Analysing Educational Dialogue Interaction: Towards Models that Support Learning

Rachel Pilkington

International Journal of Artificial Intelligence in Education (2001), 12, 1-7

Analysing Educational Dialogue Interaction: Towards Models that Support Learning

(Introduction to the IJAIED Special Issue on Analysing Educational Dialogue Interaction)

Rachel Pilkington *Computer Based Learning Unit, The University of Leeds, UK.*
R.M.Pilkington@cbl.leeds.ac.uk

The motivation for this special issue on analysing dialogue interaction was provided by the one-day workshop with the same title held at AI-Ed '99 in Le Mans. From this workshop a number of common themes and issues emerged. The call for papers for this special issue was based around these themes:

- valid and reliable approaches to identifying dialogue structures and features
- the role of dialogue in learning as evidenced by dialogue analysis
- computational models of dialogue in Intelligent Educational Systems (IESs)
- applications of dialogue analysis to Computer Mediated communication (CMC) and Computer Supported Collaborative Learning (CSCL)

Responses to the call exceeded expectation and, as a consequence, the special issue has had to be published in two parts. It seems that research in this field is now not only timely but necessary to the design of interactive systems that support learning and the evaluation of CMC in learning contexts. This editorial attempts to present an overview of the issues raised by papers in the special issue.

WHY DIALOGUE ANALYSIS AND WHY NOW?

Advances in computer based learning and the need for future developments in this area to be driven by pedagogy rather than technology has increased interest in dialogue analysis. Those working in the field are recognising the need to capitalise on different types of analysis to explore the relationship between the features of dialogues and their impact on learning. By focusing on the kinds of interaction emerging as particularly important for learning, and by abstracting formal and computational definitions of them, dialogue analysis can help bridge the gap between the empirical evaluation of interaction and the design of Intelligent Educational Systems (IESs) capable of interacting with their users. Dialogue analysis of users' interactions with such systems can also serve to evaluate their success in supporting learning. This requires triangulation with measures outside the dialogue but involves some comparison of the structure, form and strategies in dialogue with the content, focus and relevance of contributions to it.

VALID AND RELIABLE APPROACHES TO IDENTIFYING DIALOGUE STRUCTURES AND FEATURES

What sorts of dialogue analysis should we be using and what research questions can they help us with? Many authors discuss the relationship between quantitative and qualitative approaches pointing out that there are advantages and disadvantages to adopting either in isolation. There is a need for quick and effective means of analysis that can provide quantifiable data across corpora. According to Barros and Verdejo (part I of the special issue) computer mediated collaborative learning allows the recording of large amounts of data and manual approaches to fully monitor and exploit such data are "out of the question" (p. 238). On the other hand, quick quantitative approaches can generate more questions than they answer leaving us uncertain as to the underlying processes involved. For example, it is very difficult to ascertain if and when students have been addressing each other’s contributions using simple quantitative methods such as plotting the number of user accesses, the number of turns taken, or the average length of contribution. A focused in-depth qualitative analysis of part of the data is often needed (Barros and Verdejo, p.238). Moreover, further quantitative statements can be generated by applying computational processes to data after they have been qualitatively coded using computer-based tools (see also Rourke et al in part II).

The vocabulary used to describe similar types of qualitative analysis often differ (which is confusing). There is a need for some standardisation of terms in order to compare results across studies. However, some consensus seems to be emerging as to the kinds of analysis we require. For example, most qualitative coding schemes incorporate a level corresponding to the intention of an utterance, usually termed a “speech-act” or “move” and multiple such intentions may be conveyed by a single utterance (see Katz, O'Donnell and Kay in part I). Sometimes this level of analysis also includes communicative acts that are non-verbal but also convey a communicative intention (see e.g. Rosenberg & Sillince part I). Many schemes also identify patterns of such intentional acts at a higher level to try to capture either exchange structure (turn-taking patterns) or dialogue strategies (sequences of moves within a turn or over a sequence of turns that serve a particular goal). Such higher-level analyses are aimed at revealing the roles participants take, the overall structure of the dialogue or the extent to which participants engage with each other's contributions.

At a finer level of granularity than the speech act Porayska-Pomsta, Mellish & Pain (part I) look at categories of question. In common with other schemes, questions are categorised according to the ideational content that could satisfy them as a response. However, Porayska-Pomsta et al also separate out the syntactic form of the question as a separate level of analysis. This enables the effect of the surface form to be evaluated independently. In separating these levels Porayska-Pomsta et al raise a more general issue by illustrating how categorical schemes can be either too weak or too powerful to address research questions. In other words, schemes may categorise at too coarse a level to distinguish real pedagogical and communicative differences, or they may be too fine-grained to represent similarities between instances coded differently. Thus there is a need for multiple levels of discourse analysis and the selection and emphasis placed on these by the researcher depends on the researcher's particular research question. However, broadly speaking, analysis aims to triangulate aspects of dialogue structure or form with measures of content quality (or relevance of contribution) in achieving domain task-goals (or learning objectives).

Content analysis is acknowledged to be particularly difficult to perform on large data sets (see e.g. Rourke et al in part II). In general, it is more difficult to reuse content analysis schemes in different contexts and the researcher often has to generate their own set of codes that apply to a specific domain rather than adapting an existing scheme which has already been used successfully and reliably. This can limit the granularity of the analysis or the volume of data analysed in the available time. This in turn may limit the scope of conclusions drawn. However, the stages in development of content analysis schemes and other kinds of qualitative analysis "developed from scratch" remain the same:

- preparing data transcripts/protocols for coding
- deciding the units of behaviour to be coded
- coding samples of data to test the applicability of the coding scheme
- testing intra-rater reliability
- training others to use the scheme and testing the inter-rater reliability

As Rourke et al point out computer-based tools (e.g. Atlas.ti®, NUD*IST®) can be used to assist at each stage in this process including supporting and in some cases partially automating the coding of protocols and tracking the consistency of coders. However, qualitative manual coding involves subjective judgement and the reliability of dialogue analysis schemes (inter-rater and intra-rater) remains a contentious issue (see e.g. Rourke et al part II, Kneser, Pilkington & Treasure-Jones, part II). The number of categories in a scheme complicates this with schemes having a larger number of potential categories to choose from showing poorer reliability through increased chance mismatch. This makes comparison across schemes difficult unless the kappa statistic is given (see Chi, 1997).

However not all schemes report a kappa value. Rourke et al (part II) point out, in their review of 19 studies performing content analysis, that only three studies reported a kappa statistic. Six studies reported no figures for reliability at all; one study reported a reliability-coefficient and eight reported percentage agreement.

As this discussion highlights, there is a need to apply the same methodological techniques to different corpora as well as applying different schemes and techniques to the same corpora if we are to be able to make more general claims with respect to the role of interaction in learning.

THE ROLE OF DIALOGUE IN LEARNING AS EVIDENCED BY DIALOGUE ANALYSIS

What role(s) might dialogue play in acquiring domain concepts and skills and how does dialogue analysis help to inform us about these roles? Activity Theory extensions to the Vygotskian notion of "distributed mind" suggest a useful framework within which to study the importance of the dialogue community in the construction of knowledge (see e.g. Barros and Verdejo part I, Ekeblad, 1998), but many questions still need to be answered. In particular, what are the participatory mechanisms and genres of interaction that lead to effective collaborative learning? Do both individuals have to externalise their reasoning, or is it enough to pose questions prompting self-explanation? Is collaborative learning always more beneficial than whole-class tutor-led interaction? Wegerif and Mercer (1996) suggest that neither the notion of dialogue as a vehicle for resolving cognitive conflict, nor dialogue as a vehicle for externalising the reasoning of a more expert peer are sufficient to account for the positive effects of collaborative learning. However, it seems that learning gains are rarely maximised when students are not pre-trained in the collaboration process (e.g. Crook 1997).

Berzsenyi (1999) argues that students need to learn dialogue skills that are sensitive to relationships between participants in debate. Veerman, Andriessen & Kanselaar (2000) argue that students need to learn to structure and focus the content of the debate towards task goals. Many of the papers in this special issue use dialogue analysis to uncover tutoring strategies for scaffolding such skills with the aim of designing pedagogical agents capable of scaffolding them in users of their IESs.

Using even relatively simple and quantitative approaches to dialogue analysis such as counting the numbers of contributions by participants, it is possible to gain some insights into the roles students and tutor take in the collaborative process (e.g. Barros and Verdejo part I). More detailed analysis can help us to profile individual students, to see what roles they are adopting in the dialogue and tell us how actively students and tutor must engage with each other's contributions for learning to occur.

Many studies have suggested that actively participating in inquiry dialogue is beneficial (see e.g. Porayska-Pomsta, Mellish & Pain, part I; Person et al, part II) and that the quality and not the quantity of inquiry is central. Craig et al (part I) designed animated agents to simulate a tutor and student engaged in dialogue which human students listened to. Those who listened to the dialogue recalled more and asked more deep-level reasoning questions than those listening to monologue. They concluded that modelling dialogue involving asking questions could enhance the quality of student engagement with learning materials.

There is also some evidence that "successful" exchanges with respect to collaborative learning and problem-solving are more likely to include the kinds of moves identified by Wegerif & Mercer (1996) as indicative of "exploratory dialogue" such as explaining, clarifying, challenging and justifying. From detailed analysis of the co-occurrence of such speech-acts and their position within exchanges we can begin to suggest common strategies for directing and modelling the lines of reasoning which tutor and student(s) engage in to complete learning tasks successfully.

In Cook's MetaMuse system (Cook, part II) tutor-student dialogue was examined in the open-ended domain context of mentoring creative composition. Underlying sub-goals identified in the tutor's mentoring strategy included probing and targeting – to first get the student to identify a goal to work towards and then elicit self-explanations aimed at gap-filling, clarifying and giving reasons for choices made. Patterns or cycles of targeting and probing occurred as task sub-goals were opened and closed.

This illustrates that by integrating task analysis with dialogue analysis it is possible to trace when and what kinds of dialogue are prompted by which features of tasks or tools and triangulate these with the successful completion of task sub-goals. For example Luckin et al (in part II) looked at interactivity between learners as they navigated the Galapagos multimedia learning environment (MLE). By using a charting technique they tracked the focus of students' interactions (what they were discussing) when they accessed particular pages or tools they were able to compare and contrast three different guidance conditions within the MLE. They found that the type of guidance had an effect not only on the access of MLE pages and tools but also on the degree to which students discussed the task and answer construction. Moreover, writing a document using the notepad tool alongside accessing pages prompted more reflective on-task talk between students - those guided by a conventional linear path engaged less often in such talk.

Katz O'Donnell & Kay (part I) also looked at focus in dialogue at different points in the task, particularly the focus of the tutor's interactions in helping students overcome misconceptions when electronic trouble-shooting. Misconceptions were most likely to be resolved if the tutor not only discussed the misconception during problem-solving but also used probing, feedback and post-summarisation strategies to address the misconception after the problem-solving phase had ended.

From the above discussion we can begin to see how different levels of dialogue analysis can be combined to give insights into the kinds of learning taking place with and through computer based systems and so inform the design of IESs.

COMPUTATIONAL MODELS OF DIALOGUE IN INTELLIGENT EDUCATIONAL SYSTEMS (IES)

What models of dialogue do IESs require if they are to enhance learning? Many authors have as their goal the design of pedagogical agents capable of interacting effectively with their users to facilitate learning. Dialogue analysis techniques, like those described above, can be employed to suggest the features of dialogue they need to build into interaction scenarios that their agents will use to engage students (see for example Craig et al, Ravenscroft & Pilkington, part I; Cook, Rourke et al, Person et al part II).

For example AutoTutor (Person et al part II) is an animated pedagogical agent that serves as a conversational partner and whose moves (communicative acts) include some animated gestures as well as synthesised speech. An analysis of the dialogue of human tutors was used to identify tutoring moves. The pedagogical agent attempts to simulate moves preferred by human tutors. These include prompting, hinting and pumping for missing information - the "gaps" in problem solving.

Cook (part II) similarly uses descriptive dialogue models generated from empirical investigation and formalises these (as state transitions networks) to model tutoring strategies that support creative problem-solving in the domain of musical composition. MetaMuse attempts to model a mentoring approach in which creative, metacognitive and critical thinking skills are supported by the tutor in this open-ended domain. Unlike AutoTutor this system does not attempt to understand natural language and like the system by Ravenscroft & Pilkington (part I) similarly limits student input to a predefined set of moves. Ravenscroft's CoLLeGE system's "common-sense reasoning module" calls up tutoring tactics – challenge, persuade, resolve and probe which adopt sequences of moves to try to stimulate belief-revision.

In the development of such systems dialogue analysis may also be used both to validate and evaluate such pedagogical agents. Validation of models involves testing that they accurately simulate aspects of the natural dialogues that inspired their conception. This stage involves testing the model's output against human dialogues to see whether the model produces expected output. This can involve the simulation of students or a student and tutor. For example Burton, Brna & Pilkington (2000) simulate two students collaborating to solve a problem involving wiring a simple circuit. Validation of the dialogue model is by comparison of the output of dialogues at the move level of analysis with that of human students solving the same problem. Rourke et al report simulating several different kinds of virtual student to validate their model. At an earlier stage in the development process models can be formatively tested using a Wizard of Oz technique in which a human tutor is restricted to the moves the simulation is capable of. The tutor then comments on the degree to which the system interface constrains the choice of expression (see Ravenscroft & Pilkington part I).

Using constrained menu-based move entry means the user essentially pre-codes their own moves in the dialogue making it, in principle, easier to automatically track and profile the changing uses of move-types by individual users over time. However, further work is needed before evaluations with student users can tell us if these systems are more effective in meeting learning objectives.

APPLICATIONS OF DIALOGUE ANALYSIS TO COMPUTER MEDIATED COMMUNICATION (CMC) AND COMPUTER SUPPORTED COLLABORATIVE LEARNING (CSCL)

There are notable advantages and disadvantages to forms of CMC and CSCL tools – different features of these tools may create specific niche affordances for specific learning objectives. Herring (1999) reviews various genres of CMC and concludes that text-based media may have positive advantages as well as disadvantages that stem from their limited bandwidth, lag and the tendency to cause the decoupling of turns. Communicating through internet rely chat or listserve discussion groups may promote deeper and more reflective contributions than is typical in face-to-face discussion and the absence of visual cues may encourage participants who would not normally contribute in discussion to participate (Walker and Pilkington, 2000). The lack of non-verbal communication may also increase focus on the task.

However, it has equally been argued that the lack of such cues depersonalises discussion and makes it relatively difficult to provide the normal back-channels that encourage effective collaboration through positive feedback. This may create a sense of isolation. Rosenberg and Sillince (part I) looked at the absence of non-verbal communication in CMC versus face-to-face collaboration and found that non-verbal communication supported social activities that could impact on problem-solving and task completion, for example ease of requesting and getting commitment, and recognising the effort of others. They concluded that lean media such as CMC might make successful collaboration difficult because of the absence of such cues.

Veerman, Andriessen & Kanselaar (2000) argue that CMC has useful properties for developing students' debating skills but they also argue for teaching students strategies to structure/focus the debate on task goals. Veerman et al suggest encouraging students to challenge points of view during the discussion, ask for reasons and explanations or suggest counter-arguments, refocus on the issue when the discussion is diverted off-task and elaborate and explain ideas or prompt others to do so. These roles echo much of the earlier discussion on the kinds of moves adopted by the tutor when mentoring. However, before we can scaffold the use of such moves in student debate we need to first scaffold a discourse community in which students are confident to communicate. For example, Berzsenyi (1999) emphasises the need to establish the ground rules for a safe discussion: to discourage disruptive off-task behaviour and non-constructive criticism and to encourage equal and balanced participation by inviting alternative views and offering positive feedback.

By applying analyses such as exchange structure analysis (ESA) in conjunction with speech-act analysis to CMC data it is possible to compare tutor and student dialogue. In this way Kneser, Pilkington and Treasure-Jones (part II) were able to show that the tutor was successful in modeling many of the argument roles Veerman et al suggest should be encouraged in CMC discussion. However, if the tutor's aim was to encourage students to take on these roles for themselves, then there was a long way to go.

As Rourke et al (part II) point out, CMC research needs instruments and techniques that are capable of providing insightful analyses of CMC discussion if we are to move beyond naïve enthusiasm for computer conferencing towards establishing exactly what the technology can deliver.

Future work needs to investigate whether participant profiles from analyses of this type can help tutor and students reflect upon their own role in the dialogue and help them to make changes to the moves they take as a result. Tools could be developed to provide individual profiles automatically by recording the moves students make (this need not be by understanding natural language but could be by students selecting "moves" to preface dialogue turns with - see for example Robertson, Good & Pain, 1998, Ravenscroft in part I). Learning companions might then be engineered to make similar use of this data. Such agents would be able to adapt their strategies to the profile of the student they are interacting with and so encourage students to engage more effectively in discussion (see e.g. Rourke et al in part II).

SUMMARY

The goals of educational dialogue analysis are to identify the features that distinguish instructional discourse from other types of discourse and to determine what makes it effective. There is a need to continue to capitalise on different types of analysis to explore the relationship between the features of dialogues and their impact on learning. By focusing on effective patterns of dialogue emerging from the analysis of student-tutor and student-student interaction and by abstracting formal and computational models of these, dialogue analysis is helping to bridge the gap between empirical investigation of interaction and the design of Intelligent Educational Systems (IESs) that interact with students. Dialogue analysis of users' interactions with such systems is helping to validate these systems and to evaluate their success in supporting learning. The results of this research is shaping the development of the next generation of animated pedagogical agents which make use of both verbal and non-verbal communicative acts and is driving the design of more effective multimedia learning environments that encourage active engagement and interaction, both with them and around them. The development of computer based tools for performing analyses of CMC protocols is also helping us to research the effectiveness of CMC as a substitute for face-to-face discussion in distance learning and to supplement face-to-face tutoring out of hours. This special issue has brought together papers that address these themes.

ACKNOWLEDGEMENTS

Thanks are due to all those who reviewed articles for this special issue and all the IJAIED team who made this special issue possible - especially IJAIED Editor John Self for his comments. Thanks are also due to all the participants at AIED '99 Le Mans workshop on dialogue analysis - especially Helen Pain, Jean McKendree, Paul Brna, Elizabeth Delozanne, Nicolas Balacheff and Susan Lajoie for their help in organising the workshop.

REFERENCES

[Other than papers appearing in part I and II of the special issue]

Atlas.ti http://www.atlasti.de/download.htm (28/03/01)
NUD*IST http://www.qsrinternational.com/home/home.asp (9/04/01)
Berzsenyi, C. A. (1999). Teaching interlocutor relationships in electronic classrooms. *Computers and Composition*, 16, 229-246.
Burton, M., Brna, P., & Pilkington, R. M. (2000). CLARISSA: A Laboratory for the Modelling of Collaboration. *International Journal of Artificial Intelligence in Education*, 11, 79-105.
Chi, M.T.H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *The Journal of the Learning Sciences, 6* (3), 271-315.
Crook, C. (1997). Children as computer users: the case of collaborative learning. *Computers & Education*, 30(3/4), 237-247.
Ekeblad, E. (1998). Contact, community and multilogue - electronic communication in the practice of scholarship. Paper presented at the *Fourth Congress of the International Society for Cultural Research and Activity Theory. ISCRAT 1998*. Denmark, June 7-11, 1998. http://hem.fyristorg.com/evaek/writings/iscrat98/cocomu.html.
Herring, S. (1999). Interact ional Coherence in CMC. *Journal of Computer-Mediated Communication*, 4 (4). http://www.ascusc.org/jcmc/vol4/issue/4/herring.html
McKendree, J., Stenning, K., Mayes, T., Lee, J., & Cox, R. (1998). Why observing a dialogue may benefit learning. *Journal of Computer Assisted Learning*, 14, 110-119.
Robertson, J., Good, J., & Pain, H. (1998). BetterBlether: The design and evaluation of a discussion tool for education. *International Journal of Artificial Intelligence in Education*, 9, 219-236.
Walker, S. A., & Pilkington, R. M. (2000). Networked Communication and the Collaborative Development of Written Expression at Key Stage Three. *Proceedings of the Second International Conference on Networked Learning 2000* Lancaster 17th-19th April.
Wegerif, R., & Mercer, N. (1996). Computers and Reasoning Through Talk in the Classroom. Language and Education, 10 (1), 47-64.
Veerman, A.L., Andriessen, J.E.B., & Kanselaar, G. (2000). Learning through synchronous electronic discussion. *Computers and Education*, 34, 269-290.