Technological Affordances for Productive Multivocality in Analysis

Gregory Dyke, Carnegie Mellon University, USA, gregdyke@gmail.com
Kristine Lund, CNRS, University of Lyon, France, Kristine.Lund@univ-lyon2.fr
Heisawn Jeong, Hallym University, Korea, heis@hallym.ac.kr
Richard Medina, University of Hawaii, USA, rmedina@hawaii.edu
Daniel D. Suthers, University of Hawaii, USA, suthers@hawaii.edu
Jan van Aalst, University of Hong Kong, China, vanaalst@hkucc.hku.hk
Wenli Chen, Nanyang Technological University, Singapore, wenli.chen@nie.edu.sg
Chee-Kit Looi, Nanyang Technological University, Singapore, cheekit.looi@nie.edu.sg

Abstract: Productive multivocality in CSCL has been the focus of a series of workshops involving the comparison and contrasting of multiple analyses of the same datasets, with the goal of learning how different epistemologies and analysis methods of collaborative learning can complement each other and allow a more complete understanding to emerge. A prerequisite to such work is the technological ability to assist the comparison of different analyses. In this paper, we show how the Tatiana framework for manipulating analytic representations was used to compare three different analyses of a computer-mediated small group problem solving session. In particular, we draw conclusions as to the technological affordances that are needed to ensure productive multivocality and illustrate the immediate benefits provided by the Tatiana framework.

Introduction

Researchers, designers and practitioners in the Computer Supported Collaborative Learning community work within a variety of epistemological frameworks, have multiple theoretical perspectives and use a variety of tools to apply different analytical methodologies to complex datasets. Such multivocality is a strength only to the extent that boundary objects can be created and shared in order to support dialogue between disciplines and approaches. Through a series of five workshops over four years, we have attempted to address the issue of what we call productive multivocality, particularly in the analysis of small group interaction in learning situations. Over the series of workshops (Suthers et al., 2011), we have come to realize that discussions surrounding these issues need to be grounded on multiple analyses of the same dataset and that, moreover, few differences can be addressed without diving back to the primary data and examining each analyst's approach.

Such comparison of analyses is hard to perform without adequate tool support and extends the need, already highlighted by many authors, for tool-support for analysis (Sanderson & Fischer, 1994; Hilbert & Redmiles, 2000; Suthers & Medina, 2008). However, most tools remain limited to certain kinds of media or certain kinds of analyses (e.g, Greenhalgh, French, Humble, & Tennent, 2007; Fiotakis, Fidas, & Avouris, 2007; Georgeon, Mille, & Bellet, 2006). Our ability to combine analyses through interoperating tools is further contingent on the design of models of analysis (Harrer et al., 2007). This ability is not only desirable within the context of productive multivocality but is necessary if researchers are to be able to analyse as a team (Goodman, Drury, Gaimari, Kurnland, & Zarella, 2006) and in order to validate and replicate existing studies (Reffay, Chanier, Noras, & Betbeder, 2008).

Suthers, Dwyer, Medina, and Vatrapu (2010) describe a generic analytic framework for computer-mediated interactions. They explain that such frameworks can be broken up into three foundations: an empirical foundation describes the real-world observable objects the data is composed of (typically events and relationships between events); a representational foundation describes how the empirical foundation is modeled and subsequently visualized (e.g. as a table, as a graph, etc.); a conceptual foundation describes the mapping of features of the other two foundations onto epistemological concepts (e.g. ideas and uptake). The Tatiana framework (Dyke, Lund, & Girardot, 2009) provides an extensible data model and a typology of operations from which a variety of representational foundations can be generated, resulting in the Tatiana tool (based on this framework) being generic enough to afford the creation and combination of a variety of analytic representations.

In this paper, we describe three analyses on the same dataset to illustrate how the Tatiana framework can be used to create and combine different analytic representations, and to show technological affordances that must be catered for when designing analytic tools. The complexity of this dataset and the diversity of analytical methodologies and representations that were used provide the basis for a discussion about the empirical and representational foundations of these analyses. Discussion of the conceptual or epistemological foundations and the commensurability and compatibility of various analytic approaches would need to be based on the entirety of datasets analyzed over the series of workshops and is beyond the scope of this paper.

In what follows, we first present Tatiana and the core concepts onto which data and analyses are mapped within the tool. We then describe the study from which the data was collected and the form in which it was shared among analysts. Three different analyses were carried out independently from each other with an overarching goal of identifying pivotal moments for learning. For both the data and the three subsequent analyses, we describe the initial analytic representations, their affordances, and the work that was performed, both to map them onto the Tatiana framework and to coordinate them in order to compare them. Finally, we show some examples of the immediate benefits of the Tatiana tool, suggest how it should be improved, and summarize the technological and pragmatic affordances that appear to be a prerequisite to multivocal analysis.

Tatiana: a Generic Analysis Environment

Tatiana (Trace Analysis Tool for Interaction ANAlysts) (Dyke et al., 2009) is an environment (and an underlying conceptual framework) designed for manipulating various kinds of analytic representations, in particular those that present a view on event-based data, be it the original data or subsequent analyses thereof. We call these representations *replayables*, because they can be replayed in a similar fashion to a video. Tatiana replayables comprise a sequence of events and benefit from Tatiana's four core functionalities, whose extensibility differentiates it from other tools: transformation, enrichment, visualization and synchronization (see Figure 1).

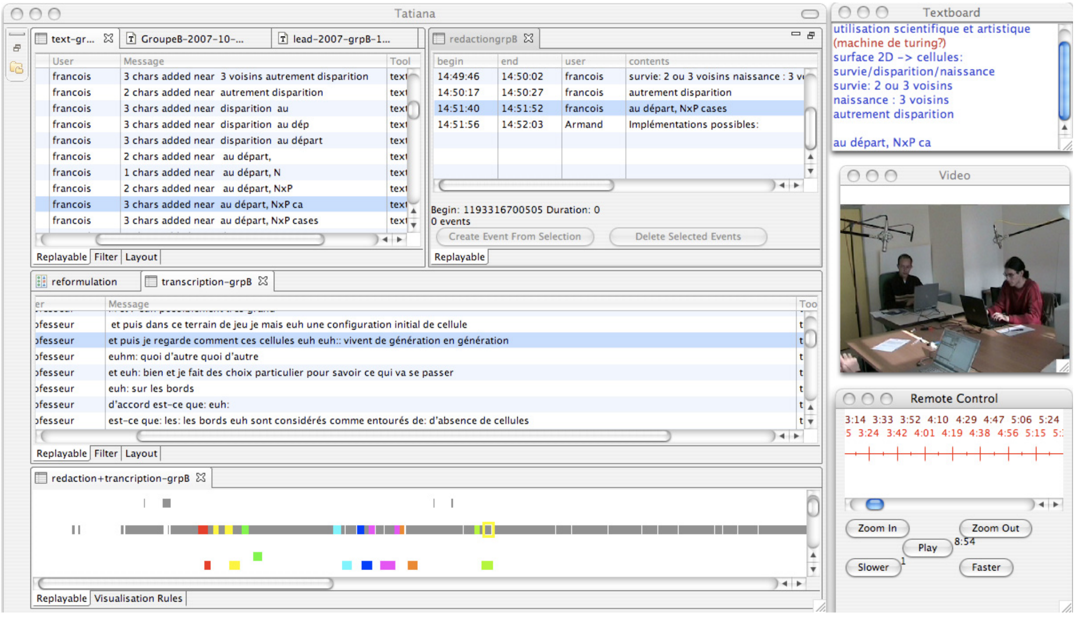

Figure 1. Various replayables visualized in Tatiana: traces of a shared text editor (top left), transcription (middle left), writing units (top center), visualization of reformulation (bottom left), synchronized with external tools, DREW replayer (top right), video player (middle right), remote control (bottom right).

Replayables can be transformed (automatically or manually) to create new replayables whose events might be a subset, an abstraction or a combination of events from other replayables. In analogy to a spreadsheet view of data, transformation can be thought of as creating, copying and deleting rows or events. In the remainder of this paper, we shall refer to events created as a combination or grouping of other events as *parent events*, since they create a parent-child relationship.

Events within Tatiana can be enriched by analytic data generated by the researcher. There are currently three kinds of enrichments supported by Tatiana: free-form annotation, categorization, and graphs. Categorization is simply a way annotating the events from a restricted list of words and can be used for coding, labeling and adding keywords. Graphs allow researchers to explicitly mark relationships between events. Conceptually, enrichments add additional columns (or properties) to events, or add links between events. Tatiana provides an extension mechanism to create new kinds of enrichment.

All replayables within Tatiana can be visualized in different viewers. Two viewers have been implemented so far: a table view in which data is presented as in a spreadsheet, with one row per event and columns for each of the event's properties; and a graphical timeline, which presents each event as a graphical object whose graphical properties (color, shape, size, position, etc.) can be set according to the properties of the event (user, tool, timestamp, category, etc.). Tatiana is extensible, allowing new kinds of viewers to be created, affording new ways of visualizing data.

Finally, all visualizations of replayables in Tatiana can be synchronized with each other and also with data viewed in external replayers such as media players. Synchronized replay means that, for example, during analysis of a video and its transcription in Tatiana, if a researcher clicks on a timestamped utterance in the table view, this action causes the replayer to move the video to this point. Conversely, when the video is replayed, the corresponding events in the transcript are highlighted one after the other. This functionality helps link different replayables together, allowing data to be looked at simultaneously from different angles.

Small Group Discussion about Fractions in Group Scribbles

The dataset we are reporting on was drawn from a three-year school-based research project using Group Scribbles (GS; Roschelle et al., 2007), which provides a shared representational space to support collaborative practices. Two fifth grade classes (about age 11) in a primary school in Singapore participated in the project. Students were asked to work in groups of four to solve the problem of dividing two pizzas equally among three children (Looi & Chen, 2010). They were first asked to work individually, either thinking about the solution or creating their private note, and then share it with the rest of the group.

Each student had an individual tablet PC with GS, which enabled them to create notes with hand-drawn sketches or typed text in their private space and then publish them to the group space. For this activity, the group space was prepared with a picture of two pizzas. Students could draw directly on the pizza (but were encouraged not to) or post notes in the public space. They could also view others' work, attach comments on other members' contributions, or move posts back to the private space.

The shared dataset focused on a group comprising four students whose names have been anonymized as Terry, Helen, Victor and Quentin. The students sat in pairs over two rows. One video camera was set to record the classroom session and another the target group. Each student also had a camera set up on the desk to capture their face. Screen capture software was used to record the contents of students' activities in GS environment. The resulting dataset comprises six videos (one of the classroom, one of the group, and one of each screen with the face video inserted into the bottom right corner). The episode that was analyzed lasted about 16 minutes.

During the episode, Victor and Helen both produce and discuss similar graphical solutions; Terry is initially confused and ends up producing a written solution that is equivalent to the graphical solutions. It is not clear to what extent Terry was unable to come up with any kind of graphical solution and to what extent his solution was influenced by Helen and Victor's solutions and by discussion with Helen. Quentin produces a symbolic arithmetic solution that is initially incorrect but gets amended later on as Quentin realizes his mistake.

Data Transcription and Subsequent Sharing within Tatiana

Transcripts of the videos were provided by Looi, Chen and Tan in a spreadsheet, with approximate time stamps (provided at one minute intervals) and with the activity organised into four columns, one for each student. Simultaneous actions (or near-simultaneous actions for which timing was not considered important by the transcriber) were transcribed in the same row, with actions taking place over a longer period of time spanning several rows. When the ordering of events was important to show potential reply-structures, these events were placed on consecutive rows. Many cells also contained screenshots to help understand the GS activity.

This kind of analytic representation is particularly interesting. It is able to condense the transcription into a reasonably small area (compared to a transcription with one row per utterance). It also accounts for the timing between events, albeit with a certain bias (as the transcriber decides whether precise information about how various actions overlap should be transcribed or not). This account can also be achieved at the relatively cheap cost of not having to find the precise beginning and end date of each event. Furthermore, the presence of screenshots helps to remove the difficulty in understanding the transcription without the aid of the video, making the transcription a remarkably standalone analytic representation. The cost for this is higher, as screen captures must be made and inserted for each event.

One goal of transcription is to remove the necessity of constantly having to replay the video. This comes at the cost of "trusting" that the transcription adequately serves as a proxy for the original video. In order to allow the analysts to view the transcription in synchronization with the videos and thus to alleviate this problem, the transcription was adapted to Tatiana before sharing. This produced four replayables, one for each student and a global replayable created through the automated merging of the four individual replayables. The events in these replayables were time-aligned with the videos, to reduce ambiguity and because Tatiana does not afford partial time-alignment. A basic typology of events was created (speech, note creation, posting a note, etc.) to enable better differentiation between them. Although no visualization in Tatiana allows screenshots to be associated with events, and the tabular view matches each event to a single row, the affordances present in the original transcription can be reproduced in Tatiana by opening in parallel the four replayables and one of the videos. However, as each video only shows the corresponding student's private space, all four videos must be opened to examine all the private spaces. The disadvantage of this mapping is that synchronization is dynamic, and static representations lose these affordances (i.e. it is no longer possible to click on a visualization in a paper and cause something to happen). One workaround is to print out the tabular view of the combined replayable, which at least provides complete timing information, in spite of displaying it in a less intuitive fashion.

Once this mapping into Tatiana was completed, the data was shared, both in Tatiana and in its original form. We now present the three analyses, two of which were performed with Tatiana and a third later adapted.

Analysis from the Perspective of Knowledge Building

Knowledge Building is the work that a community does to advance the state of knowledge in the community, such as creating, improving and using new ideas (Scardamalia, 2002). Knowledge building involves not only sharing ideas with others, but also jointly understanding and improving them. There must be a context in which such work is important, so that the new ideas are gradually used in the later work that the community does. Recent literature on knowledge building emphasizes "knowledge practices" (van Aalst, 2009; Hakkarainen, 2009). The first analysis was conducted from this perspective, entirely within Tatiana.

Because of the short length of the dataset, we did not expect evidence of all aspects of knowledge building – particularly the diffusion of the new ideas in the work of the community (a small group in this case). From the knowledge-building perspective, it is difficult to say whether learning has taken place without seeing ideas reused in multiple contexts, and by multiple students. Nevertheless it is possible to examine whether the efforts of students are limited to sharing ideas, or whether students collaboratively work to improve them and develop shared understanding. This analysis highlighted the fact that there were very few attempts to improve on existing solutions or to relate them (as all correct solutions are in some way equivalent). The problem itself was also not ideal, as it was not motivated by students' knowledge needs and was not very complicated: two of the students quickly identified solutions but did not understand why they should attempt to reconcile them.

Analytic Representations Created

The transcription was coded in Tatiana in a set of codes derived from the knowledge-building principles (Scardamalia, 2002) and some additional codes from a study that focuses on the social aspects of knowledge building (van Aalst, 2009). These codes were not mutually exclusive and each line was coded with all applicable codes. Tatiana's current implementation does not allow for more than one code per event (which might be described as *tagging*), so three enrichments based on the same coding scheme were created and most events were given between one and three codes. A further annotation enrichment was created to provide a narrative description of what was happening at each line and to justify the choice of certain codes.

Analysis of the Uptake of Representational Practices

Medina conducted uptake analysis of the GS dataset using concepts and methods from the uptake analysis framework described in (Suthers, Dwyer, Medina, & Vatrapu, 2010). This analysis was mostly performed within Tatiana. Its focus was twofold. The first part of the analysis showed, through participants' orientation to each other throughout the first few minutes that, in spite of little apparent explicit communication, the group was aware of the GS workspace as a shared medium for whose content they were accountable. In this brief period they also demonstrated an appropriation of a diverse range of modal resources (talk, gesture, and inscription), The second part of the analysis used uptake graphs and information about the context in which the dataset was produced. It showed that (1) the students were operating on representational practices developed in prior classroom activity (2), each student demonstrated a varied interpretation of these practices in the problem setting and (3), interaction between two of the students (Helen and Terry) revealed that, at times, reconciliation of interpretations can be mediated through shared, yet uniquely applied, representational and inscriptional practices (Medina, Suthers, & Vatrapu, 2009).

Analytic Representations Created

Similar to other analyses conducted under this framework, (e.g. Medina et al., 2009), the analytic implementation in the current investigation was multilevel. At an empirical level, the analysis attends to the sequential order of participants' actions in order to identify contingencies between acts and produce a contingency graph that, in turn, serves as a resource for identifying uptake. Identification of uptake relations, based on the empirical evidence provided by contingencies, produces an uptake graph as a derivative of the contingency graph. In the process of creating these graphs, a new replayable was created with its events being either a single act, already present in the original transcription, or a new event, representing a collection of related acts in the original transcript (e.g. when an idea was published in two separate notes, forming a single whole). This replayable added the affordance of being able to see how much time was taken between beginning the crafting of a new idea and its final publication.

Two graph enrichments were created, one for each analysis, making explicit the contingencies between the different acts. The replayable can be visualized on a graphical timeline, with one horizontal line per student. The arrows provided by the graph enrichment show how students re-use and respond to each other's contributions. An additional annotation enrichment was created to provide a narrative of the episode.

Analysis from a Cognitive Perspective of Group Understanding

The last analysis was performed in a spreadsheet and later adapted to Tatiana. The analysis was an attempt to further our understanding of how groups operate as a cognitive unit in artifact-mediated discourse situations (Jeong, Chen, & Looi, 2011). Jeong examined how group understanding developed by tracing contributions to the group workspace and how individual contributions evolved over verbal and non-verbal activities. This was done by splitting the session up into various activity threads around contributions that represent a group of postings about the same core idea. The contributions were further coded in terms of whether they were on-task or off-task, collaborative or individualistic, and made public or kept private.

The main findings were that the group space became crowded with many postings and comments, but still remained fragmented at the end of the activity, as students made little attempt to integrate and consolidate different postings and their relationships to each other. The analyst suggests that this could be remedied through better regulation of the group space and of the student activity. Furthermore, the students' group work was mediated by both verbal and non-verbal exchanges, e.g. a verbal question was answered by an action in GS. Thus, when group work is mediated by artifact construction, non-verbal actions must also be taken into account.

Analytic Representations Created

As with the original transcription, this analysis was particularly interesting from a representational standpoint. The original events were taken and rearranged, so that each event (or rather the cell in the spreadsheet representing this event) was placed in the row corresponding to the contribution thread it belonged to. This extracted the time-student ordering of the data into as many timelines as there were threads, thus enabled the analyst to examine more clearly how each thread developed, how many participants were involved and to distinguish off-task threads and exclusively private events more easily.

This representation can be mapped onto the Tatiana framework in any one of three ways, depending on the desired affordances. First, a replayable can be created in which the events of each contribution thread are collected together into a single parent event. In the Tatiana framework, this is a transformation that abstracts up from *individual events* to higher-level *threads*. Of the three, it is the most visually similar to the original and, via synchronization, allows events to stay in their original context while still showing to which thread they belong. Second, a categorization enrichment could enable the coding of each event according to the parent thread to which it belongs, this makes it easier to visualize thread-belonging as an extra property of the data, rather than as its central property. Third a replayable could be created for each thread, making it easier to consider threads in isolation. In the work done for this paper, only the first (initially) and second (later) was created.

Combining Analyses in Tatiana

The original idea was that the analyses (four enrichments for the first analysis, two enrichments and a replayable for the second and one replayable for the third) could be opened at the same time and visualized in synchronization in order to compare them, thus creating a common boundary object. In practice, this was not immediately possible, for several reasons.

First, the knowledge building categories were applied to all events. GS forces users to first write a note in their private space (producing a first event in the transcription) and then drag it to the public space, (producing a second). The first event would typically be coded "new idea" and the latter "community knowledge". The uptake analysis only considered events in the public space. In order to meaningfully compare the analyses, the knowledge building categories from events in the private space needed to be transferred to their public space counterparts.

Second, the replayable to which the uptake graph was applied contained events that were parents of the events in the original transcription (typically, the events of note creation, editing, publication and verbal explanation would all be grouped together into a single parent event). As enrichments (categories, annotations and links) attach to individual events, the newly created parent event was not aware of the KB categories and the children of the parent event were not aware of the links that had been attached to their parent event. Again, there needed to be some kind of transfer of enrichments for comparison to be possible.

Last, the initial replayable that was created for the group understanding analysis, with one event per thread, also did not afford easy comparison with the other analyses, since it was only possible to see which events belonged to which thread through synchronization.

To solve these problems, all three replayables (the original transcription, the one on which the uptake graph was drawn, and the threads replayable) were opened side by side along with their associated enrichments and synchronized. From these, a fourth replayable was created, which initially contained all the events of the transcription. The following procedure was then applied, event by event:

- The event was deleted (from the fourth replayable, but remained present in the original transcript), unless one of the following conditions was met:
  - It appeared (itself or by proxy via its parent event) in the uptake graph.
  - It was public, and coded otherwise than "off task" in the knowledge building analysis.
  - It was public and did not belong to an "off task" contribution thread.
- If the event was deleted and was private, its knowledge building codes and annotations were transferred to the corresponding publication event.
- If the event was not deleted, it was coded according to the thread it belonged to.

- If the event was the last publication event and child of a parent event in the uptake graph, it was added into a new graph in place of the parent event.

In order to preserve the affordance, provided by the uptake graph, of "seeing" how long a given idea takes to construct, we also then added back in the private events that lead up to a publication. The application of this process resulted in a single replayable containing only the events that had been considered pertinent in at least one analysis. Although "off task" activity, particularly in the regulation of social relations, has often been considered important in CSCL (Baker, Andriessen, & Lund, 2009), none of the three analyses focused on such events, which is why they were mostly ignored. In effect, the publication event of a note was considered as a proxy for its creation and editing on the one hand and for any parent events to which it belonged on the other, in order to conciliate the units considered in the various analyses.

In an attempt to provide a single visualization in which to examine all the analyses (in combination with synchronization with other visualizations), we used Tatiana to automatically produce a graphical timeline based on the various enrichments of the different analyses, an extract of which is shown in Figure 2. In this timeline, time progresses horizontally, each row corresponds to a contribution thread, each color corresponds to a student, squares are speech events, circles are GS-mediated events, small circles are private events, and the largest circles and squares show events which have been coded with the "core" principles of knowledge building (i.e. new ideas, idea diversity, etc.). The majority of the interaction occurs in the shown threads, with the core of the uptake between Helen (red) and Terry (blue). Potential for knowledge building can also be found in threads that do not involve multiple students (e.g. Quentin's symbolic solution in yellow).

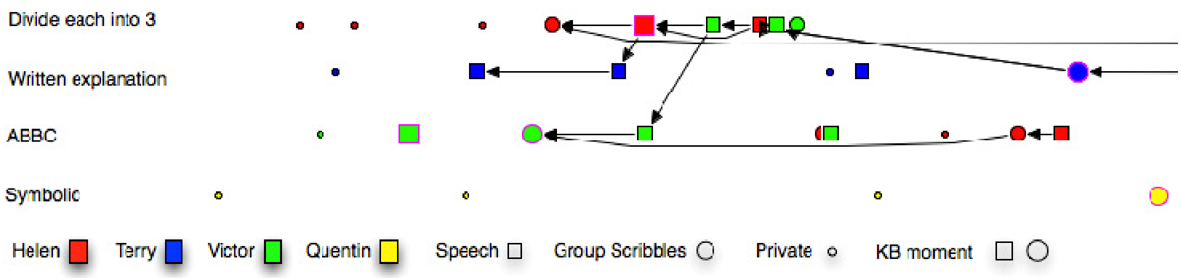

Figure 2. Graphical Timeline Showing the Combined Analyses by Thread. Four of the Nine Threads are Shown.

Benefits of Combined Analyses

Although we do not discuss the lessons learnt from these different analytic interpretations in this paper, this section briefly discusses some of the benefits that are immediately afforded by tools such as Tatiana when embarking on a multivocality or mixed-methods route.

The separation into threads carried the risk of having difficulty in identifying any transitions of ideas between threads. Opening the uptake graph on top of a graphical timeline where each line represents a thread, as in Figure 2, confirms that such transitions do occur. Conversely, the original uptake analysis showed that there was some uptake between Helen and Victor but focused on the importance uptake between Helen and Terry. However, this is not as obvious in the uptake analysis' original graph (cf. Figure 3) where the separation of users on different lines seems to indicate that the main interaction is between Helen and Victor. Separating by contribution thread, rather than by user, immediately focuses our attention on the places where uptake moves from one contribution thread to another, showing the interactions beween various students' solutions and representations. Indeed, the places where arrows cross from one thread to another in Figure 2 were identified by all three analysts as being of interest. This is hardly surprising and confirms the suggestion by Suthers and Medina (2008) that, in process analysis, trajectories of ideas should be examined to look for their transformation (which is an important concept in many epistemologies of collaborative learning and meaning-making).

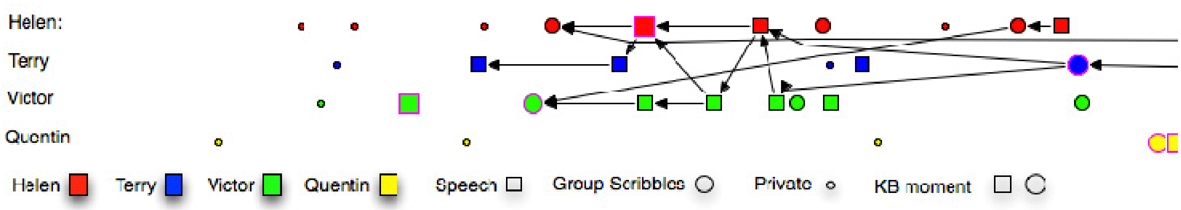

Figure 3. Same Data as in Figure 2 Above, But Organized by User, Rather Than by Thread.

This visualization also immediately highlights some differences between analytic approaches, providing the opportunity to discuss issues concerning the data (uptake between threads, important moments for knowledge building which are "ignored" by other analyses, etc.), which are connected to discussions about epistemological questions in general, grounded on these analyses, such as the relative importance of idea diversity (the mere "presence" of different ideas), collaboration in threads, or uptake across threads. Furthermore, when discussing data points, synchronization allows a return to the original data in the videos in order to verify that interpretations are not influenced by imprecision in the transcription.

Design of Tools to Support the Construction and Use of Analytic Representations

We now draw upon this case study to discuss how tools and their design can support multivocal analytic dialogue. While the previous sections show the Tatiana framework to be an important contribution as a means to reuse and combine analyses in order to create boundary objects for discussion of multiple analyses, we also highlighted some issues, which we hope to address and encourage other analysis tool designers to draw upon.

The biggest difficulty we had in combining these analyses stemmed from the unit of analysis. As explained by Suthers et al. (2010), the events of a CSCL transcript are frequently used as proxies for higher-level concepts. In other words, what is of interest to analysts are not so much events, but, for example, the ideas represented by these events. It is therefore not surprising that the proxy for an idea should be quite variable (initiation of private authorship, editing, publication, or some combination of the three). The issue can be addressed pragmatically, either by discussion after analysis (as exemplified in this paper), or by prior agreement among analysts. In either case, analysts might address what the unit of analysis should be, whether private spaces should be considered, whether the time taken to craft a contribution is important, and whether events in all modes (e.g. speech and tool mediation) should be considered equal.

A complement to this pragmatic approach would be to adjust tool support so that enrichment be transitively propagated to parent and child events under certain conditions. For example, it would make sense for the parent events in the uptake analysis to "inherit" the knowledge building codes attributed to their constituent child events. And in the case of the replayable containing contribution thread events, it would be beneficial to be able to propagate the thread an event belongs to down to the event itself in order to add the thread as a category. In general it should be easy to transition events across such near equivalent features as parent event belongingness, replayable belongingness and having identical properties (same user, same category, etc.).

During the course of this work, Tatiana's synchronization feature had to be refined. Sometimes synchronization is used to identify overlapping or synchronous events (e.g. between speech and artifact-mediated dialog); in other cases it can be used to identify the same event in different visualizations of the same replayable; a hybrid of these two is identifying the children of a parent event from one visualization to another. This only becomes a problem when there are many occurrences of simultaneous events, as is the case with the GS dataset. For this reason, Tatiana has been modified to make it possible to choose the synchronization mode: by timestamp or by identity (the latter includes parent-child event relationships).

The affordances synchronization provides are dynamic and lost in the case of static representations (such as those found in conference and journal publications). In spite of this drawback, there are so few static visual dimensions (size, shape, position and color) that it makes sense to transfer as many aspects as possible to a dynamic register, particularly for the *performance* of the analysis, as opposed to its publication.

Finally, so many analytic representations were created during the analyses and their combination that it would be useful to have meta-data describing their assumptions and purposes, along with software support for recording this meta-data. This is an extension of the recommendation by Reffay et al. (2008), that it is indispensable to include the educational and research context when sharing learning data.

Conclusion

In this paper, we described the lessons learned from the multivocal comparison of three independent analyses of the same dataset, and discussed their implications for the design of analytic tools. The original data set and the three analyses provided a diversity of analytic representations whose affordances were discussed. We then explained how these analyses were combined with the assistance of the Tatiana tool and provided some examples of the benefits of combining analyses in such a way. Finally, we discussed some of the issues raised and provided guidance for development of analytic tools beyond the existing Tatiana framework, focusing in particular on aspects such as analytic unit, propagation at different levels of abstraction, synchronization, and analytic and pedagogical context.

The current Tatiana framework proved a very positive experience as a means to create boundary objects for discussing multivocal analysis. These analyses provide further support for the hypothesis that Tatiana's provisions for generating representational foundations are generic enough to serve as the basis of boundary objects in CSCL. While there appears to be a certain influence of the conceptual foundation of an analysis on its empirical foundation (e.g. what aspect of the data will serve as proxy for an *idea*), this need not be an obstacle to multivocality, assuming analysts agree beforehand on a common empirical foundation (and its mapping onto respective representational and conceptual foundations), and are aware of the choices that might later introduce bias. Furthermore, the use of synchronization, automated transformation and enrichment reduces the need of finding a unique representational foundation for comparing analyses.

We hope to draw on this work to apply similar methods in comparing analyses in other datasets. In particular, we need to extend these methods to address analyses that are more quantitatively oriented. Having diminished the pragmatic representation-driven issues related to multivocality, we hope to be able to focus on methodological and epistemological issues to promote better understanding in the diverse field of CSCL.

References

Baker, M., Andriessen, J., Lund, K. (2009). Socio-relational, affective and cognitive dimensions of CSCL interactions: integrating theoretical-methodological perspectives. *CSCL'09*. Rhodes, Greece.

Dyke, G., Lund, K.,& Girardot, J.-J. (2009). Tatiana: an environment to support the CSCL analysis process. *Proc. CSCL 2009*, Rhodes, Greece, 58–67.

Fiotakis, G., Fidas, C. & Avouris, N (2007). Comparative usability evaluation of web systems through ActivityLens, *Proc. PCI 2007*, Patras, Greece.

Georgeon, O, Mille, A, & Bellet, B.T. (2006). ABSTRACT: a tool and a methodology for analyzing human activity mediated by a complex technical artifact. *Ingéniérie des Connaissances. Semaine de la connaissance.* Nantes, France.

Goodman, B. A., Drury, J. Gaimari, R. D., Kurland, L., & Zarrella, J (2006) Applying User Models to Improve Team Decision Making. Retrieved April 10, 2008 from http://mitre.org/work/tech_papers/tech_papers_07/06_1351/

Greenhalgh, C., French, A., Humble, J. and Tennent, P. (2007) Engineering a replay application based on RDF and OWL, *Online Proceedings of the UK e-Science All Hands Meeting 2007*, September 10-13, Nottingham: NeSC/JISC.

Hilbert, & Redmiles. (2000). Extracting usability information from user interface events. *ACM Comput. Surv.*, *32*(4), 384–421. doi: http://doi.acm.org/10.1145/371578.371593.

Harrer, A. Zeini, S. Kahrimanis, G. Avouris, N. Marcos, J. A. Martinez-Mones, A. Meier, A. Rummel, & N. Spada, H (2007) Towards a Flexible Model for Computer-based Analysis and Visualisation of Collaborative Learning Activities, *Proc. CSCL 2007*, New Jersey, USA

Hakkarainen, K. (2009) A knowledge-practice perspective on technology-mediated learning. *ijcscl 4*(2), pp. 213-231.

Jeong, H., Chen, W., & Looi, C. K. (2011). Analysis of group understanding in artifact-mediated discourse. *CSCL 2011*. Hong Kong.

Looi, C.-K., & Chen, W. (2010). Community-based individual knowledge construction in the classroom: A process-oriented account. *Journal of Computer Assisted Learning*, 26(3), 202-213.

Medina, R., Suthers, D., & Vatrapu, R. (2009). Inscriptions becoming representations. *CSCL'09*. Rhodes, Greece.

Reffay, C., Chanier, T., Noras, M., & Betbeder, M.-L. (2008) Contribution à la structuration de corpus d'apprentissage pour un meilleur partage en recherche, *STICEF journal*, Volume 15, ISSN : 1764-7223. Put online 14th October 2008, http://sticef.org.

Roschelle, J., Tatar, D., Chaudhury, S. R., Dimitriadis, Y., Patton, C., & DiGiano, C. (2007). Ink, improvisation, and interactive engagement: Learning with tablets. Computer, 40(9), 38-44.

Sanderson P., & Fisher C. (1994). Exploratory sequential data analysis: foundations. *Hum.-Comput. Interact.*, *9*(4), 251–317.

Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge. In B. Smith (Ed.) *Liberal education in a knowledge society* (pp. 67-98). Chicago: Open Court.

Suthers, D. D., Lund, K., Dyke, G., Rosé, C., Law, N., Teplovs, C., et al. (2011). Towards productive multivocality in the analysis of collaborative learning. *CSCL 2011*, Hong Kong.

Suthers, D. D. & Medina, R. (2008). Tracing Interaction in Distributed Collaborative Learning . *Annual Meeting of the American Educational Research Association* (AERA), New York, March 24-28, 2008.

Suthers, D. D., Dwyer, N., Medina, R., & Vatrapu, R. (2010). A framework for conceptualizing, representing, and analyzing distributed interaction. *ijcscl 5*(1), pp. 5-42.

van Aalst, J. (2009) Distinguishing knowledge-sharing, knowledge-construction, and knowledge-creation discourses. *ijcscl 4*(3), pp. 259-287.

Acknowledgments

The authors would like to thank the organizers and financers of the series of workshops on productive multivocality, as well as the many participants whose contributions helped structure our thoughts on these issues.