Rosé, C.P. & Lund, K. (in press). Methodological Pathways for Avoiding Pitfalls in Multivocality. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions*, Chapter 32. New York: Springer.

Chapter 32

Methodological Pathways for Avoiding Pitfalls in Multivocality

Carolyn Penstein Rosé, Carnegie Mellon University
Kristine Lund, CNRS—University of Lyon

Abstract (for e-Book ONLY!)

This chapter explores multivocality from a methodological perspective. A conceptual model is presented for thinking about multivocality and how it relates to methodological traditions. We reflect back on what we have learned through experimentation with multivocality through the five data sections of the book and draw principles for best practices that we offer to the broader research community. As a running theme throughout the chapter, and as an invitation to disseminate multivocality to the next generation of researchers in our field, we contrast the experience of expert analysts whose work is presented in the preceding data sections with the experience of students working in groups on their first discourse analysis project in the context of a Computational Models of Discourse Analysis (CMDA) class.

Methodological Pathways for Avoiding Pitfalls in Multivocality

Carolyn Penstein Rosé, Carnegie Mellon University
Kristine Lund, CNRS—University of Lyon

Introduction

In his original work on multivocality in literary criticism (Bakhtin, 1981), Bakhtin argued that the novel as a form of literature offers more of an opportunity for multivocality than other narrowly and rigidly defined forms. In this spirit, we offer a perspective on research methodology that conceptualizes the analytic cycle in such a way that provides the opportunity for multiple perspectives to speak to one another and challenge one another as we examine data that are of common interest. The earlier chapters of this book have illustrated our own journey towards multivocality and have served the purpose of illustrating potential outcomes of productive multivocality. In Chapter 34 (Lund, Rosé, Suthers & Baker, this volume), we explore the epistemological encounters that researchers had when they compared various aspects of their analyses. The perspectives of different researchers may either coexist in their natural, productive tension without being integrated, and thus remain limited with respect to their value to one another; or they may actively interact on theoretical and/or methodological levels, thus bringing new insight to the phenomenon being studied.

In this chapter we invite the reader to join us on our journey towards mutlivocality while we focus on the methods of analysis of collaborative interactions. We assume a diverse readership that may include expert analysts, steeped in their own tradition, who want to forge new partnerships to embark on their own multivocal experience. Other readers may be students just learning about research methods in order to get a better grasp on the research landscape. Still other readers may be instructors of research methods who may be looking for ideas for how to use multivocality as a teaching paradigm in those courses. In all these cases, it is important to get insight into what makes multivocality challenging. Our goal is to offer these insights. In particular, we will examine what might be considered the "dark side" of the multivocal analysis process. We do not want to present an unrealistically rosy view of our own experiences for those who will follow in our footsteps. So here we discuss potential pitfalls of the multivocal analysis process and what might be some pathways towards working around them or avoiding them altogether.

First we place this discussion within a conceptual frame. Simply put, the important running theme is that multivocal research is an intensely team-oriented sport. Pitfalls come from a tendency for researchers to fall into isolation in one way or another in their work. Pathways in this chapter will thus consistently be framed as pressures to overcome these tendencies. Experienced researchers know that research is by its very nature social, just as literary criticism is social, in other words embedded within communities of practitioners. We see this when we consider that science is the accumulation of knowledge and that theories are storehouses for the collection and integration of knowledge gained through empirical investigations. We know that a single focused research contribution by itself, no matter how insightful or high quality, is too narrow to be of significant value. It becomes valuable as it is integrated with the results of other empirical investigations that have either already occurred or occur later and relate back to it as seminal work. Despite understanding this important truth, however, we still fall prey to a tendency towards isolation.

In Bakhtin's investigations we see that it is not something inherent in epics or novels that makes one multivocal and the other not; rather it is the way they are treated by communities of practitioners that make them that way. Students learning not only research methods but also the politics and sociology of research may be even more prone to isolation that runs counter to the goals of multivocality. As a running theme, then, we will compare two distinct types of potentially multivocal analysts. At one end of the continuum, we will examine students who are just beginning to engage in research practices and who may thus treat those practices as rigidly defined forms like epics in Bakhtin's explorations. We will compare this orientation with those of more senior researchers within the ecosystem who have gained facility with the practices and have earned a position that allows them to manipulate research practices in creative and productive ways.

Communication Flow of the Multivocal Analysis Process

Now we offer a schema for thinking about our conceptual frame from the perspective of communicative processes. We review the many diverse stakeholders who take part in the multivocal analysis process either directly or indirectly, and analyze the information flow between them, since many of the pitfalls we address in this chapter occur as communication breakdowns or difficulties between stakeholders.

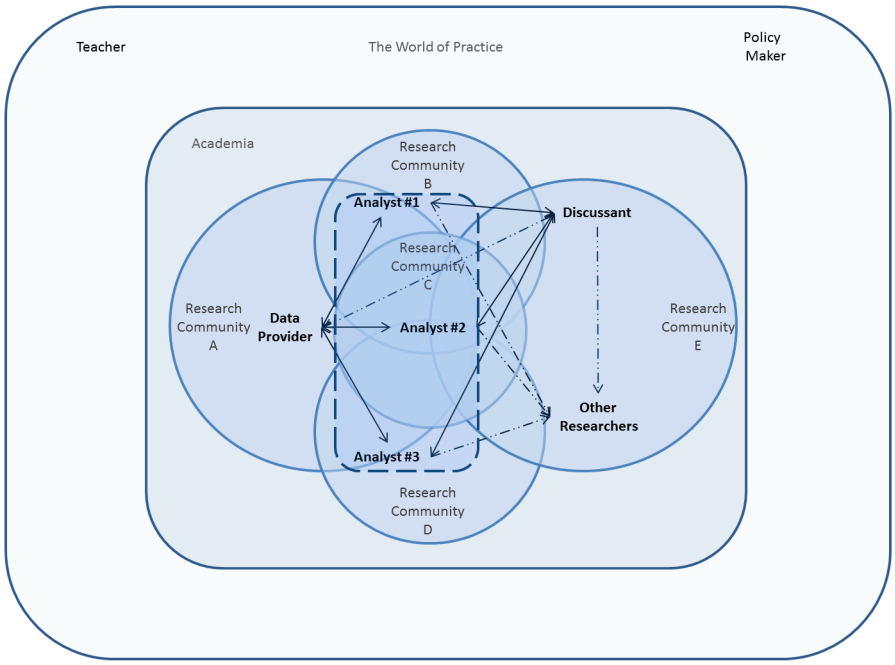

Figure 1. Information flow in an iterative multivocal analysis process. Note that only relatively infrequent formal communication processes are indicated with explicit links. The more frequent and less formal communication and coordination between the analysts is signified by the dashed, oblong shape that joins them together.

Most of the process of multivocal analysis occurs within the purview of Academia. Thus, in Figure 1 all of the direct communication links, which are represented as unidirectional (dotted lines) or bidirectional links (solid lines), are within that area. However, as discussed in Chapter 35 (Law & Laferriere, this volume), ultimately the hope is that our research will impact the world of Practice, so it is important to consider practitioners as one audience for our work, including both teachers and policy makers. That communication is aided by researchers who span both worlds, such as the authors of Chapter 35.

Communication across stakeholders is the key to fueling the iterative multivocal process. In Figure 1, informal processes of comparison between analysts are not displayed as lines. Instead their more frequent communication with one another is indicated by the oblong shaped dashed box that illustrates their joint status as an analytic team. Public presentations, which are one-way communications, are displayed with dashed lines, with an arrow indicating the direction of the communication. Direct communication involving the transfer of artifacts, including data, formal analyses, instructions, and feedback are displayed with solid lines. These communications in our work were typically two-way communications that involved discussion between both parties, and are thus indicated with bidirectional solid links.

In part because of the tendency towards distinction and isolation, the world of Academia consists of many overlapping research communities, each holding to their own epistemological and theoretical biases as well as methodological practices. Research communities are groups of researchers that share common questions and seek to build an understanding of the answers to those questions together. In order to do that, however, the researchers must come to an agreement on the criteria for consensus building, which includes evaluation of the quality of potential contributions to the shared understanding as well as methods for weighing, balancing, and reconciling apparently conflicting interpretations that come from distinct contributions. At the same time as communities become more internally coherent through this consensus building process, and as they forge their unique identity as a community, they grow in distinction from other existing communities. The multivocal process we advocate requires a concerted effort towards teamwork across communities that runs counter to the forces that drive us apart.

As illustrated in Figure 1, some research communities may overlap more than others, and they may vary by size. It is important to note, as highlighted by the overlapping circles in Figure 1, that while we as researchers may sometimes feel worlds apart from researchers working in very different traditions, we share more of the substance of our work than we might acknowledge. The similarity comes from the shared focus on the target of inquiry, in the case of this book, collaborative learning. It is the distinctions between these overlapping research communities, and the membership of the participants of that process within different ones of these communities, that makes multivocality what it is. That being said, although Figure 1 places each stakeholder participating in the formal process within a single research community, the truth is that we frequently participate in different research communities at different times. Furthermore, Figure 1 places each stakeholder in a different community from the other stakeholders, however, this need not be the case either as long as some stakeholders participate in the process as representatives of different communities than others in the process are representing at that time. Finally, Figure 1 suggests that each participant plays only one role in the process, however, sometimes the Data Provider is also one of the Analysts or the Discussant.

The multivocal analysis process begins with the Data Provider, situated within his or her research community and typically aiming to serve some specific research agenda, who collects a set of data. That Data Provider may have collected the data with the intention of sharing it with other analysts, or may have collected it specifically for his or her own purpose, and then decided later to share it. In either case, the data will typically have had a history before it is shared with the other analysts. The data along with that history is communicated by the Data Provider to each Analyst, and often the Discussant as well. Each Analyst does his or her own analysis. The Analysts may share their analyses with one another informally or formally, or may wait to share them until a formal, public presentation, which was the function of many of the workshops that played a prominent role in our process. This sharing and comparing may lead to iteration in the analyses themselves. Most of this volume focuses on that iterative process, however, we see in Figure 1 that process is only one piece of the

bigger picture. The Discussant is among the audience of the public presentation as are Other Researchers who are consumers of the research. The Discussant plays a role in synthesizing and comparing the analyses, and sometimes challenging the Analysts. The Data Provider is also a consumer of the analyses produced by the Analysts and the meta-analysis produced by the Discussant.

This schema applies equally to the very different processes that took place within each of the five data sections in this volume. Thinking about the lifecycle of an iterative multivocal analysis process, process related questions become relevant once the analysis process is under way, and the chosen answers along the way influence the path the team will take. For example, how much iteration is desirable, and when should analysts decide to stop? What are the inputs to and outputs of each iteration that allow these multiple cycles to progress? When qualitative and quantitative researchers work together, will it logically and practically make more sense for them to work in parallel or in alternating iterations? The analysts on the Fractions and Knowledge Forum dataset spent the most time discussing and comparing analyses because those datasets were chosen early on in the collaboration. In the Group Scribbles dataset, researchers struggled with a common way of referring to the data, and this hindered time spent on discussion. In both the Chemistry and Biology datasets, there were relatively few rounds of formal analysis, but many rounds of reflection and discussion in between. The extensive discussion provided the analysts significantly different lenses through which to view the data. Thus, even with a small number of iterations, in both cases, the understanding that the analysts came away with were significantly altered by the process.

Multivocal Seedlings

As a comparison to our own processes of multivocality, we will explore the experiences of researchers in training as they first experience a similar process. The hope is that this comparison will spark inspiration among both researchers in training as well as instructors of research methods. The Computational Models of Discourse Analysis (CMDA) course was designed to offer primarily first year language technologies graduate students the opportunity to learn to do multivocal corpus research in teams. For many of the students who took the course, this was their first course in any kind of research methods. The situation of this course contrasts with the teams of seasoned researchers that worked together in the five data sections of this book in many respects, which makes this course interesting as a comparison case. For example, whereas the teams that worked together for this book were assigned by the editors to specific data sets for strategic reasons relating to their distribution of analytic expertise and preferences, the CMDA teams emerged through whole class discussions about possible research directions. Because they were new to research methods, they did not come in with strong preferences in terms of their analytic approach, but as part of their orienting process, they were encouraged by the instructor to pursue distinct, complementary paths that would provide useful fodder for challenging one another as they worked together over the 9 week project. In each team, all students played the role of Analyst. In most teams one, or sometimes two, students played the dual role of Analyst and Data Provider. Frequently the data came from their research outside the course. The instructor played the role of Discussant. It is important for those who plan to undertake a similar endeavor that the role of the instructor is both to teach the methods and to help the students learn the important teamwork skills that are the heart of the process.

Potential Pitfalls of the Multivocal Process

In the five Data sections of the book, we sought to emphasize the value of multivocality. Nevertheless, our path has not always been smooth and easy, especially as these team analysis efforts have been our sometimes fledgling attempts at accomplishing something new. In

hindsight the difficulties are not surprising as we consider the extent to which multivocality can be viewed as a counter-culture, as we have hinted above. Throughout this project we were feeling our way as we went. As a result, sometimes we fell into pitfalls, which we see more clearly now in hindsight. In our efforts, we sometimes found pathways for moving forward, which we offer now as helpful hints. Here we outline the main types of pitfalls, and we expand on each along with examples from our own efforts later in the chapter.

The first two pitfalls can be thought of as operating at a more macro level, where the team is constituted, and where the team reaches out beyond itself. The other pitfalls operate at a within-team micro level, as the teams work together as teams. So we begin enumerating the pitfalls at the macro level, and then move on to the micro-level pitfalls.

The first type of pitfall, termed *Team Setup pitfalls*, occurs at the time that the multivocal analysis team is formed to analyze a specific data set. For example, a team may have been selected to represent a specific distribution of analytic approaches. However, just as we have acknowledged that Analysts may have some expertise with multiple analytic approaches, they may choose to approach the data in a different way than was intended by the one who invited them to participate. From a different angle, a lack of understanding of one analytic approach's needs or assumptions might lead to a failure to meet the preconditions of a fully satisfactory application of a method to the data.

The next type of pitfall, termed *Public Presentation pitfalls*, occurs in the public presentations that occur throughout the process wherever there is formal communication indicated in Figure 1 as links. These pitfalls can be characterized as a failure to manage the many different audiences for the public communication that exist within the full set of stakeholders. There may be a failure to respect the trust and vulnerability of the Data Provider, which is experienced as a loss of control on the part of the Data Provider. Lack of sensitivity in these public presentations may engender defensiveness or resentment that works against the intention of the multivocal process. Another Public Presentation pitfall is a failure to communicate research results clearly to those with very different expertise who may struggle to fully grasp some methods that are not familiar to them. A final Public Presentation pitfall occurs when an Analyst targets the presentation to some specific other stakeholder, rather than framing it in a way that is of general interest.

Another type of pitfall, this time a micro-level pitfall, are termed *Data Transfer pitfalls*. These occur from the perspective of the Data Provider in the transfer of data to Analysts. This includes potential failure to set expectations for the work with the Analysts that is mutually acceptable. For example, as will be illustrated below, a Data Provider may make assumptions about what is appropriate data, and these assumptions may not match those of the other analysts. Unspoken expectations and sometimes unconscious assumptions are almost sure to end in disappointment. A failure to adequately communicate important contextual information about the data, how it was collected, and how it was sampled or cleaned up prior to transfer may lead to misunderstandings that can negatively impact the ability of the Analysts to do their work. Or they may invest time and effort into analyses that they later regard as unmeaningful because they were conducted under faulty assumptions about the data. Here "cleaning up" refers to processes of reformatting the data and possibly removing some aspects that are deemed not pertinent for the analysis. Therefore, "cleaning up" only makes sense in certain analytic approaches. In others, it might invalidate the analysis altogether. Once we make the assumptions explicit, then we will know (more explicitly at least) whether the analyses that we want to do are possible or not. Another way these pitfalls may lead to a waste of Analyst time and effort is if Analysts spend time producing an analysis that has already in some ways been done before the data were shared because this prior knowledge and understanding was not communicated to them in the transfer.

Some culpability for miscommunication of context may also reside on the side of the Analysts for reading too much intentionality into choices made by the Data Provider, such as representativeness of the data selected for sharing. We refer to these breakdowns as *Analysis Transfer pitfalls*. Another Analysis Transfer pitfall may occur when an Analyst is selective in which contextual information provided by the Data Provider to take into consideration, such as ignoring heterogeneity within the dataset caused by an experimental manipulation or the inherent hierarchical nature of the data (e.g., students nested within groups, nested within classrooms, nested within teachers, nested within schools).

Within the Analytic team itself, there may also be a failure to engage productively with one another. These breakdowns are considered at length in Chapter 34, Epistemological Encounters in Multivocal Settings, and are thus not the focus of this chapter.

Pathways Around Team Setup Pitfalls

Maintaining a Diversity of Analytic Approaches

An important question to address at the inception of a multivocal research process is the composition of the analytic team. The 5 teams of expert analysts featured in this book were assembled based on interest in a particular data set, diversity in research approach, and in a few cases deliberate positioning of researchers outside of their analytical comfort zones, as we understood them. In every team, we attempted to include a mix of qualitative and quantitative approaches. Within the set of participating quantitative researchers, we included ones that make use of a diversity of different tools in their work, some of which apply statistical methods or machine learning, some of which make use of visualizations and other representations, and still others who use network analytic techniques. It was important to us that the researchers within a team would challenge one another to reflect more deeply on their approach and conclusions. Thus, it was important to pick people that were not only familiar with different techniques, but each have some commitment to something distinct from what others in the team were committed to.

While we selected researchers who were known for their work using particular methods, we did not take any sort of heavy-handed approach to managing their analyses. Thus, the analyses they provided sometimes surprised us in terms of the approaches that were taken. For example, in the Biology team, although we selected the Cress and Kimmerle team (Chapter 27, this volume) because of their sophisticated expertise in the area of quantitative methods, because of the early stage of the research that produced the data set, they did not feel that this approach was appropriate for analysis of this data, and thus took a qualitative approach instead. While this meant that the team as a whole was differently balanced than we as editors had intended, the qualitative analysis they provided was still distinct from the analyses provided by the other team members, and thus still played into the multivocal process illustrated in that data section (Chapters 25-30).

The management of the CMDA teams was looser at the team formation stage but needed to be tighter in the process stage in order to keep the processes moving forward because the students were new to corpus research. CMDA students who shared similar topic interests or research questions gravitated towards one another through whole class discussions that took place in the initial 7 weeks of the 16 week course. Usually one student within each group ended up spearheading the project idea and had some idea of where to get appropriate data. Nevertheless, while the teams themselves chose a topic focus and data, they were all required to orient their analyses to three themes, which were used to structure the 9 week projects into phases that were naturally punctuated by check points in which the teams made public presentations to the whole class and received feedback. These checkpoints served a similar purpose to the workshops that provided the impetus for the long term effort from which this

book emerged. The three themes that the students were required to orient their analyses to included: the self in relation to individual others, the self in relation to the community, and communities in relation to each other. The three themes were meant to serve two purposes. First, within phases of the project, they were meant to provide a common focus as a counterpoint to their separate analytic approaches, similar to the purpose of pivotal moments in our work. Across phases, they were meant to push the teams to see how rich interaction data is, and how it is possible to view the same data from very distinct vantage points. Within each group, the students were each required to contribute their own view of how to analyze their group's data set from the standpoint of the theme using their own chosen methods. They were then supposed to work together to integrate their differing perspectives before presenting their analysis to the class in a public presentation. This is analogous to but different from the notion of pivotal moments, which provided a common thematic focus for many of the multivocal research teams featured in this book.

Altogether there were four student teams. The student teams were meant to be heterogeneous in terms of analytic technique and expertise. Team 1 was the most heterogeneous in terms of expertise and ability level. This team focused on the 2012 Republican debates as their corpus. An emergent research question, which tied together the three themes and gave their project a united focus, was the question of in what ways and to what extent each candidate succumbed to pandering to the public. Their experiences included significant exposure to qualitative research, varying levels of quantitative research in applied linguistics and discourse analysis, and an undergraduate with very little research experience. The undergraduate played mainly a supporting role in the work. The qualitative student took a mainly grounded theory approach. The other two students explored their research question through alternative computational techniques. They spent significant time comparing and contrasting these techniques as operationalizations of pandering, and then evaluating them in light of the qualitative approach.

Team 2 was less heterogeneous in terms of methodological approach, but their work was more grounded in theory, and each of the five team members adopted a distinct theoretical framework that guided their analysis and provided contrasting explanations for learning in their corpus, which they then worked to reconcile. Their research question was what properties of interaction accounted for learning gains in a corpus of peer tutoring interactions. The framing of their question was itself quantitative, and they each gravitated towards pursuing their question using an approach in which they developed a coding scheme, worked to achieve inter-rater reliability with another team member, coded some data, possibly by hand or partly by hand and partly using machine learning techniques, and then performed a quantitative analysis, either in terms of distributions of codes or in terms of sequences of codes, to identify patterns that predicted pre to post test gains. In that sense, their team did not achieve the kind of methodological multivocality that was initially of interest. However, we see that even among researchers that share a set of analytic tools, a productive multivocality along the theoretical dimension can still be achieved. Although the theories they chose as lenses were quite distinct in terms of the independent variables, because they shared a common dependent variable and unit of analysis, they were able to explore whether these perspectives were providing alternative views on the same learning, or accounting for different learning within the same interactions.

Team 3 was unique in that they shared a common theoretical perspective, namely Good Death theory (Steinhauser, Clipp, McNeilly, Christakis, McInture and Tulsky, 2000), but explored two distinct datasets, one of which was an online cancer support forum and the other of which was a corpus of suicide notes. The questions they pursued were how concepts from Good Death theory are reflected through language behavior in the two corpora, and what is in

common and distinct in the experience of death between cancer victims and suicide victims. Three of the members of the four student team were highly quantitative in their approach, and made heavy use of machine learning in their work. The fourth student adopted an approach that was very similar to that taken by the students in Team 2 but was preceded by significant time reading whole posting histories of long time participants in the cancer support forum and gaining a qualitative sense for what was happening in the data before adopting a quantitative approach. Team 3 produced a very interesting analysis of the experience of death in an online discussion forum, which was submitted for publication soon after the end of the semester. Eventually it came to light, however, that in order to achieve this positive result, the one student in the group with the most insight into the data really drove the whole process, which allowed the other analysts to take less of a leadership role. This highlights the importance of continually monitoring group processes for breakdowns in teamwork in a multivocal process.

The fourth team was extremely homogenous in terms of approach, all taking a very quantitative approach. This team of three students came in with the least expertise of any of the teams in terms of understanding and facility with theory driven research. In their project, which focused on a Supreme Court Hearing corpus, they took a strongly atheoretical approach to their analysis. Their engagement remained at a very superficial level when it came to integration of results. In order to address this issue, in every meeting, the instructor continued to pose challenges to them regarding the interpretation of their results in order to spark more intensive engagement between their perspectives.

Like the experiences of the five expert teams from this book, there were varying levels of success at achieving the kind of exploration and interaction that was intended within teams. At each stage the instructor played an integral role at scaffolding the teamwork. At the team setup stage, the instructor pressed each student to take responsibility for playing a distinctive role on the team, and yet the instructor continually challenged the members of each team to think about how the view they were seeing through their analytic lens could speak to the other students in the group. Part of this scaffolding was the reward structure for the class, where each student received a grade that was based in a group grade for the project, but which was adjusted based on individual contribution. However, as we see with Team 3, this was not always successful. We see that there are many productive paths towards multivocality that do not require diversity in all of the dimensions. However, in the case where readers plan to similarly use multivocality as a paradigm for research methods instruction, it is important to note the ways in which they may need to intervene and scaffold the process at every stage in order for it to remain productive.

Satisfying the Preconditions of a Diversity of Analytic Methods

In our experience working on the analyses that led to this book, two tricky questions come up related to how to prepare data for a multivocal research process. First, what is an effective process for preparing data for sharing, overcoming challenges in data sharing, and specifically challenges in communication about data for secondary analysis? Second, how can we deal effectively with the fact that different analysis methods have different data needs (i.e., quantitative approaches require larger data sets, qualitative approaches require more intimate knowledge of the context from an insider's perspective whereas quantitative approaches seek objectivity)? While setting up the data may be viewed as a mundane aspect of the analytic process, it is this early stage where seeds of pitfalls are sewn. The problem starts with the tendency towards isolation referred to above. Data analysis is not frequently undertaken as a team sport. Instead, researchers are more likely to retreat into their cave with their favorite analysis tools, doing the initial exploration of their data in tandem with data preparation. It is during the exploration phase that more plans for more extensive analysis form (apart from planned contrasts directly related to the experimental manipulation if any). However, this

time of individual exploration and data manipulation might lead to work being done to set up the data that may need to be undone or redone differently later in order for the analysts to work together as a team later in the process. Important questions must be answered about how to meet the preconditions of the diversity of analytic approaches that will be used across Analysts. The later these questions are addressed within the team, the more likely it is that work done up front will have ended up including time and effort wasted on work that will need to be redone later or that was redundantly performed by more than one researcher.

Choosing and preparing data

The team that worked on the Chen and Looi Group Scribbles dataset provides the impetus for the first of these questions, which will give us the opportunity here to reflect on where assumptions about data gathering and preparing can come from. These reflections will give us the opportunity to observe how conceptual debates within the field can trigger forces towards isolation within teams that need to be overcome in order to achieve multivocaltiy. We will frame this discussion in terms of contrasts between 'naturally occurring' or 'authentic' data vs. 'contrived' or 'researcher provoked' data. In the case of Group Scribbles, the data provider made assumptions about how the data should be prepared that did not match with how the analysts wanted the data. We will see that the data provider was situated in the 'naturally occurring' data gathering paradigm, and he had additional assumptions about preparing his data for analysis, that were not shared by the other analysts.

Let us consider the assumptions behind the debate concerning 'naturally occurring' or 'authentic' data vs. 'contrived' or 'researcher provoked' data. The former stem from human interactions that would have occurred even in the absence of the data-collection activity or in other words, they pass the 'dead social scientist test'. "If the researcher got run over on the way to the university that morning (Potter, 2004: p. 612)", would the interaction have nevertheless occurred and played out in the same way? The latter type of data are relatively contrived social science data sources such as surveys, interviews, and focus groups (Speer, 2002) and although there are many differences to be discussed between this and experimental data, for the purposes of this chapter we add to this latter type, data from experiments where participants perform tasks under controlled settings. The comparison of these two data types is meant to illustrate that a researcher gathers data in order to perform specific analyses (this means that other analyses could be difficult or impossible to do) and that the way in which the data is gathered and made ready is compatible with the researcher's assumptions (this means that the data gathered could be incompatible with other researchers' assumptions).

So, how can this distinction between 'naturally occurring' and 'artificially provoked' data help us illustrate our argument that data is collected and prepared according to underlying assumptions? Speer (op. cit.) gives an insightful overview of the views of conversation analysts that stem from its ethnomethodological origins on why experimental data is problematic. She goes on to argue that just because ethnomethodologists attribute their own assumptions and higher order goals to experimental data in order to explain why it is problematic does not mean that such data is problematic for researchers in an experimental paradigm who have their own assumptions. First of all, the ultimate objective for ethnomethodologists — as developed by Harvey Sacks and colleagues — was to produce an inventory of "recognizable social actions in this culture… the aim is to find it and provide an account of it empirically and precisely, not imaginatively or typically or hypothetically or conjecturally or experimentally, and to use actual, situated occurrences of it in naturally occurring social settings to control its description" (Schegloff, 1996a: p. 167). Given this, the assumption is that if researchers use "written texts, monologues, talk or writing produced under experimental or quasi-experimental conditions" (Schegloff, 1996b, p. 468) then since "these materials are not drawn from the naturally occurring interactional environments which

seem to be the natural, primordial home for language use", the inventory of social actions will be compromised because the "primary and proximate interactional practices which undergird" the specific social action we are studying "may be largely or totally absent, often suppressed by specially designed circumstances of production" (Schegloff, 1996b, p. 468). So this is an argument that ethnomethodologists make against experimental data as analyzed in typical quantitative research methodologies, but according to their own assumptions.

If our goal as researchers is providing an inventory or a catalog of recognizable social actions as they occur naturally, it does not make sense to use imaginary, made-up examples of language use that are purported to be typical, based on the intuition that there is reason to doubt such conjecture. And it is certain that experimental conditions do *not* embody ordinary contingencies of interaction; instead, they "confront participants with quite distinctive, and potentially complicating, interactional exigencies" (Schegloff, 1999: p. 419). But researchers who are not ethnomethodologists may not be seeking to produce an inventory of recognizable social actions as they occur naturally, during ordinary conversation. Their goal may be to flesh out how experimental conditions do indeed affect language use within group interactions and to make probabilistic assertions about that, a goal foreign to conversation analysts (Golato, 2003). They may hypothesize that experimental conditions could *provoke new* language use, not usually present in ordinary interaction, but that may be beneficial for learning, for example. In that case, their data matches their assumptions and goals. Experimental data therefore escapes the criticism of not being naturally occurring, as it was never argued as being so and since experimental methods are used for different goals than conversational analysis, they can co-exist, as long as experimental researchers do not treat their interventions as "neutral resources for accessing some truth or reality beyond or beneath the data" itself (Speer, op. cit.).

This discussion about how researchers in different paradigms have different assumptions about gathering and preparing data should have made clear that once it has been assured that gathered data is both appropriate for specific analyses to be carried out and compatible with a particular set of researchers' higher objectives and assumptions, difficulties will most likely be encountered when data is shared with researchers not party to the gathering. Although other researchers may share interest for a particular type of setting (e.g. group interaction), they can hold different assumptions about it, have different goals, focus on different aspects of it and as a result want to analyze data of a specific nature different from that provided by the data gatherers.

Though the issues the team examining the Looi & Chen (Chapter 14, this volume) data experienced grew out of seemingly unsolvable debates in the field, eventually the team worked out a productive solution. In particular, the Data Providers initially furnished a synthesis of their vision of the pedagogical interaction (concerning electricity) and not a transcription of the actual interaction. If this was suitable for their own goals and assumptions, it was not suitable for example, for Lund & Robinault (Chapter 17), who requested transcriptions. It was important for their research questions to have 1) complete transcriptions of talk, 2) talk as uttered and not modified (e.g. summarized) by the person doing the transcribing, 3) correct differentiation of turn taking – both in terms of content expressed and in terms of interactional chronology in relation to other turns and finally 4) correct differentiation of speaker. In addition, since they analyzed the interaction from the view of physics didactics, it was also important for them to have knowledge of the sequence of learning activities of which the one classroom action was part: what kind of knowledge about electricity did the students have coming into the classroom? What were the specific learning goals for the classroom activity that was recorded? What kind of course was to follow? This information was not initially provided, perhaps because it was not all relevant to the data

providers' own analysis. However, the needed information was eventually offered to the analysts who needed it when the problem came to light. This required some unanticipated effort on the part of the Data Providers, however.

The interrelations of methods and data

The team that worked on the Fractions dataset provides the impetus for the question regarding the interrelations and constraints of applying any given method to a particular dataset, but also the method's relation to theory. The Fractions team was nicely poised for multivocality in terms of distribution of analytic expertise. In particular, Ming Ming Chiu (Chapter 7) was selected as an analyst because of the sophisticated statistical discourse analysis (SDA) technique that he contributed, which is striking in its contrast both to qualitative approaches as well as other, simpler quantitative approaches featured within the book that are not able to capture the sequential nature of collaborative discourse. However, a sophisticated statistical technique requires a large amount of data in order to be used appropriately and the fractions dataset was quite small; it was a transcript of one group discussion. While the typical tests for over-fitting using this approach did not show unequivocal signs of over-fitting, one must still use extreme caution when drawing conclusions from such a complex model applied to such a small amount of data.

Each type of method has its own constraints for application to data. Issues related to over-fitting are specific to quantitative approaches. A small amount of data can still provide the basis for insightful thick description using a qualitative technique. In fact, qualitative researchers may find the opposite challenge on multivocal teams. While there are well-established methods for sampling from a larger corpus to identify segments that can be approached from a qualitative standpoint, it still remains to be worked out how to integrate analyses across approaches when the quantitative analysis is applied to the whole corpus, whereas the qualitative analysis is applied only to a small portion. A qualitative analysis may be valuable even if it is not meant to illustrate a pattern that is claimed to be generalizable to the whole corpus. In fact, its value may be precisely because the scenario that is being highlighted is unusual. Therefore, the goals of the analysis from a qualitative standpoint and quantitative standpoint may be distinct, and thus the findings may require some creativity in order to integrate in a valid way. This exchange raises an important caveat that applies at the Analysis Transfer stage discussed below. Specifically, it is extremely important within multivocal teams at the time when analyses are shared across the Analysts that they take care to consider the limitations of what can be concluded from one another's analyses depending on the extent to which the preconditions for a felicitous application of those methodologies were met in the data. We observed difficulties in this regard, especially in connection with Ming Ming Chui's analysis in both Chapters 7 and 23, because the other Analysts in the community found his approach to be beyond their level of technical expertise and somewhat mystifying, and thus they found themselves less capable of evaluating or challenging this work. A potential pathway towards addressing issues like this would be for Analysts within teams to offer short tutorials to one another to build common ground prior to exchanging analyses.

How student teams dealt with preparing data and choosing methods

As mentioned earlier, there were four student teams in the CMDA course. The first question on choosing and preparing data was universally an issue for the student teams. While challenges with respect to data sharing and comparison of analyses came up for the student teams, none were insurmountable. In examining their experiences, we can learn to anticipate such issues and prepare for them in such a way that they can be dealt with efficiently. All of the data sets that students made use of required a substantial amount of time and effort to set

up for the analysis. Thus, one important lesson we can learn is to make sure when constituting analytic teams that each includes someone with expertise in data cleansing and manipulation, and that time for that preparation is taken into account when the project is planned. Furthermore, that person should take responsibility to prepare the data set for sharing before the other analysts begin their work. Such processes should not be entered into glibly, however, since as we discussed above, "cleaning up" the data is only a standard practice in some methodologies, and is inconsistent with others. So some serious discussion must be conducted with the team of analysts in order to agree on what makes sense in order to prepare for their joint endeavor that respects each represented methodology, or at least represents a compromise all of the Analysts agree to up front. Making sure this occurs effectively is the job of the instructor.

The second question on interrelating methods and data was most relevant for Team 1, where the strongest commitment to a grounded theory approach was found in one team member. The grounded theory method yielded two complementary sets of themes, one of which was related to the topics addressed by each candidate, and the other of which was related to the argumentation strategy that candidates adopted. Using these two sets of themes, when applied to debate transcripts, allowed the team to explore how candidates differed in terms of their associated distribution of strategies, but also how some kinds of strategies were more associated with some topics than others. As a comparison between the qualitative and quantitative approaches, they were able to determine whether the kinds of group level differences they saw with the automatically derived topics were similar to the ones they saw using the topics identified using a grounded theory approach. The challenge in that collaboration was that it did not become clear what the most valuable contrast between the qualitative approach and the quantitative approach would have been for constructing one integrated understanding of the dataset in the end of the project until relatively late in the process. In the case of this team, alternating between qualitative and quantitative analyses might have been a more strategic approach since the quantitative analysis cycles were quicker and were therefore able to encompass more data within shorter amounts of time.

Pathways Around Public Presentation Pitfalls

The biology data

The analysis of the Biology data is a good illustrative example of Public Presentation pitfalls. The analysis of this data was necessarily iterative since not all of the data included in the final analysis was available when the collaboration across groups began. The initial analysis also sparked a fair bit of controversy, including a question about whether multivocal analysis is even appropriate for data collected within an environment at an early stage of development.

A number of issues came up in the initial public presentation of the Biology dataset at the Alpine Rendez-Vous workshop at Garmisch-Partenkirchen in 2009. Some of these were symptomatic not only of Public Presentation pitfalls, but also of Data Transfer pitfalls, which will be the focus of the next section. The current Data chapter in the Biology section (Dyke, Howley, Kumar, & Rosé, Chapter 25, this volume) includes an explicit write up of the constraints the Data Providers were working within when collecting the data, which were not communicated to the Analysts adequately before they began their work. This information was included in the write up of that chapter as a response to this public discussion. Much of the discussion at this initial public presentation focused on what the researchers should have done differently, some of which were things the researchers did not have any control over given the context of their work, and some of which were issues they were aware of but were not the focus of their investigation. The time of this public presentation was not the appropriate time for these constraints to come to light for the Analysts. Time spent on

discussion of issues the Data providers were already aware of and issues that were beyond their control took time away from what could have been a more productive intellectual exchange.

Another issue came up in Stahl's presentation and the subsequent write up of the Stahl analysis (Stahl, Chapter 28, this volume). Here the issue was that the chapter was written in the frame of communication from Stahl to the Data providers specifically, and focused largely on advice that would be useful to the Data providers in their process but might not provide value to other researchers not specifically involved in the design process because of its level of specific focus on the prototype intervention used to collect the data. Since Stahl and the Data providers were close colleagues for whom a frank exchange of views was the norm, and in fact was quite welcome in private communication, the main issue to be considered was whether the presentation was appropriate for public consumption. That analysis focused largely on lists of things Stahl would have done differently if he were setting up the experiment. The question here is whether the value in multivocality is in license to publically present unmitigated criticism, or whether there might be some more productive exchange that can take place in public settings when the object of analysis is data from a pilot experiment. This analysis contrasts with the Cress & Kimmerle analysis (Chapter 27), where a similar focus on what could be improved in the design was offered, but it was contextualized in a theoretical framework in a way that spoke to a broader research community.

As we see, the initial analyses were critical of the collaborative environment and the design of the study that produced the data in a way that the Data Providers found tangential to the questions the study was meant to address, and the way some of this was communicated in public presentations resulted in some angst. The analysis chapters in the Biology section preserve the issues with respect to Public Presentation pitfalls discussed here in order to provide visibility to researchers interested in embarking upon a similar journey in their own work. Nevertheless, it should be acknowledged that the initial divergence of perspectives did eventually lead to a series of productive interactions between analysts that challenged the Data Providers in their conception of the research as well as challenging the analysts in their conception of multivocality. In the context of the group discussion at the Garmisch-Partenkirchen workshop, the Data Providers came to understand the underlying conceptual differences in assumptions about the ideal role of the facilitator in discussion groups that the different Analysts highlighted. This distinction eventually became the lens through which the diverse group of Analysts was able to debate and build consensus. This teaches us that while these pitfalls are prone to occur, and though they may cause some temporary friction within an analytic team, they need not cause irreconcilable difficulties in collaboration. Perhaps a take away is that Data Providers should come in to the process with an expectation that some thickness of skin and perseverance will be necessary.

After rounds of discussion and reflection, an additional analyst (Goggins & Dyke, Chapter 29, this volume) was added to contribute an extensive network analysis, to bridge the coding and counting approach in the original Howley et al. analysis (Chapter 26, this volume) and the qualitative Stahl analysis. After this interaction, a second data collection effort in an updated version of the environment provided a complementary set of data, and then eventually a third. After much reflection and discussion, the team converged in their understanding of multivocality and its role in iterative, design based research, and the multivocal process resulted in a number of observations that led to a successful redesign of the intervention. The successful intervention then produced new knowledge for the field about how conversational agents can be used to support group discussion. And the resulting agent design represented insights drawn from diverse perspectives on appropriate support for

group discussion in the field from those who may not have had the opportunity to work together apart from a desire to engage in a multivocal process.

The student teams

The student teams struggled with different issues in their public presentations than the expert teams. In particular, because of their relatively early stage of familiarity with the methods they were using in their analyses, they found it challenging to clearly articulate their analyses and findings and to compare and contrast with one another in preparation for these public presentations. Thus, although the teams presented together, and although they were instructed to present not only their own analyses but also lessons learned from interaction within their teams, many presentations came across as a patchwork. This left more work for the instructor who was acting as a Discussant to engage the teams, sometimes privately in advisory sessions with the groups prior to public presentations on some iterations, but also in the context of whole class discussion. The purpose of this scaffolding was to clarify what each analysis demonstrated and how the alternative analyses may challenge one another. These facilitated discussions served to scaffold the communication and coordination between analysts that often occurred outside of public view in the more experienced teams featured in this book. Thus it is important to note that when incorporating multivocality in an instructional setting, these public presentations may serve as valuable opportunities to learn the multivocal process itself rather than simply opportunities to communicate with a broader audience.

Pathways Around Data Transfer Pitfalls

The chemistry data

The analysis of the Chemistry data provides a convenient example of Data Transfer pitdalls. The analysis of this data proceeded in three phases. In a first phase, an initial version of all of the analyses was completed by individual researchers. All but the Sawyer analyses were presented at a workshop in 2010 at the International Conference of the Learning Sciences (ICLS). During the discussion at the workshop, the leadership theme emerged and then became a consistent thread in all subsequent analytic work by the team. Eventually, both the Rosé and Strijbos teams revised their characterization of their respective multi-dimensional coding frameworks. The workshop sparked a collaboration between the Rosé team and the Strijbos team, which proceeded in terms of informal discussions over more than a year, and then finally a formal reanalysis in time to write a chapter about their integration (Howley, Mayfield, Rosé, & Strijbos, Chapter 11, this volume). Discussions with the other analysts proceeded in parallel. Elaboration of both of those other analyses ensued, until finally the discussant used the emergent leadership theme to contrast the findings across the four analyses.

The Data Transfer pitfall that came up in this process occurred immediately subsequent to the 2010 workshop. The culpability here in the transfer may have been on the side of the Analysts some of whom may have read more in to the data sampling process than was warranted. The Data provider had chosen two discussion groups whose style provided an interesting contrast from the standpoint of the theoretical framework that motivated the data collection in the first place. The concept of leadership was not central to the contrast that the Data Providers were necessarily interested in. And the Data Providers never asserted that the particular problem solving episode that was selected for examination in order to compare the two groups was necessarily representative of every aspect of the collaboration within those groups. As the discussion comparing analyses across Analysts turned towards leadership, it was tempting to draw inferences about relationships between the students within the groups from the small amount of data that was provided. Interest in pursuing the issue further led to

a request for the Data Provider to offer more transcripts from the same groups solving other related problems. Only one Analyst actually examined this larger corpus in detail. That extended analysis revealed that the encounters examined jointly by the multivocal team were not in fact representative of consistent leadership taking within the groups. This additional analysis was eventually dropped since it was conducted only by one analyst. In hindsight it was not surprising that while the selected examples served the original intention of the Data providers in their own analysis, it was not necessarily ideal from the standpoint of analyses that focused on different issues. This simply raises a note of caution to take into consideration during the process of transferring data for multivocal analyses. If the analysts had asked more questions up front, they might have focused their questioning of the data in a direction that was more consistent with the considerations used in selecting the sample, or they might have negotiated for different sampling criteria in the data sharing process.

The student teams

With the student teams, issues with respect to data transfer came up primarily for teams 1 and 2. In the case of team 1, the Republican Debates, not all of the students within the team shared the same amount of expertise regarding the American political process, and thus some insights about how the data might be productively questioned were not shared. For example, some students who did not grow up in the United States were far less aware of the important role of the region in which a debate took place would play in terms of what could be assumed about the audience the debaters were presenting themselves to. Eventually more knowledgeable students within the groups shared their expertise with the less knowledgeable students as part of the process of working towards integration of findings. However, in hindsight, scheduling in time to have these discussions at the very beginning of the process might have allowed the team to use their time more efficiently. A more difficult issue came up for Team 2, who was using a data set that none of them had participated in collecting. In the case of the expert teams, there was a plan for the data to be shared with other analysts, and written documentation to facilitate the sharing was provided in addition to the public presentation of the data that occurred at the workshops at a key stage. In the case of the student teams, they received the data from a researcher who was remote from the process and did not anticipate what the student team would need to know in order to do its work. Thus, the team engaged in a lot of guess work about the data up front. The time lost in the process due to the guess work highlights the importance of taking the data transfer stage seriously as a critical part of the multivocal process and thus actively engaging the Data Provider.

Pathways Around Analysis Transfer Pitfalls

The fractions data

Engagement between researchers is the underlying concern of Analysis Transfer pitfalls. The Fractions team was exemplary in terms of the level of engagement of the analysts in the multivocal process. The sharing and analysis of the Fractions data began previous to the 2009 Alpine Rendez Vous in Garmisch-Partenkirchen, where three initial analyses were presented by individual researchers: Shirouzu, Chiu and Trausan-Matu. These analyses continued to evolve, already influenced by each other and re-discussed informally at a workshop in Chicago, 2010 at ICLS and more formally in preparation for a symposium in Hong, Kong, 2011 at the Conference for Computer Supported Learning (CSCL) conference. At the Alpine Rendez Vous 2011 in La Clusaz, France, analyses evolved still further, with Shirouzu, the data provider taking a kind of integrative stance, recognizing that different units of analysis and frameworks could be complementary, and both Chiu and Trausan-Matu making changes to their views as well. Discussions led to so many changes in analyses that Lund, the discussant for the Fractions section had to rewrite the chapter (Lund & Bécu-Robinault,

Chapter 17, this volume) many times in order to account for them. Nevertheless, what we see here is evidence that continued engagement between analysts can lead to progressive refinement of constructs, ideas, and conclusions over an extended period of time.

The Chiu analysis (Chapter 7, this volume) in particular highlights other relevant challenges in sharing results across very different analytic approaches. Although Chiu was indeed responsible for assuring that his method was applicable to the fractions data, certain contributions of his analysis within the context of the fractions analysis team lay on a different level than his results, per se. This can lead to challenges in communication across Analysts. For example, the pivotal moments that Chiu found prompted Shirouzu (Chapter 5) to give meaning to them in his own framework, illustrating that the results of one method can be reinterpreted within an alternative theoretical framework. Chiu claims his SDA method to be in a sense *atheoretical,* able to be used with a variety of theoretical frameworks (although this was questioned by Fujita, Chapter 24). In the fractions section, he looked for micro-creativity, but he could search for patterns of any type. These issues are not insurmountable however, they must just be carefully considered and explicitly discussed among Analysts.

The Group Scribbles data

The Group Scribbles experience stands in contrast especially with that of the Fractions dataset where there was a notable intensive exchange between analysts over multiple iterations. With the Group Scribbles team, the analysts required some prompting, sometimes by discussants, to engage deeply with the distinctions between their analyses. Data providers Chen & Looi (Chapter 14) shared the Group Scribbles data at the end of 2010 and initial analyses on the data concerning electric circuits were presented in March of 2011 at the Alpine Rendez Vous. Contrary to the other datasets, the data providers did not present their data in person, but rather sent the group documents describing it. Suthers presented the dataset at the ARV2011, and he became the discussant for this section (Chatper 19, this volume). Jeong (Chapter 18), Lund & Becu-Robinault (Chapter 17), and Medina (Chapter 16) all contributed analyses, in addition to an analysis offered by colleagues (Wee, Song & Looi) of the data providers. The team had first drafts of all analysis chapters at the end of June 2011, after some analysts obtained partial transcriptions that they put together from the videos. Some of the group members met Lund at CSCL2011 in Hong Kong in early July to discuss how our conclusions compared, but this was difficult as the analysts were analyzing different empirical material and so did not have a simple way of referring to places in the dataset that would be easy to map from one analysis to another. Until January 2012, discussion continued over e-mail through part of August 2012 when Suthers posted the discussion chapter and Jeong and Lund commented on it. But, it was only after this that Lund & Becu-Robinault succeeded in aligning the transcript they had greatly modified (in order to respect their epistemological constraints) and their pivotal moments with the data providers' original synthesis of the interaction. Unsurprisingly, the lack of a common reference to the data greatly hindered the exchange between analysts.

The Knowledge Forum data

Like the Fractions dataset, work on the Knowledge Forum data began at the 2009 Alpine Rendez Vous in Garmisch-Partenkirchen, where two separate analyses of the data were presented, namely Teplovs and Fujita as one, and Tscholl and Dowell as the other. The first round of analyses revealed some challenges with the multivocal process. In particular, the Tscholl and Dowell analysis was eventually discarded because the sample selected for up close analysis was felt to be non-representative by the data provider, and thus the conclusions drawn from the analysis did not have face validity from the data provider's perspective. The original Teplovs and Fujita analysis focused more heavily on the methodology and less

heavily on the data than the current Teplovs and Fujita chapter (Chapter 21). This prompted a discussion by Rosé that similarly focused on methods rather than substantive conclusions that might have fed back into the design based research process that produced the data. We see here that a multivocal process can get off to a slow start.

Ultimately, the outcome of this first round was disappointment to the data provider that the analyses and discussion did not necessarily further the research goals of the project nor fully appreciate the complexities of the data. In response, the Teplovs and Fujita chapter that is included in the book addresses the research goals of the project more explicitly and clearly than the initial analysis presented in 2009, and an additional analysis was invited, this time coming from within the Knowledge Building community itself by researchers who were able to fully appreciate the larger goals of the project, namely Law and Wong (Chapter 22). A further analysis was invited by Chiu (Chapter 23), which provides a purely quantitative sequential analysis of the data, in the same spirit as the Chiu analysis provided for the Fractions dataset. This second round provided a richer multivocal experience. Nevertheless, convergence is difficult. In the end, as seen in the Discussant chapter written by the data provider herself (Chapter 24), the value in the multivocal process was attributed to the impact it had on her view of the data, but not in its contribution to the design goals that prompted the data collection. In this case, one might conclude that further iteration would be required in order to provide that needed convergence in order to inform design.

The student teams

As mentioned, the CMDA students were required to orient their analysis to three related themes, each associated with a major presentation the groups were required to give to the whole class, which provided answers to these questions for them to follow. At each check point, the instructor acted as the discussant, giving the teams suggestions for how they might push their individual analyses further as well as explore comparisons and contrasts between analyses. Each check point presentation also involved time for group discussion.

Iteration was really essential for the student teams in order to spur them to do thorough analyses, since they were just learning what it meant to do a rigorous analysis. This was, of course, not an issue with the experienced researchers we were privileged to partner with in creating the 5 data sections featured in this book. When the student teams reached their first check point, none of the groups had done the integration. They spent until the wee hours of the morning of the presentation doing their own analyses, so each team presented a patchwork that was not integrated. The instructor sent them back with the feedback that at the next check point they should strive for better integration. When the teams presented their second theme analyses, the results were much stronger in every group. All but one group had several interesting stories to tell. But they still did not leave time before the presentation to think about the integration. Again, they were up until the wee hours the morning of the presentation finishing up their own analyses. The instructor then decided that since most of the groups had substantial raw material for a multivocal analysis by then, they were allowed to abandon the three themes in order to place their emphasis on the integration question for the third iteration. By the final presentation, most of the teams had come to a point of seeing value in integration and had spent substantial time working to reconcile the alternative perspectives offered by the distinct approaches their teams had pursued.

In hindsight, one might argue that trying to teach multivocality to students simultaneously as they are learning their own analytic methods is just too high of cognitive load and that this should be regarded as an advanced method, not to be entered into until one has some facility with at least one research method. An alternative option may have been to take a jigsaw approach, however, this would have been difficult in the CMDM course due to resource

constraints from the side of instructional time and teacher resources on the one hand and breadth of learning objectives on the other. Nevertheless, the CMDM course indeed struggled along the way with concepts related to conducting theory driven research at all, and found it difficult to manage their time in order to balance doing their own analysis with integrating with those of their team mates, many students commented after the fact about having benefitted from working with their team and being challenged by their team. Thus, it would certainly be possible to argue that despite the cognitive load demands, there was benefit from the additional struggle because of the broader perspective it provided. Furthermore, in working in a group on the analyses, students were able to be more ambitious in their goals for the analysis.

Lessons Learned: Reflections on Methods for Multivocality

In this chapter we have worked to abstract away from the specific processes that the five teams in the book engaged in that were illustrated within the five data sections, and have explored some practical questions in light of lessons learned from these processes. As a comparison case, we have contrasted the expert teams that worked with us on this book with a set of four student teams just learning how to do theory driven research and engage in a multivocal process in tandem.

In addition to the rules of thumb and practical suggestions that have been offered in this chapter, we can draw some conclusions in reflection. What we see is that in both the expert teams as well as the student teams, there were ways in which the multivocality proceeded successfully as planned, ways in which it did not work out so well, and ways in which it worked out differently than planned, but successfully nonetheless. It is the last of these three points that is potentially the most important, because we see that it is possible to benefit from a multivocal process even when it is not perfectly planned out to begin with, and even if it doesn't play out exactly as anticipated. Thus, one should not shy away of multivocality for fear of making mistakes. Even the student teams who were just fledgling researchers, for the most part, benefitted from being challenged to look at their data from multiple perspectives. Many of them were surprised in the end that they found out how brittle an analysis conducted from only one perspective might be and how subtle differences in operationalization even of constructs that seem to be identical when conceived from a conceptual standpoint dramatically change the claims one feels comfortable making as a result of the analysis.

Perhaps the most valuable lesson learned in all of this is the contrast between the real benefit of a multivocal analysis and more standard mixed methods approaches. Whereas there is increasing consensus about the benefits of a mixed methods approach for research findings, strength of the conclusions, and depth of insight into the target phenomena, there is still something missing from mixed methods research that is gained through multivocal methods. Whereas mixed methods approaches benefit the research, multivocality benefits the research community, forging new connections in terms of relationships and publications between researchers and their respective communities that did not exist before. This sentiment is echoed in the words of the discussant of the Knowledge Forum section, who is a researcher well experienced in mixed methods approaches prior to participation in the multivocal research process that produced this volume (Fujita, Chapter 24, this volume). In mixed methods there is one agent, so the methods are not likely to challenge each other deeply. In multivocal analysis, there is a different agent representing each method, so the dialogue can be more genuine, "multivocal", and there can be more substantial challenges.

References

Bakhtin, M. (1981). Discourse in the novel. In Holquist, M. (Ed.). *The Dialogic Imagination: Four Essays by M. M. Bakhtin*, University of Texas Press: Austin.

Blei, D. M., Ng, A. Y. and Jordan, M. I. (2003). Latent dirichlet allocation. Journal of Machine Learning Research, volume 3, pp 993-1022.Golato, A. (2003). Studying Compliment Responses: A Comparison of DCTs and Recordings of Naturally Occurring Talk. *Applied Linguistics* 24/1: 90-121, Oxford University Press.

Chiu, M. M. (this volume). Social metacognition, micro-creativity and justifications: Statistical discourse analysis of a mathematics classroom conversation. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions*, Chapter 7. New York: Springer.

Chiu, M. M. (this volume). Statistical Discourse Analysis of an Online Discussion: Cognition and Social Metacognition. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions*, Chapter 23. New York: Springer.

Cress, U., & Kimmerle, J. (this volume). Successful knowledge building needs group awareness: Interaction analysis of a 9th grade CSCL biology lesson. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions,* Chapter 27. New York: Springer.

Dyke, G., Howley, I. K., Kumar, R., & Rosé, C. P. (this volume). Towards academically productive talk supported by conversational agents. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions,* Chapter 25. New York: Springer.

Fujita, N. (this volume). Critical reflections on multivocal analysis and implications for design-based research. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs & N. Law (Eds.), Productive *Multivocality in the Analysis of Group Interactions,* Chapter 24. New York: Springer.

Goggins, S. P., & Dyke, G. (this volume). Network analytic techniques for online chat. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions,* Chapter 29. New York: Springer.

Howley, I. K., Kumar, R., Mayfield, E., Dyke, G., & Rosé, C. P. (this volume). Gaining insights from sociolinguistic style analysis for redesign of conversational agent based support for collaborative learning. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions,* Chapter 26. New York: Springer.

Howley, I. K., Mayfield, E., Rosé, C. P., & Strijbos, J.-W. (this volume). A multivocal process analysis of social positioning in study groups. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions,* Chapter 11. New York: Springer.

Jeong, H. (this volume). Development of group understanding via the construction of physical and technological artifacts. In D. D. Suthers, K. Lund, C. P. Rose, C. Teplovs & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions*, Chapter 18. New York: Springer.

Law, N., & Laferriere, T. (this volume). Multivocality in interaction analysis: Implications for practice. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs & N. Law (Eds.), Productive Multivocality in the Analysis of Group Interactions, Chapter 35. New York: Springer.

Law, N., & Wong, O.-W. (this volume). Exploring pivotal moments in students' knowledge building progress using participation and discourse marker indicators as heuristic guides. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs & N. Law (Eds.), Productive Multivocality in the Analysis of Group Interactions, Chapter 22. New York: Springer.

Chen, W. & Looi, C.-K. (this volume). Group Scribbles-supported collaborative learning in a primary grade 5 science class. In D. D. Suthers, K. Lund, C. P. Rose, C. Teplovs & N. Law (Eds.), Productive Multivocality in the Analysis of Group Interactions, Chapter 14. New York: Springer.

Lund, K., & Bécu-Robinault, K. (this volume). Conceptual change and sustainable coherency of concepts across modes of interaction. In D. D. Suthers, K. Lund, C. P. Rose, C. Teplovs & N. Law (Eds.), Productive Multivocality in the Analysis of Group Interactions, Chapter 17. New York: Springer.

Lund, K., Rosé, C. P., Suthers, D. D., & Baker, M. (this volume). Epistemological encounters in multivocal settings. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs & N. Law (Eds.), Productive Multivocality in the Analysis of Group Interactions, Chapter 34. New York: Springer.

Martin, B.(1998). The Politics of Research, in *Information Liberation,* London: Freedom Press.

Medina, R. (this volume). Cascading inscriptions and practices: Diagramming and experimentation in the Group Scribbles classroom. In D. D. Suthers, K. Lund, C. P. Rose, C. Teplovs & N. Law (Eds.), Productive Multivocality in the Analysis of Group Interactions, Chapter 16. New York: Springer.

Potter, J. (2004). Discourse Analysis. In M. Hardy & A. Bryman (Eds.), Handbook of Data Analysis (pp. 607-24). London; Sage.

Schegloff, (1996a). Confirming Allusions: Toward an Empirical Account of Action. *American Journal of Sociology*, Vol. 102, No. 1 (Jul., 1996), pp. 161-216

Schegloff, (1996b). Some practices for referring to persons in talk-in-interaction: a partial sketch of a systematics. In (Ed.) B. A. Fox, *Studies in anaphora*. Amsterdam: Benjamins publishing company. Pp. (437–485)

Shirouzu, H. (this volume). Focus-based constructive interaction. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs & N. Law (Eds.), Productive Multivocality in the Analysis of Group Interactions, Chapter 5. New York: Springer.

Speer, S. (2002). "Natural" and "contrived" data: A sustainable distinction, Discourse Studies, 4, 511–525.

Stahl, G. (this volume). Interaction analysis of a biology chat. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs & N. Law (Eds.), Productive Multivocality in the Analysis of Group Interactions, Chapter 28. New York: Springer.

Steinhauser K, Clipp E, McNeilly M, Christakis N, McIntyre L, Tulsky J. In search of a good death: observations of patients, families, and health care providers. Ann Intern Med.2000. 132:825-832.

Suthers, D. D. (this volume). Agency and modalities in multimediated interaction. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs & N. Law (Eds.), Productive Multivocality in the Analysis of Group Interactions, Chapter 19. New York: Springer.

Teplovs, C., & Fujita, N. (this volume). Socio-dynamic latent semantic learner models. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs & N. Law (Eds.), Productive Multivocality in the Analysis of Group Interactions, Chapter 21. New York: Springer.

Acknowledgments

This work was supported by the Pittsburgh Science of Learning Center and Graduate Training Grant from the Department of Education (#R305B040063) and by the French CNRS.