Technical Proposal (Volume 1)

BAA number: BAA 11-001

Title of Proposal: Towards Optimization of Macrocognitive Processes: Automating Analysis of the Emergence of Leadership in Ad Hoc Teams

Prime Offeror: (PI) Carolyn Penstein Rosé (Language Technologies Institute/ Human-Computer Interaction Institute, CMU), (CoPI) Gerry Stahl, Drexel University, (CoPI) Sean Goggins, Drexel University, (CoPI) Emily Patterson, Ohio State University, (CoPI) Marcela Borge, Penn State, (CoPI) John Carroll, Penn State, (CoPI) Andrew Duchon, APTIMA

Technical Contact: Carolyn Penstein Rosé, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, phone 412-268-7130, fax 412-268-6298, cprose@cs.cmu.edu

Administrative Contact: Radha Rao, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, phone 412-268-6592, fax 412-268-6298, rdr@cs.cmu.edu

Duration of Effort: 3 years

Table of Contents

Technical Approach and Justification

Introduction

The primary focus to date for the CKI program has been on macrocognition in small ad hoc teams with no prior social history in order to exclusively study macrocognitive phases and processes. As the CKI program begins transitioning findings to real-world settings, social processes have emerged across many of the efforts as critical for achieving effective teamwork, for both human-human and human-agent teams. Furthermore, automating intensive analytic coding efforts has been identified as an important area of opportunity to increase productivity of all CKI researchers and aid comparison of findings across studies. In order to enable dynamic support for optimal team performance, it will be beneficial, perhaps even necessary, to automate analysis of collaborative behaviors in real time. Finally, while the majority of analysis efforts have focused on single sessions of interaction, as the focus of the community turns towards longitudinal studies, such as within the APAN analysis effort, integration across analysis methodologies and units of analysis, with a particular focus on incorporating social network analysis technology, will become more important. This proposal seeks to pursue these interrelated and synergistic goals with one integrated, multi-institutional effort involving a basic research effort across four academic institutions, namely Carnegie Mellon University, Drexel University, Penn State, and Ohio State University, and an industrial partner, Aptima, Inc.

Proposed tasks:

1. Conduct automated analysis of multiple existing, coded datasets
2. Compare reliability for emerging leadership indicators from different scientific perspectives

The two proposed tasks are deeply synergistic. Task 1 provides a technological infrastructure to facilitate work on Task 2. Task 2 provides focus for continued work on Task 1, enabling the identification of key challenges faced by analysts using the technology in their basic research. In this way, we can be assured that our technological work focuses not just on what advances the fields of text mining, machine learning, and social network analysis, but that advances them in service of behavioral science that is of central importance to the CKI mission.

Overview of Proposed Work

Task 1: Automated Coding

The first major thrust of the project will be development of technological support for coding CKI corpora for which there is already research ethics approval to share. Existing CKI corpora mainly consist of strictly anonymized discussion data in a variety of formats and strictly anonymized logfiles from the macrocog testbed. Typically, audio data were collected in single 1-4 hour sessions and subsequently transcribed, but we will not be obtaining or using the audio data. In addition to process data, which have in many cases been coded by hand, there are summative assessment measures from frameworks developed within the broader CKI community by researchers such as Joan Rentsch, Steve Shope, and Sara McComb. We will make use of these to the extent that they are also fully anonymized and available for free distribution.

In order to provide technological support for analysis of CKI data, we propose to collect these existing corpora, standardize the storage format into a canonical database, and conduct a series of experiments applying existing analysis technology from Carnegie Mellon University and Aptima. The CKI Combined Canonical Corpus (C4) would standardize the raw data and results from analyzing those data (either by hand or computer). This will enable quicker application of further automated analyses and visualizations, as well as extensions to datasets outside the CKI program. Individual experiments often do not have enough data to feed statistical machine learning algorithms, so the combination of data sets from a wide variety of sources will enable the robust application of these techniques, as well as permit the comparison of different coding schemes from the CKI community and others. Of particular applicability for this effort would be the investigation of domain adaptation techniques under development in Dr Carolyn Rosé's group at Carnegie Mellon University. Domain adaptation, which is a cutting edge problem area in machine learning, would allow models trained on one dataset to generalize to new datasets with little or no coded data from the new corpora. These techniques could be further developed using the combined corpus as a testbed.

One major goal of the automatic analysis thrust would be the application of CKI analysis frameworks to the APAN dataset. Since the APAN data have not yet been coded, application of models trained using domain adaptation techniques could be crucial for the CKI program to quickly obtain an understanding of the macrocognitive processes within these data. Thus, the work related to automated analysis of the combined corpus will feed into this effort.

Task 2: Analysis of the Emergence of Leadership

The second major thrust of the proposed research is the analysis of the emergence of leadership in ad hoc teams. Ultimately, the goal will be to apply this analysis to APAN as well as to use an automated version of this coding as a key component in technology for supporting effective team processes for optimizing macrocognition in teams. This synergizes with the automatic analysis effort in that it will raise new challenges and questions for the automatic analysis effort.

Analysis of leadership has been an undercurrent in the work being conducted at Carnegie Mellon University, Penn State, Ohio State University, and Drexel University. Interesting questions related to the perception of leadership have been raised by recent investigations. OSU recently found that who provided praise following the completion of the task was a strong but infrequent indicator of leadership, that the person who identified a plan as acceptable was a frequent and usually reliable indicator, and that the person who most dominated in terms of overall utterances and being the first to speak was moderately correlated with leadership. The CSCL group at Penn State specifically asked participants to identify who they felt was their team leader for each of the 22 teams included in the last iteration of the study. When team selections were matched to researcher selections (researchers used other common indicators i.e., who led discussions, provided most decisions, etc.) they did not match up. Thus the question emerges as to what are the qualities that these individuals are tuning into in order to define their leaders and which of these qualities correlate to effective teams. The Drexel and Carnegie Mellon University teams have explored data analysis techniques for identifying leadership at different levels of granularity. Drexel researcher Sean Goggins recently explicated techniques for identifying activity type in completely online small groups and for identifying emergence and transition of group structure and leadership in online groups. Furthermore, an investigation of techniques for log analysis suggests that interactivity can be designed for through more explicit capture of semantically relevant events in log data. The team at Carnegie Mellon University is developing a multi-dimensional framework for analysis of social positioning in conversational interactions, which has recently demonstrated significant correlations with patterns in task related behavior linked with summative success metrics within computer supported collaborative learning data.

As an initial step, we plan to collect together a subset of the Combined Corpus, which we will refer to as the Leadership Corpus. This will consist of datasets collected as part of the completed work at Carnegie Mellon University, Penn State, and Ohio State University that have already been coded with the different coding schemes related to leadership developed separately at the three institutions. We will then code portions of data from all three institutions with all three analysis frameworks in order to investigate similarities and differences across these coding schemes.

Next steps include qualitative analysis techniques for investigating how leadership behaviors may be viewed differently by different members of a team. We will conduct correlational analyses to determine which codings have the best predictive value for CKI summative measures as part of a validation process. Note that no new data collection with human subjects will be conducted under this grant.

Detailed Description of Project

TASK1: AUTOMATIC CODING

Existing Infrastructure to Support Task1 Work

From a computational perspective, we aim to develop algorithms for extracting social meaning from text by means of cues identified from linguistic research. To do so, we will draw on both Carnegie Mellon University's and Aptima's expertise in computational linguistics, machine learning and artificial intelligence to extract semantic markers that also convey the meaning of the word/sentence. We will also draw on the Drexel team's expertise in social network analysis.

Rosé and her students have worked on what signals discourse level structure in on-line discussions (Wang & Rosé, 2010; Wang et al., 2008), how attitudes are communicated through blog posts (Joshi & Rosé, 2009; Arora, Joshi, & Rosé, 2009; Mayfield & Rosé, 2010; Arora, Mayfied, Rosé, and Nyberg, 2010), how perspective is communicated through conversational contributions and how conversational participants influence one another through interaction (Ai, Kumar, Nguyen, Nagasunder, & Rosé, 2010; Nguyen, Mayfield, and Rosé, 2010). From the standpoint of leveraging constructs from systemic functional linguistics, a categorical analysis is necessary, and for that we will build on Rosé's prior work in text classification (Rosé & VanLehn, 2005; Rosé et al., 2008; Ai et al., 2010b). Her recent work has focused on use of genetic programming to strategically evolve small numbers of very powerful features to increase the representational power of more traditional feature spaces for text mining without significantly increasing the total number of features (Mayfield & Rosé, 2010). We expect that the ability to strategically evolve more complex features will enable this categorical analysis.

The SIDE toolbench (Mayfield & Rosé, 2010) will provide a boundary object to facilitate the collaboration between CMU, Aptima, and Drexel. Work on development of linguistic preprocessing technology, such as various forms of structured topic models that have featured prominently in the prior work of both the CMU and Aptima teams, will be packaged as plugins for SIDE and then made available through SIDE's feature extraction panel so that it can be used to support our proposed work and that of other similar projects within the CKI community, and in the research community more broadly. The CMU and Aptima teams will then perform machine learning experiments on prepared datasets using SIDE. Each experiment will involve selection of some subset of feature extraction options. This involves making selections on a dashboard provided through SIDE. The feature extraction can then be done by SIDE at the click of a button. The full range of machine learning algorithms, feature selection algorithms, filters, and meta-classifiers that are available off-the-shelf are available through SIDE, again through the click of a button an a separate dashboard. Using the selected feature space and machine learning configuration, a model can be built on a training portion of the data and tested on a testing portion of the data. SIDE also provides an error analysis interface that will allow investigation of potential reasons for common class confusions that occur in the evaluation. The two teams will then discuss how these results indicate that changes and enhancements need to be made in the provided linguistic preprocessing technology.

Cross-validation (Efron & Gong, 1983), a well-established machine-learning evaluation methodology, will be used to validate that the detector can generalize across students and classrooms. We will use a version of this methodology, called "10-fold cross-validation", where the data is divided into 10 subsets, and each student assigned to one of the ten subsets. Then, each combination of 9 of the 10 subsets is made, and a model is trained on the 9 subsets together and tested on the tenth subset. However, an important consideration when dealing specifically with conversational data is that data from different segments of the same conversation may differ systematically from data from other conversations. Thus, if the conversational data is segmented into units that are smaller than one interaction, then the instances within the dataset will not be independent from one another. This non-independence is known to introduce certain common problems into the experimentation. First, it can lead to the models over-fitting to idiosyncratic characteristics of the specific conversations that are within the training data, since segments within the same conversation are more likely to be the same class than segments from different conversations. Thus, these idiosyncratic characteristics will appear to predict those classes.

To address this, and avoid developing class-specific models, while developing our models we will balance how much our system considers data from different students, collaborating pairs, and classrooms. When evaluating our models, we will use a modification on standard cross-validation referred to as "leave-one-subpopulation-out cross-validation". In this methodology, we divide the students in our data set by classroom or collaborating pair rather than dividing students randomly, so that the model is trained on one set of classes/pairs and tested on another set of classes/pairs. We have used this method in the past to evaluate the transfer of detectors between students (cf. Baker, 2007), pairs of students (Joshi & Rosé, 2007), and classrooms (Baker, Corbett, Koedinger, & Roll, 2005).

Another important direction for addressing this non-independence issue, which can adversely effect success at processing conversational data if not addressed explicitly, is work on domain adaption (Joshi & Rosé, submitted). Below we will discuss proposed work to this end.

Year1 Work

The focus of the year 1 work under Task 1 for the CMU, Drexel, and Aptima teams will be to further develop an infrastructure for automating analysis of collaboration data within the CKI program. This will involve both an effort to gather together existing CKI datasets and coding schemes as well as to work on automatic analysis of the APAN dataset. CMU, Drexel, and Aptima have each independently developed their own unique technologies that will become part of the integrated tool bench that will be delivered to the CKI program, along with documentation, training materials, and if desired, organized training workshops. For example, Aptima is unique in its development of visualization tools that allow analysts to interact in a tangible way with data. The CMU team has great expertise in computational linguistics and text mining, and has produced tools that make it easy to extend cutting edge language technologies and then place them as tools in the hands of novices. The Drexel team has produced a suite of social network analysis tools that focus on connecting the knowledge construction and group identity development of technologically mediated teams with changes in the group network structures over time. To date, the Drexel team has integrated data from online learning, software engineering and medical software community interaction data. As an inter-operable whole, the CMU-Aptima-Drexel integrated tool suite will place the best that technologists in the CKI community have to offer in the hands of human factors experts who have the insight to wield these tools to the most strategic ends for the program as a whole. The CMU team is very experienced in running training workshops both locally and at international conferences and symposia.

In Year 1, Aptima will take the lead in assembling a dataset referred to as C4 in the Deliverables table below, which will consist of existing CKI datasets. This will provide a common resource for the computational work that will be done at CMU, Drexel, and at Aptima.

During Year 1, the Drexel team will integrate the C4 dataset, including the APAN data, into the Drexel Interaction Warehouse in order to begin applying Goggins' knowledge construction & group identity measurement social network analysis methods and tools to the data. A key advantage of the interaction warehouse is the ability to view CKI data side by side with existing corpora, leading to the possibility of more generalized findings for CKI. Comparative analysis of interactions across these environments is consequently difficult. One might consider questions such as: How does the shape and structure of a software engineering team differ, over time, from a group of graduate students in a completely online course or teams of aid workers in Haiti? These are the types of questions the Drexel interaction warehouse sets us up to answer (Goggins, Laffey, Amelung, & Gallagher, 2010; Blincoe, Valetto, & Goggins, 2011; Goggins, Laffey, & Gallagher, 2011; Goggins, Galyen, & Laffey, 2010).

The interaction warehouse at Drexel is an analytical processing optimized database, designed to address questions of group structure through network and time series analysis of electronic trace data. Every electronic system produces a unique set of trace outputs, which the Drexel team maps into a common, meta level structure. One key principle of the interaction warehouse design is that human computer interaction, viewed as social interaction, must consider the dimensionality of what is presented on the screen. When a person views a post in a discussion board in Sakai, an online chat in Stahl's VMT environment or a bug report in Bugzilla, they are not simply looking at a single, preceding interaction. The interactions occur between the reader, responder and several interactions, depending on the specific environment in question. The Drexel interaction warehouse accounts for this by expanding the interaction record to reflect user experience (Goggins et al., 2010). We show this in figure one.

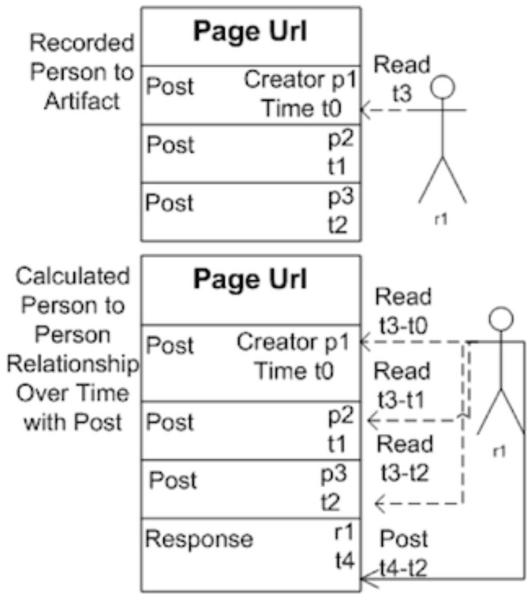

Figure 1 - Trace Data Expansion, explained in Goggins et al, 2010

Each system has unique trace data, which we map into a common, more representative structure in the interaction warehouse. For example, figure one illustrates the expansion of trace data to reflect user experience, showing an example from CANS, which keeps track of both read and post data. We are not arguing that the use of social network analysis against trace data is novel. However, we are suggesting that the quality and dimensionality of context aware trace data transformed in the Drexel interaction warehouse produces a more representative picture of social relations than cluster analysis or traces that lack these features. Illustrations of intensity of relations and breadth of relations together increase the usable information available for making social judgments. Human social intelligence requires information to make judgments in a social setting. So far, the two approaches taken by the Drexel team in this work increase the usable information in three important ways. First, the dimensionality and granularity of log data from CANS and a second project called MyLyn enables the production of high quality, ground truth verified clusters from trace data. Second, the Drexel team's enhancement of the logs to transform interaction weighting based on features of time and interaction type increases the perceived accuracy of the analysis. For example, reading is more passive and signifies a weaker tie than posting; actually working on a piece of code indicates a different type of connection between two developers than commenting in a bug tracker. Third, the first two methods combined show clusters of strong connection and a full landscape of all connection. Together, these displays of information contribute key concepts to the design of analytical tools.

A key component of the year 1 technical work will take place at CMU, specifically work on domain adaptation of machine learning models (Joshi & Rosé, under review). In order to enable effective use of machine learning models trained on one dataset on a different dataset, techniques for computing structural correspondences between corpora must be used. These allow learning more generalizable models. From a practical standpoint, that would make it possible to substantially reduce the need for hand analysis on new corpora that are collected within the CKI program. Much state of the art work in domain adaptation assumes simple correspondences between domains, which may not apply in the very different domains in which CKI data is collected. The CMU team is exploring new directions within the space of domain adaptation techniques that employ spectral processing techniques that relax assumptions of simple correspondences, and allow more indirect connections between representations within different domains, with the goal of being able to achieve farther transfer across domains that are more different from one another than ones where successful transfer has been demonstrated in past work (Fuxin et al., 2007; Dai et al., 2009).

Year2 Work

In Year 2, the CMU, Aptima, and Drexel teams will continue their work on existing CKI datasets but will also begin work on the Leadership corpus collected during Year 1 under Task 2, which combines corpora from CMU, Drexel, Ohio State, and Penn State, which have been coded with a variety of coding schemes related to leadership. Automatic analysis of the coded Leadership corpus produced in Year 1 under Task 2. Using domain adaptation techniques developed during year 1, we will explore application of macrocognitive coding schemes and leadership coding schemes to the APAN dataset as well. We expect that this will raise new challenges for the domain adaptation research, which we will work to address.

During year 2, the Drexel team will apply existing group identity and group knowledge construction coding instruments to the three corpora to inform the continued development and generalization of tools for automated network analysis. Goggins shows the relationship between group identity and group structure (Goggins, Laffey, Galyen, & Mascaro, 2011) and knowledge construction and group structure (Goggins, Laffey, & Mascaro, 2011) in two upcoming publications. The content analysis strategies used in this work will be applied to the three corpora under this proposal, and the results of that analysis – both social network and time series analysis – will increase understanding of the relationship between group identity, knowledge construction and leadership in technologically mediated teams in the CKI corpora. Existing corpora in the interaction warehouse, combined with the CKI corpora are mutually beneficial for developing generalized analytical tools.

For the three corpora in this proposal, the Drexel team also plans to map the interaction structures into the interaction warehouse and apply the techniques developed thus far to them for the purpose of identifying patterns of leadership emergence. In Goggins, Laffey & Gallagher (2011) we identified changes in leadership structure in completely online groups through analysis of electronic trace data. In Goggins, Galyen & Laffey (2010) we analyzed trace data in the manner described above to identify activity type in a completely online class. The traits of the three corpora in this study represent an important set of new structures to analyze, and provide an opportunity to develop knowledge of generalizable interaction patterns through the integration of these an other corpora using a common analysis strategy.

Year3 Work

In year 3, the Carnegie Mellon team will work with the coded corpus provided by the OSU team in order to develop automatic prediction models using SIDE to replicate the coding so similar leadership coding might be able to be applied to other CKI corpora. We will also release the analysis software and tutorial for technology supported coding for the CKI community, and will work with the program director to arrange for dissemination in the way he determines would be most effective for the community. During this year, if not before, we will plan to publish at least one journal publication, probably several, related to the many-faceted analysis of the APAN data we will perform with our technological infrastructure. We expect that interesting publications will come out of the transfer of coding schemes between corpora that will also be facilitated by this research, which will allow for a much broader exploration and validation of findings from earlier CKI studies than has been possible in the past because coding has been a primarily painstaking activity done by hand.

The Drexel team will develop at least one conference and one journal publication in year three. These publications will correspond with the release of the tools developed under this grant to other CKI researchers, including the development and presentation of training materials.

TASK 2: ANALYSIS OF EMERGENCE OF LEADERSHIP

Existing Infrastructure to Support Task2 Work

Part of the existing infrastructure that we will build on for analysis and support of emergence of leadership in ad hoc teams is a framework developed at Carnegie Mellon University for analysis of social positioning within conversation that draws from the field of systemic functional linguistics (Howley, Mayfield, and Rosé, to appear). In our work, we are taking these qualitative constructs from this work, making them precise and computationalizable, and then developing technology to automate this analysis.

The field of systemic functional linguistics is a largely descriptive linguistic tradition that provides a firm foundation in analyses of genres of writing (Martin & Rose, 2003; Martin & White, 2005), as well as face-to-face interaction (Veel, 1999), characterized in terms of the choices authors and speakers make about how to present themselves through language (Halliday, 1994). In particular, the work related to the Engagement metafunction (Martin & White, 2005), allows us to characterize a conversational contribution in terms of the propositional content communicated, the source of that content, the author/speaker's attitude towards that content, the assumed attitude of listeners towards that content, as well as the speaker's alignment or misalignment with the listeners and/or the source of the content. Extensive work applying frameworks with a similar flavor to academic writing (Hyland, 2000) allows us to view social interactions within academic writing. For example, we can view a published work as a contribution to an ongoing interaction. We can see social interactions through citation networks, and the manner in which citations are introduced within text. In earlier work (Cress & Kimmerle, 2008), wiki edits were already viewed as a form of socio-cognitive conflict. As an extension of these two ideas, we can view wiki edits from a social perspective as well. In so far as the writing that is edited already conveyed the original author's positioning within the discourse of the wiki, the edit itself can then be viewed as a social commentary on this positioning. And thus, the constructs coming from the field of systemic functional linguistics provide us with a common foundation for exploring stylistic norms of conversational behavior across genres of dialogic interactions that on the surface appear quite different. This generalizable view of social positioning viewed through conversational interactions and other types of writing will be a key insight allowing us to analyze not only the conversational interactions within the APAN dataset, but to integrate with that analysis the documents that are attached to much of the correspondence found within that dataset.

The field of systemic functional linguistics provides a wealth of constructs that can be used as lenses through which to view the stylistic choices of conversational participants, however, in order to bound our work, we will limit ourselves primarily to three, namely Martin & Rose's (2005) Negotiation system, which we have termed Exchange, Martin & White's (2003) conceptualization of Bakhtin's notion of heteroglossia (Bakhtin, 1986), and the more elaborate Engagement system, which is part of the larger system of Appraisal.

What we are concerned with in the Negotiation system is the codification of how information and goods and services are exchanged within a conversation, which is why we have termed the aspects we will discuss here as Exchange. In a collaborative learning setting, relevant goods and services frequently include helping actions. Whereas the helping coding schemes discussed in the earlier section were developed specifically to codify the process by which help is exchanged, the Negotiation system is more general, and can be applied to the exchange of all types of goods and services through conversation. For example, something as different from help exchange as a clerk requesting a customer to pay a certain amount of money can be analyzed within the same framework. What makes this system particularly valuable is the way it serves to provide a natural segmentation of an ongoing interaction into episodes where some transaction between parties within the conversation has been accomplished, i.e., either some piece of information has been exchanged, or some service has been rendered. These small accomplishments then become the building blocks for larger, and more complex accomplishments that might require a more concerted, long term effort, such as building an integrated understanding of a phenomenon, developing a plan, or solving a problem. Some of these exchanges may occur with minimal representation in terms of conversational moves. For example, a speaker could simply express an unsolicited piece of information, and it could be tacitly accepted without comment by the other members of the conversation. In this case, the exchange may require only one conversational move. However, it may not be that simple. The full exchange system allows for exchanges to transpire over multiple moves, which can all be seen as connected. The provision of the information or goods and services, and possibly the preceding request if there was one, are treated as being in a prominent position within the exchange. Other kinds of moves play a supporting role in the exchange. Each move does work towards the accomplishment of the goal of the exchange.

Within the exchange framework, speakers take up transitory speech roles within an exchange structure, in which one speaker takes up the primary role, which places the other speaker into the secondary role. It can also happen that a speaker places himself into a secondary role, which then casts the other speaker into a primary role. In group discussions, overhearers who are present but are neither cast in the primary or secondary role are by default cast into a tertiary role. Where information is exchanged, the roles are termed primary knower and secondary knower. Where goods and services are exchanged, the roles are termed primary actor and secondary actor. For example, when a clerk requests a customer to pay, that clerk is placing himself within the secondary actor role, and thus casting the customer in the primary actor role, in other words, the one who is the source of the goods and services, in this case the payment.

The Martin & White (2003) notion of heteroglossia yields further insights into the status relationship between the two speakers. Note that the general notion of monoglossia versus heterglossia from Bakhtin seems to be much broader than the notion referred to in Martin and White (see for example http://en.wikipedia.org/wiki/Heteroglossia). Heteroglossia always refers to multiple voices in a text or in a discourse community, but it can mean, for example, different registers in which someone speaks depending upon what context they are in. In Martin and White (2005), heteroglossia refers to ways in which speakers directly encode in their expression of a clause their awareness of other stakeholders in the interaction.

In Martin and White's specification of what counts as heteroglossic as a precursor to analysis using the Engagement system, which we will discuss next, three requirements must be met: First, some propositional content must be being asserted in some form, although it may be done in such a way as to communicate extreme uncertainty. Thus, questions that are framed in such a way as the reader believes the speaker was asking an honest question, for which no specific answer seems to be supposed do not count as heteroglossic. Second, an awareness must be made visible to the presence of alternative perspectives than that represented by the propositional content of an utterance. Thus, bald claims, even if they are biased, do not acknowledge alternative perspectives. For example, "Natural gas is the obviously superior choice" is undoubtedly subjective, but it is not heteroglossic. It doesn't show any awareness that someone else might disagree. If a speaker goes on to give reasons to defend the statement, however, then that speaker is showing awareness of other perspectives. These cases will be caught by the third requirement. Third, in order to count as heteroglossic, the acknowledgement of other perspectives must be expressed grammatically (e.g., through a model auxiliary like "might") or paraphrastically (e.g., "I think") within the articulation of that propositional content. If it is implicit or signaled through the discourse structure, then that is not enough to count as heteroglossic in the Martin and White sense for the purpose of feeding in to their Engagement system, although they would acknowledge it as heteroglossic "in spirit".

The construct of heteroglossia introduces the notion that the voice of the speaker is situated among other voices. But beyond that acknowledgement of the existence of other voices, what we do not see in this simple binary distinction is the manner of that positioning. The details of that positioning are further specified within Martin and White's Appraisal framework, which includes Attitude, in which feelings are revealed towards propositional content, Graduation, in which feelings are either magnified or downlplayed, and Engagement, in which a speaker positions herself in relation to the propositional content of the utterance, positions the audience in relation to the propositional content, and positions herself in relation to the audience (Martin & White, 2005). The Engagement system begins with the distinction between heteroglossia and monoglossia, which we have just discussed. Once we have determined that an utterance counts as heteroglossic, we can then further subdivide it into utterances that Contract the positions or perspectives that are treated as viable within a conversation, or conversely, ones that Expand the scope of what is treated as viable. Either way, an acknowledgement is made that more than one way of looking at the world is at play. Utterances that contract that scope, such as making an absolute assertion that leaves no room for questioning, or out right rejecting a position, are typically seen as taking a more authoritative stance than ones that expand the options, such as making a suggestion. This notion of levels of authoritativeness is one important component of expressing the positioning of the speaker in relation to the propositional content. However, it also says something about where the speaker positions himself in relation to the audience. Taking an authoritative stance casts the other speaker into a less authoritative stance. However, when this system is further subdivided, we see other options for positioning. For example, a Distancing move, in which the source of authority is ascribed to a third party, allows an authoritative statement to be made, which may contract options, but does not interfere with the positioning between the speaker and the audience. The speaker remains committed to the authoritative proposition, but is not responsible for it.

Year1 Work

The focus of the project in Year 1 will be on Task 1, however, we will get started with the work on Task 2 by formalizing the three dimensional framework described above into coding manuals, with formal human agreement scores associated with them. We are already starting to work on this under current funding. We will assemble a Leadership corpus consisting of corpora already collected at CMU, Drexel, Ohio State, and Penn State where work on analysis of leadership has already taken place from different perspectives. As a bridge allowing comparison across frameworks, we will use the 3 dimensional conversational framework described above, applied semi-automatically or fully automatically, as enabled by our in progress technological infrastructure, to ideally the whole corpus, or if not, then selected portions of each dataset within it. The newly annotated data along with the coding manuals will be the year 1 deliverables for Task 2.

Year2 Work

During Year 2, we will increase the participation of the OSU team. The OSU team will recode data from an existing corpus and deliver that coded corpus along with an evaluation of human coding reliability. An existing corpus exists at OSU from research on small, ad hoc teams previously funded by the CKI program. This corpus was collected from a laboratory study with 12 three-person teams of study participants conducted a challenging, face valid task of optimally moving troops and supporting materials to an attack location securely, economically, and within the least amount of time possible. Six teams were in a face-to-face condition and six were physically distributed with audio platform (SKYPE) support. Each session lasted 90 minutes and has been completely transcribed into Microsoft Excel format.

In the scenario used for the data collection, each team was tasked with the mission to transport troops and cargo to a desired location while optimally satisfying time, cost, and safety constraints. Each participant was given different information critical to task completion. The task was to transport 15,000 kilograms of cargo and 100 troops to the desired location in under 2.5 hours, while also minimizing cost and maximizing security. The team could choose the route and vehicles used in the mission. Each analyst had unique information about the safety, cost, and speed/distance of the vehicles/routes along with added intelligence information.

OSU will recode the existing corpus using the three different coding schemes related to leadership developed separately at the three institutions. We will then compare our analyses with data analysis done at the other three institutions with all three analysis frameworks in order to investigate similarities and differences across these coding schemes. This research will result in a completed recoding of the existing corpus by the end of year two, as well as generate inter-rater reliability of leadership indicators that are examined in order to identify the most reliable measure.

The CMU, Drexel, and OSU teams together will work on an integrated analysis of the Leadership Corpus, developing a report related to comparisons across frameworks within that corpus. This report will include correlations between leadership coding from conversational level with CKI summative measures. As a byproduct of this work we will continue to refine all relevant coding manuals, and the updated manuals will be another deliverable this year. Comparison across frameworks as well as reliability measures together will allow us to work towards a validation of all of these frameworks as providing synergistic, reinforcing, and complementary views of leadership.

Year3 Work

During year 3, OSU will support the writing of the publications led by the other team members related to Year 2 OSU team work as well as lead the writing of a journal article publication planned for submission to the Human Factors journal. This publication will provide a synthetic literature review of different approaches to manually and automatically identifying leaders from verbal transcript (chat) data in teams and provide insights on trade-offs for using various approaches, including reliability, ease of obtaining the measures, and what research questions and software design implications are best supported by each indicator.

The Penn State team will join in for year 3 to provide another angle on the study of emergence of leadership in ad hoc teams:

The Penn state team has already conducted research on leadership in teams, which we will leverage in this proposed project. Here we describe the dataset that they have already collected and will work with on this project:

Collaborative Task, Roles, and the Design of Solutions. A collaborative information analysis scenario was constructed involving a series of campus laptop thefts. Two hundred and twenty two pieces of information on people, locations, and social relationships were given to each team. The team was asked to organize and evaluate this information to create a list of eight suspects (Phase 1), identify primary suspects and motives for a set of possibly-related thefts (Phase 2), and finally predict the next crime that would occur (Phase 3). The "campus crime" scenario is appropriate for the population from which we drew our experimental participants (college students). It was crafted to be analogous to other information analysis tasks, specifically to the Special Operations Reconnaissance scenario developed by the Office of Naval Research.

The three participants in each team were told that each was an expert in a specialized information gathering field and that they had used their expertise to gather information to help police solve a crime. They were given briefing information to explain their role in the team's problem solving, and they were presented with a packet of information gathered for their team role. Team members were told that they needed to work together to solve past thefts and predict a future one. The participants were also told that in order to be successful they had to share information and make decisions together as each held important pieces to the puzzle. Two parallel versions of the scenario were created to ensure the security of solutions to the scenario. The two versions were identical except for the names of people of interest were replaced with a different set of names.

The scenario was divided into three parts, each consisting of a specific decision making task. Participants had a set amount of time in which to complete each decision-making task: 50 minutes for part one, 45minutes for part two, and 30 minutes for part three. At the end of each task, participants needed to come to a joint decision and write down a team recommendation or answer for that phase.

The three parts of the scenario and the solutions to each of the task was constructed considering the three aspects of crime: motives, means, and opportunities. The participants were expected to solve the tasks following the common steps of crime investigation, obtained from our search on crime investigation literature and interviews with subject matter experts (e.g., local police). In the first part of the scenario, the teams were asked to narrow down a list of 26 persons of interest to a list of the eight most likely suspects (8 points). The solutions were designed as the eight persons of interest who were near the two of the crime scenes when the thefts occurred (opportunity). In part two, participants need to identify thieves for each of four thefts (4 points), the instigators of each theft (8 points), motives for stealing the laptops (4 points), and whether there were connections among the four thefts (1 point). Regarding the solutions to this part of the scenario, the thieves were the persons of interest who either (1) were near the crime scene at the time of the theft and had motives based on the social relationships given, or (2) were near the crime scene at the time of the theft and had suspicious relationship with someone else who might have motives. The instigators were the persons of interest who had motives and were related to someone near the crime scene (means and motives). Finally, in part three, participants were asked to predict the date, time, and location of a future theft (3 points). To solve this part of the task, participants needed to identify the people whose schedule overlapped with the victim and possessed a map with the location where the victim was during the overlapping time. Participants were given additional information at the beginning of part two and three. Information received in previous task(s) can be used for later task(s).

Upon arrival, each participant was randomly assigned one of the three roles (Interview Analyst, Web Analyst, or Records Analyst). Participants play the same role throughout the three-part experiment. The 222 pieces of information were equally distributed among the roles. Each team member was provided with unique information that other specialists lack. For example, records analyst had information on bank transactions, receipts, and class schedules, Web analyst had information gathered from Facebook, Twitter, Ebay, and other online resources, and Interview Analyst had information gathered from questioning persons of interest and people they know, and tailing persons of interest in order to determine regular routines or contacts. In each part of the scenario, the three participants were provided with information particular to their role and were told that only by integrating information and resources will they be able to provide accurate recommendations, or in other words, successfully complete each task.

Procedures and Data Collection. Participants responded to a pre-experiment background survey online before participating in the lab session. Upon arrival at the lab, they were instructed to pick one of the three seats around the table. In front of each seat were Mission Statement and General Instruction, phase one role document, a questionnaire, a laptop computer, and tools (e.g., notepads, post-its, pens, markers, highlighters, and rulers) that they can use to analyze the information provided. The Mission Statement and General Instruction given to each role were identical. It contained the goals of each phase as well as information about the four thefts that occurred on campus (date, time, location, and victim). Part one role documents contained information unique to each role in part one. The laptop contained electronic copies of the role documents given to each role. Participants were instructed to use the laptop computers only for search function, reading the documents, and responding to the surveys administered throughout the study.

Participant were given ten minutes to read the Mission Statement and General Instruction and phase one role document without talking to other team members. A questionnaire asking each participant to write down suspicious people of interest and their rationale for selection was given at the beginning of the task. Participants were instructed to respond to the questionnaire while reading the documents. Five extra minutes were given for completing the questionnaire when needed.

After reading the documents and responding to the questionnaire, participants were asked start the conversation by introducing their roles and the type of information they have to each other as well as the suspects that they wrote down.

Task specific role documents as well as team answer sheets were handed to the participants by experimenter at the beginning of each part of the scenario. Participants were asked to write down their answers to questions on the answer sheet provided. After participants wrote down their answers as a team, each team was provided with correct answers and five minutes to reflect on the correct answers. The goal of the reflection session was to bring each team to similar levels of knowledge before starting the next part of the experiment.

Several surveys were administered during the lab session. After their completion of phase one, participants answered a confidence survey in which participants were asked to rate their confidence on each of the 26 people of interest being the actual thief. After phase two, a mid-experiment survey containing questions about team processes (e.g., psychological safety and team cognition) was administered. Finally, after phase three, a post-survey was administered. The post-survey contains measures of team processes and outcomes. After responded to the post-experiment survey, participants were debriefed and thanked. These surveys will provide external validation for the corpus coding efforts.

Selection of High- and Low- Performing Teams. In order to accurately categorize and assess representations it was necessary to understand the context surrounding their creation and the individual's motives and rationales for creating them. For this reason we decided to select ten teams for microanalysis to transcribe and examine as we analyzed their group artifacts, i.e., their collaborative visual representations. Rather than simply picking our teams based on average performance across the three parts of the scenario, we decided to prioritize consistent performance, i.e., how consistently they remained in the top (high performing) or bottom (low performing teams) across the three parts of the scenario. In this way it would be more likely to select teams whose patterns of interaction might be more consistent. So we designated each team with an "H", high, or "L", low, label for each part of the scenario; mid performing teams remained blank. We then selected teams with the most "H" designations for our high performing teams and teams with the most "L" designations for our low performing teams. In cases where more than one team had the same frequency of "H" or "L" designations across the three parts of the scenario, average percentage accuracy across all three parts was used as a secondary filter.
For the selection of high-performing teams, three teams received three "H" designations (Team 2, 7, and 13) and two received two "H" designations (Team 14 and 15). Therefore, these five teams were selected as high-performing teams. None of these five consistently high performing teams received an "L" designation in any of the three phases.

For the selection of low-performing teams, four teams received two "L" designations (Team 1, 17, 18, and 21) and were selected as part of the consistently low performing teams. The last low-performing team needed to be selected from the ten teams that received one "L" designation. Using average percentage accuracy as a secondary filter, Team 8 was selected as the fifth lowest performing team because it had the lowest average percentage accuracy across the three parts of the scenario (see figure 3 for graphical representation of performance ranges of the ten selected teams).

The Penn State group's proposed work focuses on identification of different leadership styles:

*Identification of leadership style differences between high and low accuracy teams***.** We will use coded transcripts of the ten selected teams to determine leadership styles for each team, i.e., patterns of management around orchestration of the task and decision-making. We will then see whether there are any relationships between particular leadership styles or patterns of interaction and team performance. The aim of this research is to identify specific behaviors that may be associated with breakdowns in team problem solving. In this way we can red-flag teams before they experience major breakdowns and develop ways to correct problematic behaviors.

Analysis of individual characteristics and relationships with particular leadership styles. We will use individual background surveys in order to determine whether relationships exist between individual characteristics, team composition, and leadership styles. The aim of this research is to identify if particular team compositions may be more prone to problems, conflict, error, or leadership issues.

As always, the analysis work from the Penn State team will go hand in hand with the computational work under Task 1, with the goal of supporting analysis with existing technology, driving the development of new technology by needs of analysts, and automating as much of the analysis as possible.

Project Schedule and Milestones

	Year 1	Year 2	Year 3
Carnegie Mellon University	* Coding Manual for three dimensional analysis scheme for social positioning * Annotated version of Leadership corpus using Social positioning coding manual * Results applying domain adaptation techniques (Joshi & Rosé, submitted) and feature evolution techniques (Mayfield & Rosé, 2010) to C4 compiled by Aptima	* Report related to integration of analysis of leadership from Behavioral team. * Report correlating leadership coding from conversational level with CKI summative measures * Updated coding manual for three dimensional analysis scheme for social positioning (validated with Rentsch and McComb assessment instruments and reliability ratings) * Preliminary results applying leadership and macrocog coding schemes to APAN dataset	* Release of software and tutorial for technology supported coding for CKI community * Publication related to automated analysis of APAN * Publication related to automated analysis of emergence of leadership in ad hoc teams * Publication related to new technology for domain adaptation, evaluated on Combined corpus
Drexel University	* Mapping and integration of coded data from three corpora into Drexel's interaction data warehouse * Integrate measures of group identity and knowledge construction derived from existing corpora.	* Completed identification of structural and semantic patterns related to leadership behavior. *Coding of three corpora for knowledge construction and group identity, using 2 raters. *Comparison of leadership patterns with existing corpora	* Publications related to leadership and group knowledge construction and group identity behaviors from the three corpora. * Publications related to visualization of knowledge construction and group identity indicators. * Release of network analysis and leadership identification scripts and development of training materials.

Drexel University	* Mapping and integration of coded data from three corpuses into Drexel's interaction data warehouse, which includes measures of group identity and knowledge construction.	* Completed identification of structural and semantic patterns related to leadership behavior. *Coding of three corpora for knowledge construction and group identity, using 2 raters. *Comparison of leadership patterns with existing corpora i	* Publications related to leadership and group knowledge construction and group identity behaviors from the three corpora.
Aptima	* Development of C4 database * Results applying Aptima analysis technology to C4	* Results applying Aptima analysis technology to APAN dataset * Results applying technology to classified datasets	* Visualization of analysis of APAN data
Ohio State University	* Leadership Corpus consolidated from CMU, Drexel, Ohio State, and Penn State corpora	* Completed recoding of existing corpus to calculate inter-rater reliability of leadership indicators	* Support writing of publications * Lead publication related to leadership indicators
Penn State			* Identification of leadership style differences between high and low accuracy teams. * Analysis of individual characteristics that correlate with particular leadership styles. * Publication related to findings from year two with implications for designing activity awareness tools.

Reports

The work effort defined in the tasks will result in deliverable products of quarterly technical and financial progress reports and a final report.

Management Approach

This proposal brings together an interdisciplinary team of researchers from both qualitative and quantitative approaches to behavioral research as well as expertise in social network analysis, machine learning and language technologies to address theoretically interesting scientific questions related to leadership and group functioning in ad hoc teams while making a very practical contribution to the CKI community in terms of technological support for coding corpora.

PI Dr. Carolyn Rosé has been actively involved in research on robust processing of natural language for almost two full decades and has nearly 100 peer reviewed publications in this area. Furthermore, her team has been at the forefront of work on dynamic support for collaborative learning using this technology for automated, real time analysis of collaborative behaviors for triggering support in a context-sensitive manner, and have received a variety of awards and award nominations for this work at international conferences such as ACM SIG-CHI, Computer Supported Collaborative Learning, Artificial Intelligence in Education, and Intelligent Tutoring Systems. Her research group has produced two tool sets for supporting easy application of text mining technology, which cumulatively have over 3,000 downloads from over 75 countries and have produced an architecture for dynamic support for collaborative learning that can be easily tailored for use with alternative automatic analysis techniques and alternative support agent technologies. It can easily be integrated with a variety of collaboration environments, typically with only a couple of days of work.

CoPI Gerry Stahl is a leading researcher and theoretician in computer-supported collaborative learning (CSCL). He has presented at every CSCL conference and founded the *International Journal of CSCL*. Trained in computer science, human-computer interaction, artificial intelligence, cognitive science and philosophy, he is a tenured Associate Professor at the College of Information Science & Technology, Drexel University.

CoPI Dr. Sean Goggins is an Assistant Professor at Drexel University's iSchool. His research is focused on the identification and description of group development in real world, socio-technical environments. His work integrates network analytic techniques with ethnography and case study methods. Dr. Goggins research questions center on the knowledge creation abilities, information use, identity development and structure of technologically mediated groups. Dr. Goggins has applied his integrated methods to the study of groups in software engineering, education, health care and small business networks. His publications are available in competitive conference proceedings of the ACM, IEEE, and ICLS. He currently has papers under review, which further explicate his emerging, mixed methods approach to describing and understanding completely online group development. Dr. Goggins analyzes data in existing ONR corpora, including VMT, on a project with Dr. Rose' and Dr. Stahl. Dr. Goggins is a member of the Group Cognition lab at Drexel, and conducts his research primarily at the small group unit of analysis.

CoPI Dr. Emily Patterson is an Assistant Professor at The Ohio State University. Her professional focus is in applying human factors knowledge methods to improve the design of complex, socio-technical settings, and particularly in human-computer interaction and computer-supported cooperative work. She has served as Principal Investigator on a number of federal grants and contracts, including for the Office of Naval Research, the Air Force Research Laboratory, and intelligence agencies. In 2004, Dr. Patterson received the Alexander C. Williams, Jr., Design Award from the Human Factors and Ergonomics Society. Dr. Patterson currently serves as Associate Editor for IEEE Systems, Man, and Cybernetics Part A, and formerly on the Editorial Board for the Human Factors journal. She has conducted a number of service activities to champion more effective national policy-making, including improving the rigor of intelligence analysis. Dr. Patterson holds a PhD in Industrial and Systems Engineering from The Ohio State University and is a member of the Human Factors and Ergonomics Society.

CoPI Dr Marcela Borge and **CoPI Dr John Carroll** are on the faculty at the Laboratory for Computer-Supported Collaboration and Learning (CSCL) in the College of Information Sciences and Technology at Penn State, which addresses a wide range of challenge areas in which people collectively and individually use information technology to learn and solve problems. Our work currently focuses on software and information design, end-user programming and design, design rationale, creativity in design, training and instructional design, case-based learning and collaborative learning, open source software, e-science, web-based collaborative systems, online communities, wireless community networks, decision support, support for information analysis, community health applications, geospatial information systems, equity and access to computing and information technology, usability engineering methods, and theories in human-computer interaction.

CoPI Dr. Andrew Duchon is a Distinguished Scientist on the Communications Dynamics Team in Aptima's Analytics, Modeling and Simulation Division. He is a project manager, researcher, and software engineer applying statistical natural language processing to projects requiring the analysis of text or verbal information to gain better understanding of a situation, for example, team communications to assess performance, news trends to forecast terrorist activities, and resumes to choose which experts should undertake a mission. He has authored four patents pending related to this work. Prior to Aptima, Dr. Duchon was a co-founder and Director of Technology at Simpli.com, a semantically based search engine. He has also worked as a private consultant developing statistical models for the finance industry, an NSF postdoctoral fellow at Brown University studying virtual reality, and a researcher at the Advanced Telecommunications Research Institute in Japan developing neural networks for speech recognition. Dr. Duchon holds a Ph.D. in Cognitive Science from Brown University, and a B.A. in Psychology and the

Integrated Science Program from Northwestern University. He is a member of the Association for Computational Linguistics.

Carnegie Mellon University will be the lead institution, and Dr. Rosé will take the lead in coordinating the effort across four academic institutions and one industrial partner.

Technology research related to automated coding and social network analysis will take place at Carnegie Mellon University, Drexel University, and Aptima. The automated analysis team will coordinate by phone meetings twice a month. Behavioral research including both qualitative investigations and quantitative experimental studies will take place at Ohio State University and Penn State with input from Carnegie Mellon University. The Behavioral Research team will also conduct phone meetings twice a month. The whole team will have a phone meeting together every two months and an in person whole group meeting once per year. In order to raise awareness in the broader research communities that the PIs are actively engaged in, the team will organize at least one workshop at an international conference sometime during the three year term of the project.

All human data to be used in the project will have been collected in a manner approved through IRB review at all collaborating institutions. Previously collected data from other CKI projects have already been approved by the IRBs at the institutions where it was collected. We will obtain permission from our institutions and from the institutions where the data was originally collected in order to ensure that the highest ethical standards of data handling are met.

Impact of Proposed Work

The proposed project will have impact in two areas. First, it will result in development and testing of analysis methods, automated tools, dynamic models and empirically grounded theory for the understanding of group processes of macrocognition (aka group cognition) in ad hoc teams confronted by non-standard problems. Second, it will result in implemented interventions leveraging this analysis technology to support optimization of macrocognitive group processes.

ONR Mission Relevance

The proposed project is directly responsive to the ONR CKI Program focus on analyzing group processes involved in team decision making in tactical teams, and especially the current focus on increasing efficiency in analysis of group processes through automation and using such technology to optimize group processes through technological support. The project develops tools for analyzing, theorizing and modeling group processes involved in team decision making in small ad hoc groups collaborating on complex problem exploration/analysis/solving.

Bibliography

Ai, H., Kumar, R., Nguyen, D., Nagasunder, A., Rosé, C. P. (2010). Exploring the Effectiveness of Social Capabilities and Goal Alignment in Computer Supported Collaborative Learning, in *Proceedings of Intelligent Tutoring Systems.*

Ai, H., Sionti, M., Wang, Y. C., Rosé, C. P. (2010). Finding Transactive Contributions in Whole Group Classroom Discussions, in *Proceedings of the International Conference of the Learning Sciences.*

Arguello, J. & Rosé, C. P. (2006). Museli: A Multi-source Evidence Integration Approach to Topic Segmentation of Spontaneous Dialogue, *Proceedings of the North American Chapter of the Association for Computational Linguistics*

Arora, S., Joshi, M., Rosé, C. P. (2009). Identifying Types of Claims in Online Customer Reviews, *Proceedings of the North American Chapter of the Association for Computational Linguistics*

Arora, S., Mayfield, E., Rosé, C. P., & Nyberg, E. (2010). Sentiment Classification using Automatically Extracted Subgraph Features, *Proceedings of the NAACL HLT Workshop on Emotion in Text*.

Baker, R.S., Corbett, A., Koedinger, K., Roll, I. (2005) Detecting When Students Game The System, Across Tutor Subjects and Classroom Cohorts. *Proceedings of User Modeling 2005*, 220-224.

Baker, R.S.J.d. (2007) Modeling and Understanding Students' Off-Task Behavior in Intelligent Tutoring Systems. *Proceedings of ACM CHI 2007: Computer-Human Interaction*, 1059-1068.

Bakhtin, M.M. (1986) Speech Genres and Other Late Essays. Trans. by Vern W. McGee. Austin, University of Texas Press.

Blincoe, K., Valetto, G., & Goggins, S. (2011). *Leveraging Task Contexts for Managing Developers' Coordination*. Proceedings from ACM International Conference on Software Engineering, Hawaii.

Cress, U.& Kimmerle, J. (2008). A systematic and cognitive view on collaborative knowledge building with wikis, *International Journal of Computer Supported Collaborative Learning* 3:105-122.

Cui, Y., Chaudhuri, S., Kumar, R., Gweon, G., Rosé, C. P. (2009). Helping Agents in VMT, in G. Stahl (Ed.) *Studying Virtual Math Teams*, Springer CSCL Series, Springer.

Efron, B., Gong, G. (1983) A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation. *The American Statistician, 37* (1), 36-48.

Fuxin Li, Jian Yang, Jue Wang (2007). A Transductive Frameowrk of Distance Metric Learning by Spectral Dimensionality Reduction, in *Proceedings of the 24th International Conference on Machine Learning.*

Goggins, S., Laffey, J., & Mascaro, C. (2011). Context-Aware Event Log Analysis in CSCL: Emergent Grammars, Social Networks and Knowledge Construction. *International Journal of Computer Supported Collaborative Learning*, *In Development*.

Goggins, S., Laffey, J., Galyen, K., & Mascaro, C. (2011). Group Awareness in Completely Online Learning Groups: Identity, Structure, Efficacy and Performance. *Journal of Computer Mediated Communication*, *In Preparation*.

Goggins, S., Galyen, K., & Laffey, J. (2010). *Network Analysis of Trace Data for the Support of Group Work: Activity Patterns in a Completely Online Course*. Proceedings from ACM Group 2010, Sanibel Island, FL.

Goggins, S., Laffey, J., & Gallagher, M. (2011). Completely Online Group Formation and Development: Small Groups as Socio-Technical Systems. *Information Technology & People*, *under review*.

Goggins, S. P., Laffey, J., Amelung, C., & Gallagher, M. (2010). *Social Intelligence In Completely Online Groups*. Proceedings from IEEE International Conference on Social Computing, Minneapolis, MN.

Gweon, G., Jeon, S., Lee, J., Rosé, C. P. (in press). Diagnosing Problems in Student Project Groups, in Puntambekar, S., Erkens, G., Hmelo-Silver, C. (Eds.) *Analyzing Collaborative Interactions in CSCL: Methods, Approaches and Issues*, Springer.

Halliday, M. A. K. (1994). *An Introduction to Functional Grammar*, Edward Arnold

Howley, I., Mayfield, E. & Rosé, C. P. (to appear). Linguistic Analysis Methods for Studying Small Groups, in Cindy Hmelo-Silver, Angela O'Donnell, Carol Chan, & Clark Chin (Eds.) *International Handbook of Collaborative Learning,* Taylor and Francis, Inc.

Hyland, K. (2000). *Disciplinary Discourses: Social Interactions in Academic Writing*, The University of Michigan Press

Joshi, M. & Rosé, C. P. (2007). Using Transactivity in Conversation Summarization in Educational Dialog. *Proceedings of the SLaTE Workshop on Speech and Language Technology in Education*

Joshi, M., Wang, Y., Rosé, C. P., Wilkerson, J. (2008). A Needs Analysis for Support in a Game Based Learning Environment, *Proceedings of the International Conference of the Learning Sciences*

Joshi, M. & Rosé, C. P. (2009). Generalizing Dependency Features for Opinion Mining, *Proceedings of the Association for Computational Linguistics*

Joshi, M. & Rosé (submitted). Target-driven Efficient Semi-supervised domain adaptation, submitted to *Empirical Methods in Natural Language Processing*.

Kang, M., Chaudhuri, S., Kumar, R., Wang, Y., Rosé, E., Cui, Y., Rosé, C. P. (2008). Supporting the Guide on the SIDE, in *Proceedings of Intelligent Tutoring Systems (ITS '08)*

Martin, J. R. & White, P. R. (2005). *The Language of Evaluation: Appraisal in English*, Palgrave

Martin, J. R. & Rose, D. (2007). *Working with Discourse: Meaning Above the Clause*, Continuum

Mayfield, E. & Rosé, C. P. (2010). An Interactive Tool for Supporting Error Analysis for Text Mining. In *Proceedings of the North American Chapter of the Association for Computational Linguistics*

Mayfield, E. & Rosé, C. P. (2010). Using Feature Construction to Avoid Large Feature Spaces in Text Classification, in *Proceedings of the Genetic and Evolutionary Computation Conference*

Nguyen, D., Mayfield, E., & Rosé, C. P. (2010). An analysis of perspectives in interactive settings, in *Proceedings of the KDD Workshop on Social Media Analytics*.

Rosé C. P., & VanLehn, K. (2005). An Evaluation of a Hybrid Language Understanding Approach for Robust Selection of Tutoring Goals, *International Journal of AI in Education 15(4).*

Rosé, C. P., Wang, Y.C., Cui, Y., Arguello, J., Stegmann, K., Weinberger, A., Fischer, F., (2008). Analyzing Collaborative Learning Processes Automatically: Exploiting the Advances of Computational Linguistics in Computer-Supported Collaborative Learning, submitted to the *International Journal of Computer Supported Collaborative Learning 3(3), pp237-271*

Sionti, M., Ai, H., Rosé, C. P., Resnick, L. (in press). A Framework for Analyzing Develpoment of Argumentation through Classroom Discussions, in Niels Pinkwart & Bruce McClaren (Eds.) *Educational Technologies for Teaching Argumentation Skills*, Bentham Science.

Veel, R. (1999). Language, Knowledge, and Authority in School Mathematics, in Christie, F. (Ed)*Pedagogy and the Shaping of Consciousness: Linguistic and Social Processes*, Continuum.

Wang, H. C., Rosé, C.P., Cui, Y., Chang, C. Y, Huang, C. C., Li, T. Y. (2007). Thinking Hard Together: The Long and Short of Collaborative Idea Generation for Scientific Inquiry, *Proceedings of Computer Supported Collaborative Learning*

Wang, Y. C., Joshi, M., Cohen, W., Rosé, C. P. (2008). Recovering Implicit Thread Structure in Newsgroup Style Conversations, *Proceedings of Computer Supported Cooperate Work*, Note

Wang, Y. C., Joshi, M., Cohen, W., & Rosé, C. P. (2008). Recovering Implicit Thread Structure in Newsgroup Style Conversations, *Proceedings of the International Conference on Weblogs and Social Media,* Seattle, WA.

Wang, Y. C. & Rosé, C. P. (2010). Making Conversation Structure Explicit: Identification of Initiation-response Pairs in Online Discussion, In *Proceedings of the North American Chapter of the Association for Computational Linguistics.*

Wenyuan Dai, Ou Jin, Gui-Rong Xue, Qiang Yang, Yong Yu (2009). EigenTransfer: A Unified Framework for Transfer Learning, in *Proceedings of the 26th International Conference on Machine Learning.*

Witten, I. H. & Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques*, second edition, Elsevier: San Francisco