Modeling Comprehension Processes via Automated Analyses of Dialogism

Mihai Dascalu (mihai.dascalu@cs.pub.ro)
Department of Computer Science, University Politehnica of Bucharest
Bucharest, 060042 Romania

Laura K. Allen (laurakallen@asu.edu)
Danielle S. McNamara (dsmcnama@asu.edu)
Department of Psychology, Arizona State University
Tempe, AZ, 85287-2111 USA

Stefan Trausan-Matu (stefan.trausan@cs.pub.ro)
Department of Computer Science, University Politehnica of Bucharest
Bucharest, 060042 Romania

Scott A. Crossley (scrossley@gsu.edu)
Department of Applied Linguistics/ESL, Georgia State University
Atlanta, GA, 30303 USA

In press - Completed 2017

Dascalu, M., Allen, K. A., McNamara, D. S., Trausan-Matu, S., & Crossley, S. A. (in press). Modeling comprehension processes via automated analyses of dialogism. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar (Eds.), *Proceedings of the 39th Annual Meeting of the Cognitive Science Society* (CogSci 2017). London, UK: Cognitive Science Society. To be published with acknowledgment of federal support.

Author's Note

The work presented in this paper was partially funded by the University Politehnica of Bucharest through the "Excellence Research Grants" Program, UPB – GEX Contract number 12/26.09.2016, by the EC H2020 project RAGE (Realising and Applied Gaming Eco-System) http://www.rageproject.eu/ Grant agreement No 644187, as well as funding to Arizona State University (IES 305A130124, IES R305A120707, NSF 1417997, NSF 1418378, ONR 12249156, ONR N00014140343). Any opinions or conclusions expressed are those of the authors and do not represent views of the IES, NSF, or ONR.

Modeling Comprehension Processes via Automated Analyses of Dialogism

Mihai Dascalu (mihai.dascalu@cs.pub.ro)
Department of Computer Science, University Politehnica of Bucharest
Bucharest, 060042 Romania

Laura K. Allen (laurakallen@asu.edu)
Danielle S. McNamara (dsmcnama@asu.edu)
Department of Psychology, Arizona State University
Tempe, AZ, 85287-2111 USA

Stefan Trausan-Matu (stefan.trausan@cs.pub.ro)
Department of Computer Science, University Politehnica of Bucharest
Bucharest, 060042 Romania

Scott A. Crossley (scrossley@gsu.edu)
Department of Applied Linguistics/ESL, Georgia State University
Atlanta, GA, 30303 USA

Abstract

Dialogism provides the grounds for building a comprehensive model of discourse and it is focused on the multiplicity of perspectives (i.e., voices). Dialogism can be present in any type of text, while voices become themes or recurrent topics emerging from the discourse. In this study, we examine the extent that differences between self-explanations and think-alouds can be detected using computational textual indices derived from dialogism. Students ($n = 68$) read a text about natural selection and were instructed to generate self-explanations or think-alouds. The linguistic features of these text responses were analyzed using *ReaderBench*, an automated text analysis tool. A discriminant function analysis using these features correctly classified 80.9% of the students' assigned experimental conditions (self-explanation vs. think aloud). Our results indicate that self-explanation promotes text processing that focuses on connected ideas, rather than separate voices or points of view covering multiple topics.

Keywords: comprehension; discourse analysis; dialogism; polyphonic model; self-explanation; think-aloud

Introduction

Research on text comprehension suggests that skilled and less skilled readers differ in the frequency and type of strategies they employ while processing texts (Millis, Magliano, & Todaro, 2006; Oakhill & Yuill, 1996). Skilled readers, for instance, generate more inferences while reading, which allows them to establish connections between information in the text and their prior knowledge (Kintsch, 1998). Although not all readers naturally make these connections while reading, students can be prompted to generate inferences through instructions to *self-explain* (McNamara, 2004). Self-explanation is a response to text or discourse that is directed toward oneself, with an explicit purpose to construct meaning from the text. Explanations are statements generated aloud, through text, or silently to oneself, that go beyond the information provided explicitly in the text to explain the ideas, their relations, and their underlying meaning.

In the context of text comprehension, self-explanation can improve readers' understanding of complex topics (McNamara, 2004, in press). From the point of view of theories within the field of text and discourse comprehension, the benefits of self-explanation have been attributed to increased bridging and elaborative inferences (i.e., making connections to prior ideas in the text or to prior knowledge) and to increased causal inferences (e.g., making connections to causal events; (Allen, McNamara, & McCrudden, 2015).

Instructions to self-explain can be contrasted with those to think-aloud, which ask readers to report whatever thoughts are available to them while reading a text (e.g., readers report these thoughts after reading each sentence). Asking a reader to think-aloud reveals their use of comprehension strategies, including any inferences they generate while reading, but does not alter a reader's natural comprehension processes (McNamara & Magliano, 2009).

In this study, we compare the comprehension processes associated with self-explanation and think-aloud from the lens of dialogism. *Dialogism* refers to a wider perspective of dialogue which is assumed to be present within any verbal or non-verbal language activity. Dialogism originates from the Russian philosopher and philologist Mikhail Bakhtin (1981, 1984) who proposed that there is an implicit and multi-voiced dialogue underlying sense-making, communication, actions, and interactions (Linell, 2009). Accordingly, *voices* represent distinct positions, points of view, or ideas, that impact the nature and outcome of a discourse. Multi-voicedness, may drive to *polyphony*, which is a central concept within dialogism, and a focus of the *polyphonic model*, which is essential to this study (Trausan-Matu & Rebedea, 2009).

Dialogue traditionally refers to communication between two or more individuals. Indeed, within the context of

Computer Supported Collaborative Learning, dialogism has been considered better suited as a theoretical framework for multi-party conversations than classic Natural Language Processing theories that focus on phone-like interactions between two interlocutors (Trausan-Matu & Rebedea, 2009).

Bakhtin, however, stressed the point that *all text* is multi-vocal, wherein our speech (all of our utterances) is filled with others' words (Bakhtin, 1986, p. 89). In this view, the concept of dialogue can be extended to include a wider variety of language activities. For instance, it may also refer to an internal dialogue within oneself or a dialogue amongst inner voices contrasting and debating ideas (Marková, Linell, Grossen, & Salazar Orvig, 2007, ch. 6).

The polyphonic model is a generalization of Bakhtin's ideas in the sense that voices may not be only associated to an individual person. Voices may also be themes, or recurrent topics emerging from the discourse. They enter in inter-animation patterns which generate a polyphonic weaving characterized by a multitude of voices, each with its individuality, but which give birth to a coherent whole (Trausan-Matu & Rebedea, 2009).

In this study, we apply the polyphonic model to analyze the presence of voices and their interactions within students' think-alouds and self-explanations. Think-aloud, by definition, is the externalization of the inner voice of the student, including voices that correspond to ideas, justifications, and assumed positions. Expressions of the text comprehension process by a student, in the form of think aloud or self-explanation, can be expected to include positions, reasons, ideas, all which may be viewed as voices.

We have operationalized the polyphonic model within the *ReaderBench* framework, which provides automated text and conversational analysis (Dascalu, 2014; Dascalu, Dessus, Bianco, Trausan-Matu, & Nardy, 2014; Dascalu, Trausan-Matu, McNamara, & Dessus, 2015). In *ReaderBench*, voices are identified using Natural Language Processing (NLP). *Lexical chains* (Galley & McKeown, 2003), or sequences of repeated or related words, are merged into *semantic chains* by using relatedness calculated using Latent Semantic Analysis (Landauer & Dumais, 1997) and Latent Dirichlet Allocation (Blei, Ng, & Jordan, 2003). The previous semantic models play an important role in our polyphonic model as they are used to identify voices through semantic relatedness, thus highlighting cohesive contexts.

In addition, polyphonic *inter-animation* considers relations between voices, or points of view, along two dimensions: *longitudinal* along time, and *transversal* across time, using voices' co-occurrences within and across text segments (Trausan-Matu, Stahl, & Sarmiento, 2007). The longitudinal dimension follows the continuation of ideas throughout the discourse, similar to a voice's individual melodic line in music. Simultaneously, voices co-occur in a vertical manner and, as in polyphonic music, this generates specific discourse contexts consisting of potential dissonances that need to be solved toward consonances. This transversal effect, or voice overlap, supports the integration process that can create both unity across various themes, as well as differences or variations in points of view.

Specifically, we examine the extent to which differences between participants' expressions of self-explanation and think-aloud can be detected using the computational text analyses provided by ReaderBench, and in turn, how this theoretical perspective informs our understanding of text and discourse comprehension.

Discourse Analysis within the Polyphonic Model

The polyphonic model can be used to analyze discourse in both conversations and plain texts (Trausan-Matu & Rebedea, 2009). Bakhtin (1984) stated that polyphony occurs in any text, similarly to polyphonic music, composed under counterpoint rules. That means that there is a multitude of voices, each with its own individuality, whose sum comprises a coherent whole: "the voices of others become woven into what we say, write, and think" (Koschmann, 1999, p. 308). Meanwhile, the polyphonic discourse should also bring novelty, voices should inter-animate in order to foster creativity. Following this perspective, the polyphonic approach to discourse analysis identifies voices in text and then investigates how voices are woven and how they inter-animate (Trausan-Matu, Stahl, et al., 2007).

Our automated process of voice identification starts by building lexical chains that are merged into semantic chains through semantic relatedness (Dascalu et al., 2015). Lexical chains can be identified using a disambiguation graph in which nodes are word instances having assigned their most probable sense, while weighted edges are semantic distances from *WordNet* (Galley & McKeown, 2003). However, this approach is inherently limited because it only includes words from the same part of speech. Thus, we have used an iterative agglomerative hierarchical clustering algorithm that begins with the identified lexical chains as groups of clustered words and uses the semantic similarity between lexical chains as a distance function (Dascalu et al., 2015). If the semantic relatedness value is greater than an imposed threshold or if identical lemmas are identified in two word clusters, the latter are automatically merged.

Voices emerge as central topics of each text and rely on the occurrences of the underlying cohesive and semantically related words. The *longitudinal* dimension of voices becomes the context in which the voices span throughout the entire discourse. In contrast, the *transversal* dimension highlights different co-occurrence and inter-animation patterns of voices present within the same textual element, i.e., sentence or paragraph.

After voices are identified, a cohesion (or utterance) graph is constructed from the links between utterances (Dascalu, 2014; Dascalu, Dessus, Trausan-Matu, Bianco, & Nardy, 2013; Trausan-Matu, Dascalu, & Dessus, 2012; Trausan-Matu, Rebedea, Dragan, & Alexandru, 2007) Within the cohesion graph, utterances are the nodes and links consist of adjacency pairs, repetitions, or lexical and semantic chains, which are detected using NLP.

As such, voices can be identified as threads in the graphs (Trausan-Matu, Dascalu, & Rebedea, 2014). Each utterance has an inner voice that inter-twines with other voices from the same thread or from different ones, but with less strength. Any new utterance in a dialogue is expressed as a voice, including its degree of interconnection with other utterances, relevance within the discourse, and potential impact within the overall discussion. Examining different semantic chains within the same textual fragment (sentences or paragraphs) reveals the transversal dimension of voice inter-animation.

Current Study

This study comprises an analysis of a corpus of self-explanations and think-alouds previously described in Allen et al. (2015). University students ($n = 68$) read a text about natural selection and were randomly assigned to one of two conditions related to their reading instructions: self-explanation ($n = 33$) and think-aloud ($n = 35$). Students in the self-explanation and think-aloud conditions were prompted to generate typed responses on 16 occasions (i.e., on 16 of the 41 sentences). The self-explanation instructions asked students to explain the information they had just read to themselves, whereas the think-aloud instructions asked students to state whatever they were thinking. We aggregated students' 16 text responses (their self-explanations or think-alouds) following the procedure described in Allen et al. (2015).

The individual files were then analyzed using *ReaderBench*. We calculated 29 voice indices related to:

a) *span* (distances between word occurrences within the same semantic chain),
b) *recurrence* (average and standard deviation in terms of distance between two consecutive words pertaining to the same voice, measured in number of in-between words from the initial text),
c) *coverage* of these semantic chains (average number of contained concepts per sentence or paragraph), and
d) information theory entropy (Shannon, 1948) based on the probability that a voice appears in a given text segment.

The previous indices relate to the longitudinal dimension of our analysis, while voice inter-animation relates to the transversal effect, which is computed in terms of *co-occurrence patterns*. As operationalization of the transversal dimension, we rely on: a) counting the number of concepts pertaining to different voices, but present in the same text segment, and b) pointwise mutual information (PMI) that measures the degree of association between voice distributions (Dascalu et al., 2015). These dialogic indices provide insights in terms of a text's overall cohesion, as voices help build a higher cohesion through lexical and semantic relatedness.

Self-explanation (SE)	Think-aloud (TA)
In our lives, there are so many types of people around us to our lives colorful. also, in our daily lives, we meet different people who have different story to tell. some of them are happy, wealthy. Some of them have to worry about how to survive in this society. they are components to make our lives fascinates. Life around us is fascinating because of the force of nature. those creatures around us are differently designed. Some of them are capable of seeing stuffs because they are given an unique thing-eye. that's one of the things to make them special, to make their lives fancy. Life fascinates us because we have eyes. And eyes have precise arrangement so that eyes make us see things. This is also true for other organs, they are complexly design to make our body function. Organs are not designed in advance for a specific purpose. organs are formed by the activities people do in their everyday-lives. organs are formed for the destination to make people survive, to make people's body function well. the two animals with cloudy lenses must give their next generation cloudy lens. and the generation of cloudy lens animal and clear lens animal will be hard to tell. because the offspring are given clear lens due to those who gave birth to it. Because a replicator is something that can make a copy of itself, with most of its traits duplicated in the copy, including the ability to replicate. The offspring's parents survived and pass this replicator to it, so that the offspring's eyes are the same as its parent's.	The surroundings we cannot to change, but we can our heart to adopt. In my mind, the human also as one of the animals in the world, we have only different from the other animals because we have a thought. The eyes is difference with the other organs. The animals eyes may be less important than other organs. The author cannot to believe that the organs must have been designed in advance for a specific purpose is right. Used an example to support his ideal. According to the example, I feel that the offspring has clear lenses and can see well which is incorrect. in some way, their eyes has different with their parents' eyes. That's mean the their another eyes is usedness. the better vision can help these animals to reproduce and get better generation. It's to point out the ideal which is the author want to explain. The living surrounding makes natural selection in order to get better next generation. that's mean we can change or direct the selection to product. replicators try to use-up material to find their the great copies and energy to power replication. the most of the copying is worse that causing less efficient just the less of the copying is better and useful, back to the viewpoint. the apparent well engineered body is result by the replicator to make the natural selection. Organisms is the standard by the natural selection.

Figure 1. Sample inter-animation of voices within a self-explanation and a think-aloud protocol.

Table 1: MANOVA results.

Index	Self-explanations M (SD)	Think-alouds M (SD)	*F*	*p*	Partial Eta Squared
Average span of lexical chains	**2.10** (0.29)	1.69 (0.22)	45.325	<.001	.407
Average paragraph voice co-occurrence	**6.00** (2.22)	2.99 (1.51)	43.296	<.001	.396
Average sentence voice co-occurrence	**3.08** (0.81)	1.84 (0.94)	33.212	<.001	.335
Standard deviation of paragraph voice co-occurrences	**2.93** (1.08)	1.93 (0.97)	15.984	<.001	.195
Average sentence entropy of voices	**1.32** (0.23)	1.14 (0.25)	10.202	<.01	.134
Average span of voices	**6.37** (1.64)	5.25 (1.41)	9.184	<.01	.122
Standard deviation of paragraph voice mutual information (PMI)	0.58 (0.07)	**0.65** (0.13)	5.755	<.05	.080
Percentage of words that are included in lexical chains	**0.10** (0.02)	0.09 (0.03)	5.291	<.05	.074
Standard deviation of distributions per paragraph	**0.75** (0.04)	0.72 (0.06)	4.994	<.05	.070
Standard deviation of sentence voice co-occurrences	**1.51** (0.35)	1.39 (0.47)	4.497	<.05	.064

Figure 1 presents an example of inter-animation of voices within a self-explanation and a think-aloud verbal protocol. Several threads that can be considered as voices, ranging from simple word repetitions (i.e., "organs") to semantically related concepts (i.e., "eye – lens", "generation – parent – offspring", "copy – replication – duplicate – replication"), co-appear and inter-animate. Additional voices can be identified in both texts, but it is important to observe differences in terms of distributions: denser, more cluttered and more elaborated voices are present in self-explanations versus more varied and more spread-out voices in think-alouds.

In addition, the texts exhibit different discourse structures: longer, more elaborated and more cohesive paragraphs in self-explanations versus shorter, more condensed phrases introducing multiple ideas in think-alouds. These latter discourse specific traits also directly influence the distribution of voices within the underlying analysis element (paragraph or sentence) as the chance of voice co-occurrence inherently increases in longer texts (e.g., self-explanations).

Results

Statistical analyses were conducted to assess the extent to which the dialogic indices related to voices and in turn, accurately classified students based on their experimental condition. Because *ReaderBench* reports raw voice counts, indices were also checked for multicollinearity with text length. Any index that was highly collinear ($r > .90$) with text length was removed. We then conducted a MANOVA to identify which indices exhibited significant differences between the self-explanation and think-aloud conditions. Indices that were highly collinear ($r > .90$) were flagged, and the index with the strongest effect size in the MANOVA was retained while the other indices were removed (see Table 1).

Longer spans of both lexical chains and voices, as well as higher paragraph and sentence voice co-occurrences, are indicative of longer, more elaborated texts (i.e. self-explanations). Self-explanations also have higher standard deviations of co-occurrence patterns at both sentence and paragraph levels which reflect a greater variety of voices, as well as a more diverse and unequal overlap of voices. Higher voice entropy at sentence level and higher standard deviation of voice distributions at paragraph level also support the latter finding. Moreover, self-explanations have a slightly higher coverage of words that are integrated in longer semantic chains, thus denoting a more connected discourse. In contrast, think-alouds exhibited a higher standard deviation of paragraph voice pointwise mutual information. This was specific to the generation of new ideas which may or may not be intertwined with other voices. This result is indicative of a wider spread of synergic effects - either new, more isolated voices, or ones that inter-animate more with the previous voices.

Based on the MANOVA, we selected the four indices with the strongest effect sizes to enter into a stepwise discriminant function analysis (DFA): Average span of lexical chains *(SE > TA)*, Average paragraph voice co-occurrence *(SE > TA)*, Average sentence voice co-occurrence *(SE > TA)*, and Standard deviation of paragraph voice co-occurrences *(SE > TA)*.

Table 2: DFA results.

	Type	Predicted Group Membership		Total
		SE	TA	
Original	SE	25	8	33
	TA	5	30	35
Cross-validated	SE	25	8	33
	TA	6	29	35

Note: SE = self-explanation; TA = think-aloud

The DFA yielded a significant model, $\chi 2(df = 1, n = 68) = 34.243, p < .001$, correctly allocating 55 (25 + 30) of the 68 students (accuracy = 80.9%, see Table 2). To test the stability of our model, we conducted a leave-one-out cross-validation analysis, which also yielded an accuracy of 79.4%. The measure of agreement between the actual instructional group and that assigned by our model produced a weighted Cohen's Kappa of .616, demonstrating substantial agreement.

Discussion and Conclusions

In the current study, we examined the differences from a dialogism perspective between self-explanations and think-alouds generated in response to a text. In previous research on this dataset (Allen et al, 2015), we examined the causal and referential cohesion differences between self-explanation and think-aloud. The results of the latter study indicated that *causal* cohesion, but not *referential* cohesion differentiated students who were in the self-explanation and think-aloud conditions. In the current study, we build on this prior research by examining how textual indices related to dialogism relate to students' processing of text based on their text reading instructions.

Our results indicate that students who self-explained the text generated longer voices (lexical or semantic chains with a higher span) that inter-animate more (higher voice co-occurrences at sentence and paragraph levels). This suggests that students who were prompted to self-explain responded to the text by maintaining semantic connections of the concepts within the text.

We interpret these results to indicate that self-explanation promotes specific comprehension processes that are fundamentally different from responses generated during think-aloud protocols, evidenced by students' generation of more conceptually related and cohesive text responses, rather than multiple, separate "voices" or points of view covering multiple topics. Previous research indicates that self-explanation can enhance students' understanding of complex concepts; however, it is less clear how these instructional differences manifest in the linguistic properties of students' text responses.

By adopting a Natural Language Processing approach, this study examines on-line comprehension processes at a more fine-grained level and also contributes to a better understanding of how these processes may be automatically detected via computational linguistic analyses. The polyphonic model, built on dialogism and integrating advanced NLP techniques, represented a viable alternative to analyze students' discourse and differentiate among instructional settings.

As a limitation of our approach, voices need to account for more than word clustering based on distance, lexical and semantic overlaps which are currently used to operationalize our polyphonic model. In addition, many voice indices are multicollinear with text length. We need to develop methods to normalize raw voice score to help control for text length constraints. Voices represent a generalization of emerging topics and should consider the corresponding sentiment valences in order to create a clearer perspective whether convergence or dissonances are encountered in the discourse. In order to address this issue, opinion mining techniques will be integrated in a follow-up iteration of our implemented model.

The dialogical framework offers new perspectives in the context of this study because both self-explanation and think-aloud be perceived as different kinds of dialogues. Self-explanations on the one hand include positions, reasons, ideas, all of which may be viewed as voices while the overall discourse can be regarded as an 'internal dialogue within the self' (Linell, 2009, ch. 6). On the other hand, think-alouds are more condensed, centered on generating new ideas and can also be perceived a 'dialogue between ideas' (Marková et al., 2007, ch. 6), a dialogue in which the debating voices are ideas. However, in both cases reflexive and cognitive processes are needed in order for students to express themselves.

As Linell (2009) stated, dialogues occur 'in and through words.' Certainly, there is more to dialogue and communication – for example, gestures, facial expressions, emotions, movement, all play crucial roles in dialogue; but here, because it is printed text, we can only examine words. Nonetheless, dialogism is tightly connected to the notion of sense-making as meaning is constructed by interacting with others and with the world, as well as with oneself via internal dialogue. As such, dialogism has a strong connection to cognition, which is sometimes ignored. Figure 2 represents this viewpoint. Cognition reflects prior and intrapersonal (individual) information and knowledge about the world, while meaning is constructed through social interactions and language within the dialogical context. Communication, both explicit interpersonal dialogue and implicit to oneself (i.e., internal dialogue), becomes a facilitator in terms of interaction, therein generating meaning in discourse.

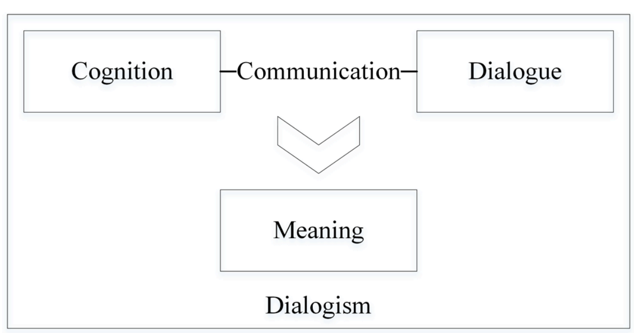

Figure 2. Dialogical framing and interdependencies with core concepts.

Acknowledgments

The work presented in this paper was partially funded by the University Politehnica of Bucharest through the "Excellence Research Grants" Program, UPB – GEX Contract number 12/26.09.2016, by the EC H2020 project RAGE (Realising and Applied Gaming Eco-System) http://www.rageproject.eu/ Grant agreement No 644187, as well as funding to Arizona State University (IES 305A130124, IES R305A120707, NSF 1417997, NSF 1418378, ONR 12249156, ONR N00014140343). Any opinions or conclusions expressed are those of the authors and do not represent views of the IES, NSF, or ONR.

References

Allen, L. K., McNamara, D. S., & McCrudden, M. T. (2015). *Change your mind: Investigating the effects of self-explanation in the resolution of misconceptions.* Paper presented at the 37th Annual Meeting of the Cognitive Science Society (CogSci 2015), Pasadena, CA.

Bakhtin, M. M. (1981). *The dialogic imagination: Four essays* (C. Emerson & M. Holquist, Trans.). Austin and London: The University of Texas Press.

Bakhtin, M. M. (1984). *Problems of Dostoevsky's poetics* (C. Emerson, Trans. C. Emerson Ed.). Minneapolis: University of Minnesota Press.

Bakhtin, M. M. (1986). *Speech genres and other late essays* (V. W. McGee, Trans.). Austin: University of Texas.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research, 3*(4-5), 993–1022.

Dascalu, M. (2014). *Analyzing discourse and text complexity for learning and collaborating, Studies in Computational Intelligence* (Vol. 534). Cham, Switzerland: Springer.

Dascalu, M., Dessus, P., Bianco, M., Trausan-Matu, S., & Nardy, A. (2014). Mining texts, learner productions and strategies with ReaderBench. In A. Peña-Ayala (Ed.), *Educational Data Mining: Applications and Trends* (pp. 335–377). Cham, Switzerland: Springer.

Dascalu, M., Dessus, P., Trausan-Matu, S., Bianco, M., & Nardy, A. (2013). *ReaderBench, an environment for analyzing text complexity and reading strategies.* Paper presented at the 16th Int. Conf. on Artificial Intelligence in Education (AIED 2013), Memphis, USA.

Dascalu, M., Trausan-Matu, S., McNamara, D. S., & Dessus, P. (2015). ReaderBench – Automated Evaluation of Collaboration based on Cohesion and Dialogism. *International Journal of Computer-Supported Collaborative Learning, 10*(4), 395–423. doi:10.1007/s11412-015-9226-y

Galley, M., & McKeown, K. (2003). *Improving word sense disambiguation in lexical chaining.* Paper presented at the 18th International Joint Conference on Artificial Intelligence (IJCAI'03), Acapulco, Mexico.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, UK: Cambridge University Press.

Koschmann, T. (1999). *Toward a dialogic theory of learning: Bakhtin's contribution to understanding learning in settings of collaboration.* Paper presented at the Int. Conf. on Computer Support for Collaborative Learning (CSCL'99), Palo Alto.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: the Latent Semantic Analysis theory of acquisition, induction and representation of knowledge. *Psychological Review, 104*(2), 211–240.

Linell, P. (2009). *Rethinking language, mind, and world dialogically: Interactional and contextual theories of human sense-making*. Information Age Publishing: Charlotte, NC.

Marková, I., Linell, P., Grossen, M., & Salazar Orvig, A. (2007). *Dialogue in focus groups: Exploring socially shared knowledge*. London, UK: Equinox.

McNamara, D. S. (2004). SERT: Self-Explanation Reading Training. *Discourse processes, 38*, 1–30.

McNamara, D. S. (in press). Self-explanation and reading strategy training (SERT) improves low-knowledge students' science course performance. *Discourse Processes*.

McNamara, D. S., & Magliano, J. P. (2009). Self-explanation and metacognition: The dynamics of reading. In J. D. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of Metacognition in Education* (pp. 60–81). Mahwah, NJ: Erlbaum.

Millis, K., Magliano, J. P., & Todaro, S. (2006). Measuring discourse-level processes with verbal protocols and Latent Semantic Analysis. *Scientific Studies of Reading, 10*(3), 225–240.

Oakhill, J., & Yuill, N. (1996). Higher order factors in comprehension disability: Processes and remediation. In C. Cornaldi & J. Oakhill (Eds.), (pp. 69–72). Mahwah, NJ: Erlbaum.

Shannon, C. E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal, 27*, 379–423 & 623–656.

Trausan-Matu, S., Dascalu, M., & Dessus, P. (2012). *Textual complexity and discourse structure in Computer-Supported Collaborative Learning.* Paper presented at the 11th Int. Conf. on Intelligent Tutoring Systems (ITS 2012), Chania, Grece.

Trausan-Matu, S., Dascalu, M., & Rebedea, T. (2014). PolyCAFe–automatic support for the polyphonic analysis of CSCL chats. *International Journal of Computer-Supported Collaborative Learning, 9*(2), 127–156. doi:10.1007/s11412-014-9190-y

Trausan-Matu, S., & Rebedea, T. (2009). Polyphonic inter-animation of voices in VMT. In G. Stahl (Ed.), *Studying Virtual Math Teams* (pp. 451–473). Boston, MA: Springer.

Trausan-Matu, S., Rebedea, T., Dragan, A., & Alexandru, C. (2007). Visualisation of learners' contributions in chat conversations. In J. Fong & F. L. Wang (Eds.), *Blended learning* (pp. 217–226). Singapour: Pearson/Prentice Hall.

Trausan-Matu, S., Stahl, G., & Sarmiento, J. (2007). Supporting polyphonic collaborative learning. *Indiana University Press, E-service Journal, 6*(1), 58–74.