania, e-mail: daniela.stamati5@gmail.com
[2] Lecturer, Dept.of Computer Science, University POLITEHNICA of Bucharest, Romania, e-mail: mihai.dascalu@cs.pub.ro
[3] Prof., Dept.of Computer Science, University POLITEHNICA of Bucharest, Romania, e-mail: stefan.trausan@cs.pub.ro

representation of a concept in order to fully understand it. Loose connections may lead to misinterpretation and possibly wrong decisions [4].

In this context, we find it appropriate to introduce idea generation techniques that concentrate on combining existing ideas and on reshaping reachable ones into new forms of comprehension. There are multiple ways to enhance and stimulate comprehension, but one of the most powerful and frequently used techniques consists of collaborative brainstorming [5]. By putting together people with different backgrounds and thinking patterns, we can introduce a conceptual diversity from which more ideas can emerge.

Computer Supported Collaborative Leaning (CSCL) is a newly adopted pedagogical approach concerning the way people can learn together with the help of computers [6]. As simple this statement may seem, it carries complex intricate theories build upon psychology, sociology and computer science. CSCL is developed on learning models which emphasize that knowledge is the result of a collaborative effort. The social order dictates the way we perceive things and it is also a nearly perfect environment of delivering information due to peoples' inclination of assimilating it faster in a social context. Assigning faces, events and a physical context to a concept often has a bigger impact that theoretical evidence.

The advantages of computer fostered creativity are mainly focused on its flexibility. Starting from a sound theoretical background from section 2, we introduce in section 3 the general morphological analysis (GMA) as the main model for stimulating creativity. While section 4 is mostly focused on the integrated semantic models from our discourse analysis framework - *ReaderBench* [7, 8] -, section 5 presents in detail the implemented GMA model applied on chat conversations, the obtained results and the validation experiment. The paper ends with conclusion and potential extensions of our computational model.

2. Theoretical background

The process of idea generation can be modeled to fit various patterns, each of them explaining this concept in a different way. Nevertheless, the common ground for all patterns is to create the proper set of concepts and underlying links for combining thoughts and assimilating information gradually, up until the moment the mind is ready to reach revelation.

Divergent and convergent thinking

Psychologists have shown that during idea generation there are two different and yet close approaches to get a new idea in our reach: divergent and convergent thinking [9]. *Divergent thinking* plays an important role in stimulating the autobiographical events. It combines the data stored in the semantic memory, which keeps track of facts, referred as knowledge, with data retrieved from the semantic memory, which is a record of our personal experiences. Even with the same degree of knowledge in a certain domain, subjects would suggest different solutions to a problem by filtering the information through episodic memory. In contrast, *convergent thinking* is centered on providing a single, goal-directed answer to a given problem.

The main difference between the two approaches relies in their linearity. We may associate divergent thinking with finding multiple solutions to a problem, while convergent thinking consists of combining facts in giving the answer to a question. However, convergent thinking plays a very important role in filtering and analysing ideas. Overall, these approaches are both met on an individual level, but also in collaborative brainstorming.

Idea generation stages

An active idea retains around the entire network of concepts familiar to the subject. Depending on how the network is navigated, new concepts can emerge. Disregarding the nature of the idea generation process, the following stages are encountered: preparation, incubation, illumination and verification [10]. *Preparation* consists of gathering intellectual resources involving planning, research and setting up the problem solving mind-set. This is the most time consuming operation that also involves the most cognitive effort. The next step, known as *incubation* does not involve any active intellectual effort, but it is crucial because it triggers subconscious thinking, when the brain creates the mandatory connections for an idea to emerge. It was observed that sometimes productivity can be enforced by starting several tasks and leaving them unfinished while we turn to others. *Illumination* is the culmination of previous cognitive efforts, but it is not a cognitive effort by itself since it occurs in a flash.

Following illumination, *verification* is used to validate the solution and to trigger a new iteration of the previous stages, if necessary. Each of these steps is mandatory, but sometimes they don't follow the same order. If there are not enough resources, a learner can go from incubation back to preparation, or sometimes seamlessly skip incubation because (s)he already has the necessary information in reach. It is important to understand that incubation still occurs, but it may not occur in the current iteration. Seeking a solution to a problem is a long-term process that often starts before being acquainted with the problem itself.

Dialogism and polyphony

As defined by Bakhtin [11], dialogism describes every level of expression from live conversations to complex cultural expressions as an ongoing inter-animation of voices in which new statements are built upon previous utterances and can predict anticipate responses. Extending the ideas introduced by Bakhtin for novels to conversations, we may say that the latter are modeled by three important components: polyphony, counterpoints and ventriloquism.

Firstly, *polyphony* refers to the multiple voices overlapping within a conversation. A voice is built on an idea and does not necessarily relate to a single participant. It could be generated by an utterance or even a word which develops its fingerprint during the conversation. Voices can be spread by a group of people and also can originate on an individual level. In correlation to polyphonic music in which voices may be played by a single human interpret, a single participant may spread several ideas [12]. When Bakhtin [13] introduced this concept, he described it as independent voices that play along the same melody. Bakhtin claims that polyphony is the only existing method that can lead someone to the absolute "*truth*", because only by fully understanding and accepting the whole set of associations around a concept one can definitely understand it: "*In order to understand, it is immensely important for the person who understands to be located outside the object of his or her creative understanding—in time, in space, in culture. For one cannot even really see one's own exterior and comprehend it as a whole, and no mirrors or photographs can help; our real exterior can be seen and understood only by other people, because they are located outside us in space, and because they are others*" [14]. In other words, it is sometimes important to detach ourselves from our own mind-set in order to view the entire landscape of a problem.

Generation of different ideas is often a result of *ventriloquism*, which is described as borrowing others' solutions and molding them into individual perspectives (ventriloquism is defined, in general, as speaking with another's voice). One can adopt a new idea if there is a trace of his/her own personality and experience in it. By accepting an idea, we later on filter it though our own personality and put to it an individual understanding.

The *counterpoint* is the technique used in polyphonic music to enable the achievement of superposing several melodic lines that each has its own 'personality' while a coherent, harmonious whole is kept. We can say, in the idea of our approach that counterpoint is a property of ideas to "sing" by their own pitches while a coherent goal is achieved. Every melodic line has a distinct character and bares with it the personality trace of the voice who uttered it. This determines the originality of every idea and facilitates creativity within the chat conversation. An important property of counterpoint is that it builds an antithesis of ideas, making it easier to compare and analyze them: *"In polyphony, several voices jointly construct a melody (or a story, or a potential solution in the textual-chat case) while other voices situate themselves on a differential position, identifying dissonances (unsound, rickety stories or solutions). This polyphonic game may eventually make clear the correct, sound solution."* [15]

3. General Morphological Analysis

General Morphological Analysis (GMA) [16] is a problem solving technique that supports dynamic sets of inputs and outputs of non-quantifiable problems relying on judgmental processes and internal consistency, rather than causality. GMA rather assists judgment than providing a final answer to a problem. The algorithm was developed by Fritz Zwicky [17] who applied it in studying astronomy and rocket propulsion systems. Today, it is used in a wide range of domains such as economy, political analysis and various aspects of futurology.

One of the biggest challenges in using this algorithm is to express complex real-world problems into the GMA model. The general idea is to identify inputs and outputs and to enforce rules on which inputs can lead to a corresponding output configuration. GMA builds problem outputs as sets of combined input parameters where each combination describes a different possible outcome. The approach begins with defining a set a parameters or dimensions of a problem and assigning each a relevant set of values. Inputs and outputs are modeled using a morphological box or a hypercube, also known as “Zwicky Box” [17]. Each of the dimensions represents the entire set of values for a corresponding parameter. The intersection of the parameters marks out a particular configuration of the problem. Naturally this provides a solution for modeling problems with a limited set of variables; hence a more flexible method of representation is needed. Therefore, in most cases for problems with a larger number of dimensions, a morphological field format representation is used (see Table 1). Note that the highlighted fields describe the similar potential configurations as in Zwicky's box [18].

Table 1

GMA Problem space

Parameter 1	Parameter 2	Parameter 3
P1.1	P2.1	P3.1
P1.2	P2.2	P3.2
P1.3	P2.3	P3.3
P1.4	P2.4	P3.4
P1.5	P2.5	P3.5
P1.6	P2.6	P3.6

Zwicky [18] also introduced the principle of contradiction and reduction in order to lower the number of formerly possible configurations. The main idea is to consider mutual incompatibilities between some parameter values within the initial set. For example, when building a social model, some configurations can be mutually exclusive due to legal regulations. There are three types of inconsistencies that can occur within the input set. The most common ones are logical inconsistencies, which describe parameters of a problem that cannot coexist. Normative constraints are based on regulations, on ethical or legal grounds, whereas empirical constraints describe highly improbable situations.

Cross consistency assessments are crucial not only for reducing the dimension of the problems, but also for delivering more accurate outcomes. Table 2 shows that there is no possible way for the parameter 1 to have the value P1.2, while the parameter 2 has the value P2.3. The same assumption can also be made for the other grayed out fields in the table.

Table 2

Cross Consistency Assessment

		Parameter 1				Parameter 2			
		P1.1	P1.2	P1.3	P1.4	P2.1	P2.2	P2.3	P2.4
Parameter 2	P2.1								
	P2.2								
	P2.3		~~X~~						
	P2.4								
Parameter 3	P3.1								
	P3.2				X				
	P3.3								
	P3.4	X							

The final step in GMA consists of the analysis of the possible outcomes. It provides a synthesis of well-defined configurations and dependences exposed in an unbiased way. The outcome set will contain only consistent combinations, leaving it in the end to the user to choose between them.

4. *ReaderBench*, a discourse analysis framework

ReaderBench [7, 8] is an environment that enables the assessment of a wide range of learners' productions and their manipulation by the teacher. It allows the assessment of three main textual features: cohesion-based assessment, reading strategies identification and textual complexity evaluation, which have been subject to empirical validations. *ReaderBench* can be also used to perform chat analysis in order to highlight important features of the conversation such as participant involvement, collaboration assessment, voice inter-animation and other dialogism related characteristics [7, 19].

In a nutshell, the integrated semantic models consist of Latent Semantic Analysis (LSA) [20], Latent Dirichlet Allocation (LDA) [21] and semantic distances [22] from WordNet – a lexicalized ontology [23], all later on described in detail. LSA and LDA both rely on several training sets for building their semantic spaces and both interpret texts as bags of words, disregarding the order of words. The advantage of these approaches is that the training sets can be easily altered if the focus of the research changes, which make them easily adaptable. The downside is the poor performance in detecting synonymy as synonyms are rarely collocated within a textual fragment, as well as polysemy because all word senses are collided within a singular written form. In order to support the previous cases, WordNet was integrated as it was built based on a taxonomy tree of synsets retaining a large set of semantic relationships between concepts. Moreover, all the three semantic similarity measures are integrated within an aggregated cohesion score [24].

Latent Semantic Analysis (LSA)

LSA [20] starts by building a Term-Document matrix that reflects the structure of each document used within the training corpora. The columns are the analyzed documents, while the lines are the concepts that occur in these texts. Term frequency - inverse document frequency (see Eq. 1) is afterwards applied on this matrix in order to improve the adequacy of the importance scores assigned per concept (e.g., stop words - "the", "at", "in", etc. - are assigned a lower rating due to their frequent use).

$$TF.IDF = \frac{N_{i,j}}{N_j} * \log \frac{N_{Doc}}{N_{Doc_{i,j}}} \tag{1}$$

Where:

$\boldsymbol{N_{i,j}}$ = the number of occurrences of the word “i” in document “j”
$\boldsymbol{N_j}$ = the number of words in the document "j"
$\boldsymbol{N_{Doc}}$ = the number of documents analysed
$\boldsymbol{N_{Doc_{i,j}}}$ = the number of documents in which the words "j" appears

The resulted matrix is considerably large, but sparse, which justifies the use of singular value decomposition. Thus a reduced dimensional representation of the matrix is found that emphasizes the strongest relationships, reduces the noise and infers new associations between words through generic concepts. Cosine similarity is used as the measure of relatedness between two vectors, either words or documents (see Eq. 2).

$$sim(c_1, c_2) = \cos(v_1, v_2) = \frac{v_1 * v_2}{||v_1|| * ||v_2||} \tag{2}$$

Where v_1 and v_2 are the vectors corresponding to the concepts c_1 and c_2 from the share matrix of LSA

Latent Dirichlet Allocation (LDA)

LDA [21] is a generative probabilistic model applied on text corpora. Each document is modeled as a mixture of underlying topics with corresponding weights. Therefore, in LDA texts are considered bags of words that can contain a limited set of topics. Each word inside the document can be assigned to various topics with a determined probability. One of the approaches of determining the topic distribution over a set of documents is using the collapsed Gibbs sampling which starts off by randomly assigning each word from the training corpora to one of the *K* imposed topics, whereas incremental re-assignations increase accuracy:

Jensen-Shannon divergence (see Eq. 3) is used to measure the similarity of two probability distributions. It provides the similarity measure of two concepts by using the LDA measurements of topic distributions.

$$JSD(P||Q) = \frac{1}{2}D(P||M) + \frac{1}{2}D(Q||M) \quad (3)$$

Where:

$\boldsymbol{M}$ = $\frac{1}{2}(P+Q)$

$\boldsymbol{P}$, $\boldsymbol{Q}$ = probability distributions

$\boldsymbol{D}$ = Kullback–Leibler divergence (see Eq. 4)

$$D(P||Q) = \sum_i \ln\left(\frac{P(i)}{Q(i)}\right) * P(i) \quad (4)$$

Semantic distances in lexicalized ontologies

WordNet [23] is a large lexical database of English developed by Princeton University. Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept. WordNet interlinks not only word forms, but groups together words based on their meanings. Each synset is linked to other synsets by a set of “conceptual relations”. These relations include hypernyms, hyponyms, meronyms, holonims, troponyms and entailments. Based on the obtained relationships among concept, multiple semantic distance measures can be applied in order to determine the relatedness between concepts (see Table 3).

Table 3

Semantic distances in WordNet

Name and reference	Formula	Description
Path length	$l(c1, c2)$	The shortest path between two concepts/synsets in the taxonomy tree.
Leacock-Chondorow [25]	$sim_{LC}(c_1, c_2) = \frac{-\log\,(len(c_{1,}c_2\,))}{2 * D}$	The path length is normalized by the overall depth *D* of the ontology.
Wu-Palmer [26]	$sim_{wp}(c_1, c_2) = \frac{2depth(lcs(c_{1,}c_2\,))}{depth(c_1) + depth(c_2)}$	Conceptual similarity is a scaled metric perceived in comparison to a global depth.

5. GMA implementation using *ReaderBench*

There are few implementations of the GMA method. Two commercial software systems are *MA/Carma*^{TM} (Computer-Aided Resource for Morphological Analysis) and its predecessor, *MA/Casper* (http://www.swemorph.com/). A recent implementation extends GMA with methods derived from it (closeness to ideal point, HDDM, Relevance trees) [27].

ReaderBench provided the bases for a new implementation of GMA. It offers support for extracting important concepts discussed within a chat conversation which were further on integrated in the problem space of GMA, next to the nearest neighbors in their semantic networks (see right sidebar from Fig. 1 in which the user can select the topics of the conversation (s)he wishes to further expand upon).

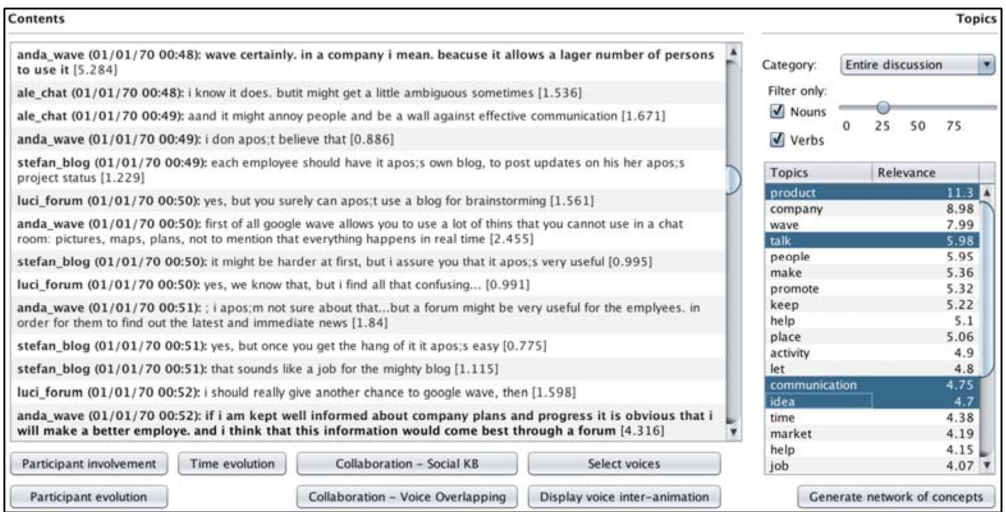

Fig. 1 *ReaderBench* view of a conversation

The entries in the problem space were then compared in order to determine the concepts which had a high semantic cohesion and originated from different words in the input set. These pairs of words defined the cross consistency assessments that ultimately defined the solution space. Concepts in the output set had a high connectivity to the conversation, but were not addressed yet.

The results further presented were built based on the analysis of an online conversation about the adoption of the most suitable brainstorming environment for a company. After extracting the most relevant topics of the conversation, the user was asked to select the topics (s)he would like to investigate further. Supposedly the topics “tell”, “answer”, “question” and “talk” were selected, the algorithm would first present the closest concepts to the ones selected which also shape the problem space of GMA (see Fig 2).

tell	answer	question	talk
narrate	question	answer	mouth
say	reply	enquiry	speak
state	response	interpretive	utter
	solution	interrogation	
		query	

Fig. 2 Initial Problem Space

Table 4 presents the closest concepts to the previous entry set (i.e., “tell”, “answer”, “talk” and “question”) and corresponding degree of similarity in order to argue for the concepts presented in Fig. 2. The proposed concepts are extracted with *ReaderBench* by determining the words with the highest degree of similarity to the selected concepts - synonyms from WordNet and *k* nearest neighbors from LSA and LDA semantic spaces (*k* has been experimentally set to 5).

Table 4

Similarities example

Concept	Determined similarities	Accuracy level
Tell	Narrate	100%
	Say	100%
	State	100%
Answer	Question	90%
	Reply	100%
	Response	100%
	Solution	100%
Talk	Mouth (v)	100%
	Speak	100%
	Utter	100%
Question	Answer	90%
	Enquiry	100%
	Interpretive	100%
	Interrogation	100%
	Query	100%

By using the weighted average of the similarity measures of Jensen-Shannon (LDA), Leacock-Chondorow (WordNet) and cosine similarity (LSA), the cross consistency assessments were determined. Only entries with a similarity measure above 0.6 are considered eligible for entering the cross consistency assessments set (see Table 5).

Table 5

Similarity measures for cross consistency assesment

Concept 1	Concept 2	Similarity measure
Narrate	Mouth	.66
Narrate	Speak	.70
Say	Answer	.75
Say	Speak	.70
Say	Utter	.84
State	Answer	.72
State	Utter	.76
Answer	Speak	.60
Answer	Utter	.75
Question	Enquiry	.81
Question	Query	.84
Reply	Query	.72
Answer	Speak	.61
Answer	Utter	.75

Contrary to the original GMA approach, entries that are highly compatible with each other are marked out, i.e. correlated concepts. The correlations are determined automatically (see Fig 3), but the user is free to alter the results by this modifying the dependencies in the solution space. By highlighting the correlated concepts in the entry set, tight relationships between initial topics are observed. These relations will further enable the possibility of finding learning materials close to the concerns of the conversation, but which still bring a degree of novelty.

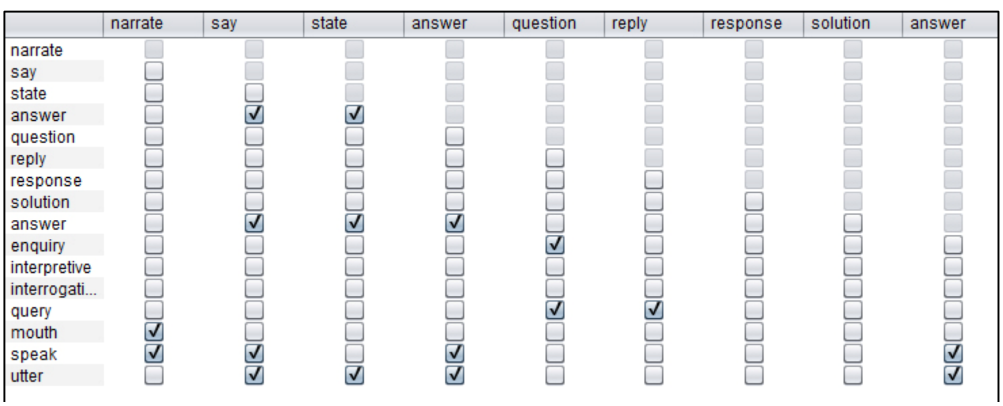

Fig. 3 Partial view of a cross-consistency assesment depicting all relevant connections within the solution space

Fig. 4 displays one of the possible outcomes in the solution space. In this case, the user is presented with the option to investigate two more concepts related to the one selected (narrate). By triggering a Wikipedia search the program will suggest articles concerning these topics.

tell	answer	question	talk
narrate	question	answer	mouth
say	reply	enquiry	speak
state	response	interpretive	utter
	solution	interrogation	
		query	

Fig. 4 Solution space

In order to evaluate the adequacy of our approach, we ran a series of surveys on 39 undergraduate students in our university regarding the adoption of creativity stimulation in e-learning solutions. It was found that most subjects use collaborative learning to some extent. 54% of the students prefer text driven tools in the preparation phase, while the rest prefer collaborative learning with a tutor or peer involvement, while the majority (89%) would like to have their learning results validated. Overall, we have come across a high acceptance degree since most students use or have been using e-learning software and have shown a great interest in testing a more adaptable solution that would be able to meet their needs as individuals. Moreover, students stated that they often study in groups and have high productivity rates during these sessions.

One of the most interesting parts of the surveys consisted of the direct feedback and concerns coming from respondents such as: “What if alternatively to CSCL, ‘Gamified Computer Supported Collaborative Learning’ was developed?”, “Will computer supported learning be more accurate if the undergone experiments are observed/monitored by a tutor?”, “Will a computer be able to interpret the human speech accurately enough in order to make valuable creativity suggestions?”.

6. Conclusions and further development

Today technology becomes a necessity and new types of learners are ready to embrace the newly developed tools. This brings big challenges, but also limitless possibilities of experiencing new learning strategies and accommodating the learning environment to meet the needs of groups, as well as individuals.

Beyond acting as a facilitator, Computer Supported Collaborative Learning helps students learn in a relaxed environment that is sensitive to their needs. By combining the current computational power with the stimulating environment of brainstorming, a powerful solution for fostering human creativity through CSCL is created. Overall, we can consider our model a joint and balanced approach, both personal through the use of manual annotations, as well as automatic through the use of natural language processing tools, in which the constituent elements catalyze one another.

In terms of future developments, we envision developing comprehensive educational experiments that will make use of the implemented mechanisms.

Acknowledgments

The work presented in this paper was partially funded by the Sectorial Operational Programme Human Resources Development 2007-2013 of the Ministry of European Funds through the Financial Agreement POSDRU/159/ 1.5/S/134398.

R E F E R E N C E S

[1]. *W. Kintsch, T. A. van Dijk*, "Toward a model of text comprehension and production", in Psychological Review, **vol. 85**, no. 5, 1978, pp. 363–394.
[2]. *G. Militaru*, "A critical evaluation of innovation and imitation processes: A conceptual approach", in Scientific Bulletin, University Politehnica of Bucharest, Series C, **vol. 73**, no. 4, 2011, pp. 273–286.
[3]. *T. A. van Dijk, W. Kintsch*, Strategies of discourse comprehension, Academic Press, New York, NY, 1983.
[4]. *I. Tapiero*, Situation models and levels of coherence, Erlbaum, Mahwah, NJ, 2007.
[5]. *S. Trausan-Matu*, "Computer support for creativity in small groups using chats", in Annals of the Academy of Romanian Scientists, Series on Science and Technology of Information, **vol. 3**, no. 2, 2010, pp. 81–90.
[6]. *G. Stahl*, Group cognition. Computer support for building collaborative knowledge, MIT Press, Cambridge, MA, 2006.
[7]. *M. Dascalu*, Analyzing discourse and text complexity for learning and collaborating, Studies in Computational Intelligence, Springer, Switzerland, 2014.
[8]. *M. Dascalu, P. Dessus, M. Bianco, S. Trausan-Matu and A. Nardy*, Mining texts, learners productions and strategies with ReaderBench, in Educational Data Mining: Applications and Trends, A. Peña-Ayala Ed. Springer, Switzerland, 335–377, 2014.
[9]. *K. D. Vohs, R. F. Baumeister*, "Handbook of Self-Regulation: Research, theory, and applications", The Guildford Press, New York, NY & London, 2011.
[10]. *G. Wallas*, The Art of Thought, Jonathan Cape, London, UK, 1926.
[11]. *M. M. Bakhtin*, Problems of Dostoevsky's poetics, University of Minnesota Press, Minneapolis, 1984.
[12]. *S. Trausan-Matu*, "Polyphonic Design, Conduct, Experience, and Evaluation in CSCL Chats", in Annals of the Academy of Romanian Scientists, Series on Science and Technology of Information, **vol. 7**, no. 2, 2014, pp. 21–34.
[13]. *M. M. Bakhtin*, The dialogic imagination: Four essays, The University of Texas Press, Austin and London, 1981.
[14]. *M. M. Bakhtin*, Speech genres and other late essays, University of Texas, Austin, 1986.
[15]. *S. Trausan-Matu, G. Stahl and J. Sarmiento*, "Polyphonic Support for Collaborative Learning", in proceedings of the Groupware: Design, Implementation, and Use, 12th International Workshop (CRIWG 2006), Medina del Campo, Spain, Springer, pp. 132–139, 2006.

[16]. *T. Ritchey*, General Morphological Analysis: A general method for non-quantified modeling. Swedish Morphological Society, 1998.
[17]. *T. Ritchey*, "Fritz Zwicky, Morphologie and Policy Analysis", in proceedings of the 16th EURO Conference on Operational Analysis, Brussels, 1998.
[18]. *F. Zwicky*, Discovery, Invention, Research - Through the Morphological Approach, The Macmillian Company, Toronto, 1969.
[19]. *M. Dascalu, S. Trausan-Matu and P. Dessus*, "Validating the Automated Assessment of Participation and of Collaboration in Chat Conversations", in proceedings of the 12th Int. Conf. on Intelligent Tutoring Systems (ITS 2014), Honolulu, USA, Springer, pp. 230–235, 2014.
[20]. *T. K. Landauer, S. T. Dumais*, "A solution to Plato's problem: the Latent Semantic Analysis theory of acquisition, induction and representation of knowledge", in Psychological Review, **vol. 104**, no. 2, 1997, pp. 211–240.
[21]. *D. M. Blei, A. Y. Ng and M. I. Jordan*, "Latent Dirichlet Allocation", in Journal of Machine Learning Research, **vol. 3**, no. 4-5, 2003, pp. 993–1022.
[22]. *A. Budanitsky, G. Hirst*, "Evaluating WordNet-based Measures of Lexical Semantic Relatedness", in Computational Linguistics, **vol. 32**, no. 1, 2006, pp. 13–47.
[23]. *G. A. Miller*, "WordNet: A lexical database for English", in Communications of the ACM, **vol. 38**, no. 11, 1995, pp. 39–41.
[24]. *M. Dascalu, P. Dessus, S. Trausan-Matu, M. Bianco and A. Nardy*, "ReaderBench, an environment for analyzing text complexity and reading strategies", in proceedings of the 16th Int. Conf. on Artificial Intelligence in Education (AIED 2013), Memphis, USA, Springer, pp. 379–388, 2013.
[25]. *C. Leacock, M. Chodorow*, Combining local context and WordNet similarity for wordsense identification, in WordNet: An electronic lexical database, C. Fellbaum Ed. MIT Press, Cambridge, MA, 265–283, 1998.
[26]. *Z. Wu, M. Palmer*, "Verb semantics and lexical selection", in proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics, ACL '94, New Mexico, USA, ACL, pp. 133–138, 1994.
[27]. *A. Oprisan, S. Trausan-Matu*, "Creativity Stimulation Tool", in Annals of the Academy of Romanian Scientists, Series on Science and Technology of Information, **vol. 6**, no. 1, 2013, pp. 63–83.