

A Polyphonic Model, Analysis Method and Computer Support Tools for the Analysis of Socially-Built Discourse

Ștefan TRĂUȘAN-MATU

“Politehnica” University of Bucharest

E-mail: stefan.trausan@cs.pub.ro

Abstract. The paper presents the polyphonic model of discourse, an associated analysis method and computer support tools based on them, following a socio-cultural approach and focusing on discourse built socially, for example in inter-ethnic debates. Natural Language Processing and Social Network Analysis methods were used for analyzing discourse in order to provide visualizations, feedback, and proposing conducting inter-animations toward a equilibrium-based polyphonic weaving.

1. Introduction

Many social debates are reflected in written texts, especially in journals, newspapers, either on paper or online, allowing the usage of Natural Language Processing (NLP) tools for their analysis. Even face-to-face or video (e.g. television) debates may be transcribed in a textual format, therefore being available also for an automated analysis. However, the analysis of the threading of ideas in multiple documents, of inter-textuality [13], and especially of the explicit or implicit debates and dialogues is yet at beginning in NLP, especially due to the difficulty of the involved tasks and the limitations of the NLP technology [29].

Nevertheless, text mining technology, a NLP approach mainly based on statistical learning techniques [11] [16] offers facilities for knowledge extraction even from a large number of documents, which may be applied for analyzing the construction of social discourse. Sentiment analysis (or opinion mining), a recent sub-domain of text mining may be used for statistically analyzing peoples' attitudes starting from texts belonging

to the above enumerated media types. Time series analysis may also be performed in combination with text mining [1].

Usually there is a distinction between the discourse models of conversations (spoken, transcribed or written in instant messenger chat) and written text in books, essays or journal papers. From a shallow perspective, the former involves in general two speakers which enter into a dialogue, debating around a given subject, even “fighting” sometimes to take the floor or to impose an idea. In the second case, in a simplistic (but common) analysis, a “line of ideas” is searched. However, as Mikhail Bakhtin emphasized, everything may be considered as a dialog, including novels, essays, and even words, which may be seen as being affected by the echoes of past utterances [2]. In his vision, not only the physical voices of the participants are present in language-based phenomena of communication. In a generalized sense, many other voices may be identified in any text, in face-to-face dialogs, or in its written transcription, and even in our thoughts: “the voices of others become woven into what we say, write, and think” [12]. This is similar to a ventriloquism (someone speaking with another person’s voice) process [2] [3] or heteroglossia, defined as “another’s speech in another’s language, serving to express authorial intentions but in a refracted way” [2] [23].

The analysis of socially-built discourse and especially of the debates around sensible issues like inter-ethnic relations is very well suited for an analysis considering multivoiceness, heteroglossia and ventriloquism. Therefore, a model and analysis tools based on this point of view is beneficial. A source of a metaphor for such a model is music and, in particular, the polyphonic music, in which a number of melodic lines (or voices) coexist and inter-animate according to some (counterpoint) rules along longitudinal and transversal dimensions, as we will discuss later: “When there is more than one independent melodic line happening at the same time in a piece of music, we say that the music is contrapuntal. The independent melodic lines are called counterpoint. The music that is made up of counterpoint can also be called polyphony, or one can say that the music is polyphonic or speak of the polyphonic texture of the music.” [18]. Starting from the polyphony metaphor we have constructed a model of discourse that was used for developing NLP a series of dialog analysis tools.

The paper continues with a section that introduces the polyphonic model of discourse. The third section presents details of the steps of the polyphonic analysis method and of the implemented support tools. The last section introduces conclusions and future directions.

2. The polyphonic discourse model

Starting from the polyphonic perspective in linguistics, initiated by [2] [3], we introduced a polyphonic model for analyzing discourse and a polyphonic method for analyzing discourse in texts. The polyphonic model is well suited for the analysis of the interethnic discourse, in which we should start from the premise that at least two voices may be present and any discourse, even an essay contains multiple voices. Based on the polyphonic model and method we have developed software tools for learning analytics [24] [23] [25] [8] [9], creativity fostering [20], news analysis [1], intertextuality detection [10] and speech genres detection [22].

The polyphonic discourse model was presented in detail in several other papers [26] [28] [23]. In this section we will describe its main ideas and components, emphasizing how it may be applied for analyzing social debates.

2.1. Discourse

We consider discourse as a socially achieved artifact: “discourse lives on the boundary between its own context and another, alien, context” [2]. It is mainly mediated in different language-based ways, either written or spoken communication: newspaper or journal articles, scientific papers, books, discussions on electronic forums, social networks, instant messenger (chat), interviews, and speeches at meetings or on television or on the web (e.g. youtube). However, it may be mediated also by non-verbal acts [21].

In Natural Language Processing is considered that discourse is different in the cases of text (for example, essays) and conversations [11]. Moreover, conversations are usually regarded only as phone dialogs, including only two interlocutors. A main metaphor used in conversations is “taking the floor”, that means that they are analyzed in the direction of identifying one line of discussion at a time, idea that is similar to (if not derived from) the case of discourse analysis in texts (essay-like), in which also a narrative line or idea thread is searched.

An unitary discourse model for facilitating its analysis and the implementation of analysis systems is beneficial for tackling the complexity of the phenomena. In this paper we will consider only textual mediation and processing using NLP technology. However, we see discourse in an extended way, as having an important holist (gestalt) dimension. We start from a musical metaphor for including subjective aspects like musicality (need of an esthetic feature), interestingness (not boring), coherence and structure. In the computational linguistics case, these correspond to opinion mining issues, cohesion/coherence, complexity and narrativity.

2.2. Voices and utterances

The polyphonic model is constructed around the idea of the “voice” concept, which is central also in Bakhtin’s writings [2] [3]. Voices manifest their existence by emitting utterances, by entering in dialog with other voices and by their echoes: “dialogic orientation of discourse is a phenomenon that is, of course, a property of any discourse. On all its various routes towards the object, in all its directions, the word encounters an alien word, and cannot help encountering it in a living, tension-filled interaction. Only the mythical Adam, who approached a virginal and as yet verbally unqualified world with the first word, could really have escaped from start to finish this dialogic inter-orientation with the alien word that occurs in the object. Concrete historical human discourse does not have this privilege: it can deviate from such inter-orientation only on a conditional basis and only to a certain degree.” [2].

In a generalized sense, following Bakhtin’s perspective, we consider that an utterance may range from a word or an idea to a reply in a chat or even to an entire book. Words play a essential role in this framework: “The idea is a living event ... In this

respect the idea resembles the word, with which it forms a dialogical unity. Like the word, the idea wants to be heard, understood and “answered” by other voices, from other positions. Like the word, the idea is by nature dialogical” [3]. As we said before, we have extended the scope of utterances also to non-verbal acts (Trausan-Matu 2003).

An utterance is produced by a voice, but it may contain any number of voices, in a kind of echoing or ventriloquism phenomena. For example, Bakhtin says that every word is also partly other’s word [3].

We do not see a voice as associated to a single human participant in a dialog. This is also the case in music, where, for example, the same instrumentalist may play several different voices at the same time at piano, violin, etc. At the opposite side, multiple instrumentalists in an orchestra, grouped in sections (violin I, violin II, cello, etc.) usually play a single voice, distinct from other sections. In a social context we may say that a person may include several voices in what s/he utters and, on the side, a group of persons may express a common voice.

Nr	Ref	Time	User	Text
17		10.26.25	tim	You discussed about a topic separation
18	15	10.26.37	adrian	First of all, the reply method is cumbersome
19	17	10.26.50	john	yes.. because we did not like the way the topics were presented in concert chat
20	18	10.26.56	john	yes !!
21	20	10.27.04	john	i hate double-clicking !
22	20	10.27.18	tim	and how can we find topics ?
23	18	10.27.26	adrian	What bothers me is the linear presentation of the discussin
24	23	10.27.43	john	Yep
25	18	10.27.46	adrian	and double-clicking too
26		10.27.54	tim	You mean u want something like a chat forum ?
27	24	10.27.58	john	and the reply -to facility is supposed to help you
28	18	10.28.15	adrian	i'd like a tree presentation more
29	18	10.28.38	adrian	or maybe multiple chat columns, for each chat sub-thread
30	27	10.28.58	john	but it is really difficult to use in real-time, because there are so many topics discussed which intertwine each other
31	28	10.29.18	john	i subscribe to a tree-like presentation form
32	P 30	10.29.20	adrian	yes, that's why a clear separation of topics is needed
33	31	10.29.47	adrian	this is easy to implement, no problem here :)
34	30	10.29.49	tim	You need also a clever visual representation
35	30	10.30.05	tim	you'll need also a clever visual interface
36		10.30.22	tim	Who decides the topics ?
37	33	10.30.33	john	i suppose you are referring to the visual representation , right ?
38	37	10.30.45	john	What i would like is a clever way to separate the topics :)
39	38	10.30.59	john	not just doing ot myself, manually
40	37	10.31.00	adrian	Yeah
41	39	10.31.44	adrian	When you start a new thread (a new message, non-related to other message), the app can assume a new topic
42	39	10.31.46	john	i would like the application to be able to detect w topic change all by itself
43	42	10.32.01	tim	That right

Fig. 1. In addition to obvious voices (tim, adrian and john), repeated words (e.g. topics, presentation, double clicking) become also voices.

In our theoretical framework we consider a voice in a general sense, not reduced to the physical, acoustical dimension. We rather consider it as a distinct, differential position with persistence and interference with other voices. We consider that, for

example, an utterance, that means a word, especially if it is repeated (see Fig. 1), an idea, a reply, a book or even a non-verbal act, as mentioned above) may become a distinct voice through its echoes and influences in the subsequent utterances. Of course that we consider as voices also the participants to a conversation or even groups of persons (for example, minorities), because they represent distinct positions, with persistence and that interfere with other voices, be they other persons, groups or voices in a general sense (for example, ideas or replies that influences them).

A voice should be viewed differentially, in contrast to other voices played in the same time. Differential positions may be generated also by echoing (the superimposition of the echoes of other voices), by ventriloquism, by mimicking other voices and therefore contrasting to their actual particularities. The differential positioning of voices may enter in a kind of a game or in a counterpointal framework, specific to polyphonic music, inducing inter-animation phenomena: “The deconstructivist attack (...) – according to which only the difference between difference and unity as an emphatic difference (and not as a return to unity) can act as the basis of a differential theory (which dialectic merely claims to be) is the methodical point of departure for the distinction between polyphony and non-polyphony.” [15].

In addition to the differential feature of voices, manifested on a transversal direction, another basic feature is their persistence in time, the fact that they have a longitudinal dimension, a history. Through their specific, particular features, voices may get differential positions displaying sometimes a kind of personhood.

2.3. Inter-animation and polyphony

The most important aspect related to voices is their interaction on their two dimensions: transversal and longitudinal, entering in an inter-animation phenomenon. In both the transversal and longitudinal dimensions voices may be dissonant (differential, divergent) or consonant (on a conjunctive position).

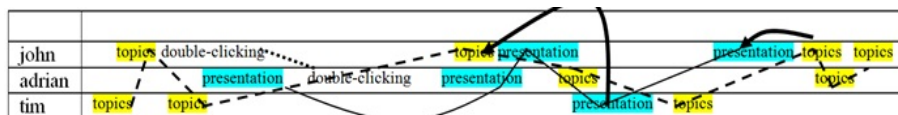


Fig. 2. The interanimation of the voices emphasized in Fig. 1. The thick arrows indicate inter-animation patterns.

Similarly with Bakhtin, we remark that centrifugal and centripetal forces concur in the development of voices, in the direction of divergence and respectively convergence, in an action-reaction phenomenon. For example, in music, moments of dissonance may occur, but they induce the expectation of consonances. However, at a social scale this dynamic equilibrium might not be obtained without an intervening force (voice).

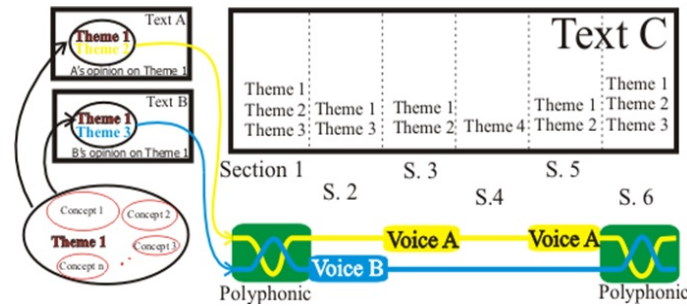


Fig. 3. Polyphonic weaving in inter-textuality [10].

The inter-animating voices along the transversal-longitudinal directions facing the dissonance-consonance game restricted by the centrifugal-centripetal forces may enter into a kind of polyphonic texture. Mikhail Bakhtin has analyzed this polyphonic weaving in Dostoevsky's novels, characterizing it as "a plurality of independent and unmerged voices and consciousnesses" [3].

In each debate there are at least two opposite themes, two voices put face-to-face, similar to the note-counter-note in the musical polyphonic counterpoint.

Some of the interactions which appear in the polyphonic weaving may be abstracted in classes of inter-animation patterns in which voices co-occur or are in a cause-effect relation. The identification of such patterns may be the starting point for discourse analysis [27].

3. The polyphonic analysis method and the computer-based support

The polyphonic model presented in the previous section was the basis for a discourse analysis method and of a series of computer-based systems for several purposes. Domains in which the model and the method were used are Computer-Supported Collaborative Learning [26] [27] [28], creativity fostering [20], analysis of text complexity [9], discourse analysis [26] [23], automatic summarization [27] [6], inter-textuality detection [10].

The polyphonic analysis method is trying to uncover the polyphonic weaving of various types of socially constructed discourses, including non-verbal utterances [21]. In this aim it follows several steps, not all of them being mandatory, with the exception of the first two. For the majority of the following seven steps computer-based systems were developed for providing (semi-)automatic assistance.

1. Identify voices
2. Detect inter-animation
3. Rhythm analysis
4. Detect variation and modulation

5. Analyze involvement and strength
6. Coherence and complexity evaluation
7. Feedback provision and polyphony conducting

Voices' identification may be different depending on the nature of utterances. In the case of textual utterances (documents, conversations or discussion forums), methods that may be applied are those specific in text mining: the identification of concepts using Tf*Idf [16] or the identification of semantic groupings: Latent Semantic Analysis - LSA - [14], Latent Dirichlet Allocation - LDA - [4] or lexical chains construction [11]. All the computer systems which will be presented below use one or more of these methods. For example, PolyCAFe [25] and ReaderBench [7] use Tf*Idf and LSA. ReaderBench uses also LDA and lexical chains.

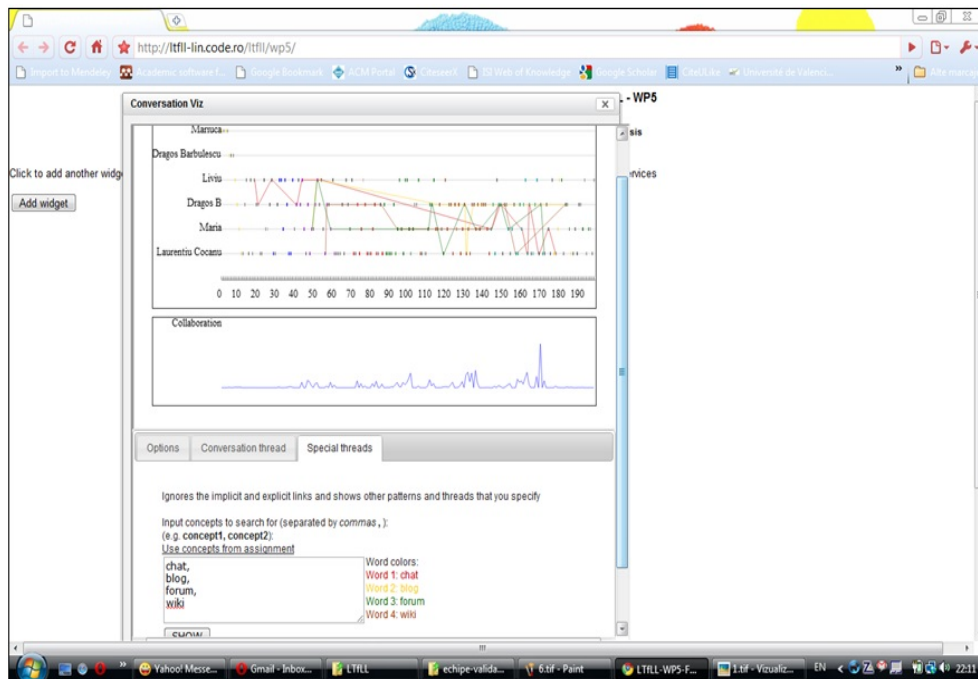


Fig. 4. Visualization of the voices (of the participants – horizontal places small rectangles for emitted utterances –, and of four repeated words associated with different colors – and of the strength of the collaborative strength of the utterances in a conversation through PolyCAFe.

Inter-animation detection is implemented in PolyCAFe and ReaderBench based on NLP techniques for the identification of speech and argumentation acts [11]. Cue-phrases identification and machine learning techniques were used in this aim. Under

development is a system using time series analysis for detecting inter-animation patterns among voices in series of news articles [1].

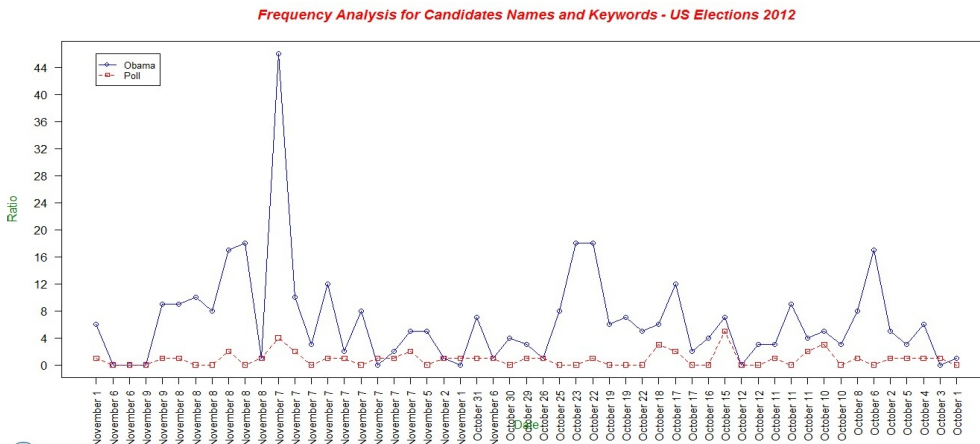


Fig. 5. Time-series analysis of a news articles related to US Elections in 2012 [1].

Visualization facilities for *Rhythm analysis* in chat conversations are provided in PolyCAFe. Rhythm is also analyzed and visualized in a system implemented by Chiru et al. [5].

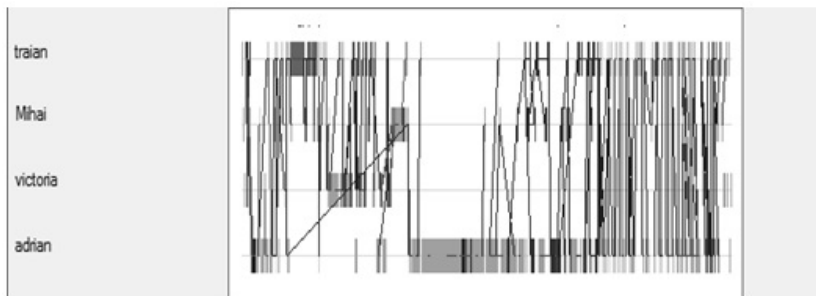


Fig. 6. Visualization with PolyCAFe of a rhythmic segment in the second part of a conversation.

The *analysis of degrees of involvement and strength* of utterances and voices was performed using a combination of techniques: Social Networks Analysis [8], semantics (the content of important concepts) and a polyphony-specific technique: the analysis of the voices echoes [6]. In addition to these methods, for the *[Coherence and complexity evaluation]*, LSA is also used in ReaderBench [7]. *Feedback* is provided by PolyCAFe starting from the results obtained in the analysis of the degrees of involvement and of the strengths of the utterances.

- [3] BAKHTIN M., *Problems of Dostoevsky's Poetics*, Ed. and trans. C. Emerson, Theory and History of Literature Series, vol. 8, University of Minnesota Press, Minneapolis, 1984.
- [4] BLEI D.M., NG A.Y., JORDAN M.I., *Latent Dirichlet Allocation*, Journal of Machine Learning Research, **3**(4–5), pp. 993–1022, 2003.
- [5] CHIRU C.-G., COJOCARU V., TRĂUȘAN-MATU Ș., REBEDEA T., MIHĂILĂ D., *Repetition and Rhythmicity Based Assessment Model for Chat Conversations*, Marzena Kryszkiewicz, Henryk Rybinski, Andrzej Skowron, Zbigniew W. Ras (Eds.): *Foundations of Intelligent Systems – 19th International Symposium, ISMIS 2011*, Warsaw, Poland, June 28–30, 2011. *Lecture Notes in Computer Science*, 6804, Springer, pp. 513–522, 2011.
- [6] DASCALU M., *Analysing Discourse and Text Complexity for Learning and Collaborating*, PhD Thesis, 2013.
- [7] DASCALU M., DESSUS P., TRAUSAN-MATU S., BIANCO M., NARDY A., *ReaderBench, an Environment for Analyzing Text Complexity and Reading Strategies*, 16th Int. Conf. on Artificial Intelligence in Education (AIED 2013), Memphis, USA, 2013.
- [8] DASCALU M., TRAUSAN-MATU S., *Ch.A.M.P. Sistem pentru evaluarea si modelarea contributiei participantilor la un Chat*, Revista Romana de Interactiune Om-Calculator, **2**(2), pp. 131–146, 2009.
- [9] DASCALU M., TRAUSAN-MATU S., DESSUS P., *Towards an integrated approach for evaluating textual complexity for learning purposes*, 11th Int. Conf. in Advances in Web-Based Learning (ICWL 2012), Sinaia, Romania, 2012.
- [10] GHIBAN I., TRAUSAN-MATU S., *Network Based Analysis of Intertextual Relations*, *Proceedings of WorldCIST'13 – The 2013 World Conference on Information Systems and Technologies*, 2013, Advances in Intelligent Systems and Computing, 206, Springer, pp. 753–762, 2013.
- [11] JURAFSKY D., MARTIN J.H., *Speech and Language Processing, An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*, Second Edition, 2009.
- [12] KOSCHMANN T., *Toward a Dialogic Theory of Learning: Bakhtin's Contribution to Understanding Learning in Settings of Collaboration*, in C.Hoadley and J. Roschelle (eds.), *Proceedings of the Computer Support for Collaborative Learning Conference*, Stanford, Laurence Erlbaum Associates, 1999.
- [13] KRISTEVA J., *Desire in language: a semiotic approach to literature and art*, Columbia University Press, 1980.
- [14] LANDAUER T.K., FOLTZ P.W., LAHAM D., *An introduction to Latent Semantic Analysis*, Discourse Processes, **25**(2–3), pp. 259–284, 1998.
- [15] MAHNKOPF C.S., *Theory of Polyphony*, in Mahnkopf C.S., Cox F., Schurig W. (eds), *Polyphony and Complexity*, Wolke Verlags Gmbh., Hofheim, Germany, 2002.
- [16] MANNING C., SCHUTZE H., *Foundations of Statistical Natural Language Processing*, MIT Press: Cambridge Mass, 1999.
- [17] MUSAT C., VELCIN J., TRAUSAN-MATU S., RIZOIU M.-A., *Improving Topic Evaluation Using Conceptual Knowledge*, *Proceedings of IJCAI 2011*, pp. 1866–1871, 2010.
- [18] Polyphony (2005) <http://cnx.rice.edu/content/m11634/latest/>, retrieved on 4th May, 2005.

- [19] TRAUSAN-MATU S., *Dialogistica lui Bahtin: polifonie, ventrilogism, carnavalesc, inter-animare*, in Constandache G.G. (Ed.) *Comunicarea - sugestie si influenta*, Editura Sigma, pp. 39–44, 2009.
- [20] TRAUSAN-MATU S., *Computer Support for Creativity in Small Groups using chats*, *Annals of the Academy of Romanian Scientists, Series on Science and Technology of Information*, **3**(2), pp. 81–90, 2010.
- [21] TRAUSAN-MATU S., *Collaborative and Differential Utterances, Pivotal Moments, and Polyphony*, in D. Suthers, K. Lund, C. P. Ros & N. Law (Eds.), *Productive multivocality*, New York, NY: Springer, 2013.
- [22] TRAUSAN-MATU S., KARATZAS K., CHIRU C., *Environmental Information Perception, Analysis and Communication with the Aid of Natural Language Processing*, in *Proceedings of the 21st International Conference on Informatics for Environmental Protection Environmental Informatics and Systems Research*, 2007.
- [23] TRAUSAN-MATU S., REBEDEA T., *Polyphonic Inter-Animation of Voices in VMT*, in Stahl G. (Ed.), *Studying Virtual Math Teams*, pp. 451–473, Boston, MA: Springer US, 2009.
- [24] TRAUSAN-MATU S., REBEDEA T., DRAGAN A., ALEXANDRU C., *Visual Analysis and Assessment of Collaborative Learning in Chat Conversations*, *Proceedings of WBL 2007*, Pearson-Prentice Hall, 2007.
- [25] TRAUSAN-MATU S., REBEDEA T., *A Polyphonic Model and System for Inter-animation Analysis in Chat Conversations with Multiple Participants*, in A. Gelbukh (Ed.), *Computational Linguistics and Intelligent Text Processing*, Vol. **6008**, pp. 354–363, Springer Berlin / Heidelberg, 2010.
- [26] TRAUSAN-MATU S., STAHL G., ZEMEL A., *Polyphonic Inter-animation in Collaborative Problem Solving Chats*, Drexel University, Philadelphia, 2005. Retrieved from <http://mathforum.org/wikis/uploads/Stefan.Interanimation.doc>
- [27] TRAUSAN-MATU S., STAHL G. SARMIENTO J., *Polyphonic Support for Collaborative Learning*, in Dimitriadis Y.A. et al. (Eds.): CRIWG 2006, *Lecture Notes in Computer Science* **4154**, pp. 132–139, ISSN 0302 9743, Springer-Verlag, Berlin, Heidelberg, New York, 2006.
- [28] TRAUSAN-MATU S., STAHL G., SARMIENTO J., *Supporting Polyphonic Collaborative Learning*, *e-Service Journal*, **6**(1), pp. 59–74, 2007.
- [29] WINOGRAD T., *Thinking machines: Can there be? Are we?*, Report No. STAN-CS-87-1161, Stanford, 1987.