Discourse, computation and context – sociocultural DCLA revisited

Conference or Workshop Item

Discourse, Computation and Context – Sociocultural DCLA Revisited

Simon Knight[1] and Karen Littleton[2]
[1] Knowledge Media Institute
[2] Centre for Research in Education & Educational Technology
The Open University, Milton Keynes, MK7 6AA, UK
+44 1908 654672
{firstname.lastname}@open.ac.uk

ABSTRACT
This paper expands the sociocultural analysis of earlier discourse centric learning analytics (DCLA) to discuss the pedagogic functions of discourse, and the implications of these functions for DCLA. Given the importance of discourse for learning [13], and the potential of computers to (a) scaffold effective discourse and (b) give meaningful feedback on such discourse, it is important that DCLA are well theorised. Sociocultural theory emphasises context, and discourse "in action" in its analysis. If DCLA wishes to ground itself in such theory, work will need to be done to address these aspects of discourse in computational analysis. Given the potential of DCLA to provide support for educational talk – an important aspect of learning – research should be conducted to further develop DCLA approaches to such talk.

Categories and Subject Descriptors
K.3.1 [**Computers and Education**]: Computer Uses in Education – *collaborative learning.*

General Terms
Measurement, Documentation, Design, Human Factors, Theory,

Keywords
Learning analytics; epistemology; pedagogy; educational assessment; discourse analytics; social learning analytics

1. INTRODUCTION
Discourse Centric Learning Analytics (DCLA) provide opportunity to explore: the discursive elements of learning; the ways that learners construct meaning individually, with teachers, and collaboratively; and the sorts of language which might be most associated with positive learning outcomes.

"Wherever education is taking place, commonality – a shared perspective – is key, and dialogue[1] is the tool used to create such a perspective [2]" [11]. This shared perspective has been termed "common knowledge" [2], the body of shared contextual knowledge which is built up through discourse and joint action, and forms the basis for further communication. Thus, in this perspective, "common knowledge" forms a key constitutive part of context for speakers in a discourse, as well as being a fundamental aspect of education – in which a mutuality of understanding is crucial.

Indeed, the strong consensus among researchers is that in a variety of contexts, high quality dialogue is associated with learning (see the collection edited by Littleton and Howe [13]). That research shows that, "Engaging children in extended talk which encourages them to 'interthink' and explain themselves…stimulates both their subject learning, and general reasoning skills [15,17,18,25], as well as their social and language skills [31]" [11].

[1] Dialogue and discourse are used interchangeably in this paper

One socioculturally motivated example which DCLA has engaged with is discourse which is described as 'exploratory'. 'Exploratory talk' has been shown to be associated with positive educational outcomes. In exploratory talk:

> *Partners engage critically but constructively with each other's ideas. Statements and suggestions are offered for joint consideration. These may be challenged and counter-challenged, but challenges are justified and alternative hypotheses are offered. Partners all actively participate, and opinions are sought and considered before decisions are jointly made. Compared with the other two types, in exploratory talk knowledge is made more publicly accountable and reasoning is more visible in the talk.* [16:59]

Similar characterisations of effective dialogue have emerged from the work of other researchers across a range of ages [20,24]. Computational linguistic approaches have had some success in applying this educational research to the identification of a variety of discourse markers indicative of high quality dialogue, or written text. In discourse analysis, markers for such talk include explanatory terms such as *for example; I think; because/'cause; if; also*. Researchers working on Learning Analytics have shown that this sort of talk can be identified in computer-mediated-communication (CMC) contexts [3,4,12].

Given the potential of such systems to provide support for high quality discourse – important in its own right, and as a means to develop learning and reasoning – identifying problematic, or high quality dialogue to provide scaffolding may have potential to support learning through such systems. This paper remains agnostic on the pedagogic scaffolding for such an approach (assessment, scaffolding, feedback, corrections, etc.) instead highlighting some aspects of sociocultural approaches which may present challenges.

This short paper first introduces a concern regarding the nature of context, both for DCLA and sociocultural research. It explicates this with respect to discussions in those communities, highlighting some corresponding concerns. It goes on to offer one perspective on sociocultural and web technologies with relation to the importance of context, suggesting this is an area in need of attention. This is particularly highlighted with respect to the apparent functions of dialogue in pedagogic contexts.

2. THE CHALLENGE OF CONTEXT
Educational researchers within the sociocultural tradition would highlight the importance of dialogue as not only constitutive *of* context (that is, representing context), but constituted *in* context (that is, creating context).

Recently Littleton and Mercer [14] consider the complexity of common knowledge context as both historical and dynamic :

> *Successful interthinking requires partners to have, and to develop, a foundation of common knowledge to underpin their*

discussions. We have distinguished two types of common knowledge, both of which can be important. The first of these is accumulated through the activities of a group, as members develop a shared history. They have knowledge in common because it has been generated by their joint activities and associated conversations. It is the kind of common knowledge which allows a teacher to refer only briefly to the content of a previous lesson and expect students to have some recollection what it had been about. We have called this dynamic common knowledge, because it is produced by the dynamics of the group's own extended activity. The second type, which we call background common knowledge, is that which any established member of a community of practice can take for granted as being shared with other members and does not therefore need to be explained from first principles. It is the kind of knowledge which enables any two physicists, Beatles fans or people who have grown up in the same town to take certain kinds of understanding for granted, even if they have never met before.

This distinction highlights the need to understand that context should not only be assumed from the state of the dialogue at any particular point (assuming dialogue represents context), but rather, we should also explore the ways in which the context changes over time as a feature of the dialogue (assuming dialogue involves the co-construction of context).

It is in part due to this consideration that sociocultural researchers have emphasized the use of both qualitative and quantitative methods, in which – in contrast to some other qualitative methods – the quantitative data is taken to aid the understanding of the qualitative, as opposed to the converse. It is thus that such researchers often include excerpts of talk, concordance analysis, and other contextual markers such as cohesive ties (see below) in their reporting. Another technique is drawn from 'systemic functional linguistics', which takes it that types of text have contexts by being members of a particular genre, which is revealed through the way such texts are written[2] – thus, context is imbued into texts at the time of writing. In sociocultural discourse analysis, this assumption is adapted from that of 'texts' to the co-construction of context through dialogue in which "'context' is created anew in every interaction between a speaker and listener or writer and reader. From this perspective, we must take account of listeners and readers as well as speakers and writers, who [co]create meanings together" [19:21]. It is thus that sociocultural researchers may seek to understand the temporal aspects of context, as involving continuity across talk, by looking for repetition of words, synonyms and ways of approaching problems, to understand how "speakers can jointly, co-operatively create cohesion in…their speech" [19:62].

This focus on the repetition of words raises an interesting problem for DCLA. Sociocultural researchers are interested in the purposes of dialogue – pointing out that it is more than just an exchange, transfer, or transaction of private thoughts (see below). In particular, in learning contexts we are interested not only in the deployment of particular terms, but of their *effective deployment* in contexts which indicate an understanding of the inferential properties of the term to other concepts. Analytics which explore key words in abstracted ways may obscure the misuse of terms, or – moreover – their simple copying from texts which students use for contextualizing purposes, such as task instructions. Describing what is happening is not the same as understanding what is being done. For example, the same question might be asked at the beginning and end of a lesson, while serving different functions in each instance.

The importance of context is also a familiar topic in Natural Language Processing (NLP). In accord with their psycho-social analogues in the educational world (sociocultural theory, discursive psychology, many varieties of humanistic psychology) much emerging work has eschewed the focus on cognitivist models which seek to understand the *beliefs* and *intentions* of agents, instead focusing on attempts to understand the *contextual* and *action-based* nature of talk, as a thing "to do" rather than its role in the abstracted expression of underlying beliefs. This is the shift "from the view of language as a tool of representing the world to its view as a means of interacting with the world." [22:39] (see also the rest of this Special Issue of *Pragmatics and Cognition* devoted to discussing the work of Robert Brandom).

It is thus that researchers working in NLP and semantic web have proposed a shift away from attempts to *represent* non-dynamic properties of objects (pages, OER, dialogue, etc.) towards a pragmatic understanding of the semantic properties of entities as contextually salient [28]. This salience of context may be seen both in the sense of context's salience to the semantic content of an expression, and in the sense that the semantic content of an expression is salient to the context of use; expressions are tied to commitments and entitlements [9] – they exist within a particular context, and they create context as enactive speech acts.

This interest in context for discourse has come both in the form of an interest in the "Pragmatic Web" [1,28] and focus on means for tracking context, and pragmatic semantics in NLP (see e.g. the special issues introduced by [8,23]). In these works – particularly [23] – the emphasis is on language not as a means of exchange, but as a mode of doing, and creating common knowledge.

This reflects the nature of sociocultural approaches to discourse analysis in which context is taken as both a feature of, and a feature *in* the discourse. That is, discourse both exists within a particular context and mediates it, and it creates context, as a dynamic, collaborative, discursive property.

3. DISCOURSE, COMPUTATION, AND CONTEXT – AN EXAMPLE

This emphasis on the *purposes* of language, and its multi-faceted deployment, is particularly salient to education – particularly if we seek to design analytics to identify pedagogically functional language in discourse episodes, and support students' use of such language in their learning.

In the sociocultural educational – particularly school based – context , "With respect to direct pedagogical functions (as opposed to social functions such as behaviour management), dialogue seems to serve several purposes" [11]. is a preliminary sketch of how these pedagogical functions (column 1) might be used to drive DCLA design, outlining the properties of the functions (column 2); the types of content seen in each of these cases (column 3); and the linguistic objects which would form the basis of any NLP technique (column 4) exploring the respective functions.

[2] See Halliday, Hasan and Christie [5]

Function (from [11])	**Focus**	**Content**	**DCLA**
supporting individuals' subject learning	advancement of subject knowledge	Propositional, semantic content of what is said. Appropriate reference to entities and relationships from the curriculum.	Techniques for recognising named entities, esp. ontology-driven techniques for domain specific entity-relationship extraction (e.g. history; biology) [26,30]
supporting psychological development – the development of oral language and reasoning skills	development of argumentation skills or dispositions	Pragmatic, rhetorical content – how it is said. Use of language to signal the making of claims/discourse moves in dialogue or writing	Techniques for extracting rhetorical forms (e.g. exploratory talk [3,4,12]. ; unresolved problems; new findings; contrasting results [27])
promoting whole class and small group understanding or commonality	the interactional nature of collaborative dialogue	Who says what to whom? The actors constituting the interpersonal context	Techniques for identifying interlocutors in dialogue, and more diffuse social networks of strong/weak ties via different channels [6,21,29]
enabling sharing of ideas that can be improved together (both whole class and small group)	the concepts, themes, and ideas which emerge through the dialogue, cohesive ties and creation of external artifacts as 'improvable objects' [32], as resources for continuity and progression (as aspects of 'common knowledge)	The context in which it is said and its outputs (the discourse itself as a cognitive, social – and when persistent, material – resource)	The temporal and interactional nature of the above 3 rows, and the nature of the shifts between states – the *pragmatic* context. (Possibly analyzable through techniques such as Epistemic Network Analysis [33])

Table 1 – Dialogue Functions, Focus, Content and DCLA

The aim of the table is to highlight the multiple ways in which any segment of dialogue may be probed, the different properties of such analysis, and the ways in which such analysis might interact[3]. This is particularly important if, through our data collection, task setting, and analytic methods, we reify a particular perspective on the purposes of dialogue which neglects the multi-functionality of dialogue. This is true in both the context of summative and formative assessments of dialogue quality. In the former case, we should be careful to avoid giving feedback on only one aspect of dialogue, of obscuring the nature of the dialogue in attempts to reduce data to manageable, neat categories (thus enabling visualization), and of failing to appropriately scaffold both teachers and students in their use of talk for learning. In the latter case, alongside those concerns the issue of performativity is raised – wherein those aspects of a curricula which are directly assessed, become those aspects which are most focussed on in classroom contexts. Indeed (as with any assessment instrument) there is the risk that as certain forms of DCLA become widespread, teaching and assessment become driven by those facets of dialogue that have been made visible by these techniques.

While DCLA of various types may be useful where deployed for particular purposes [10] the sociocultural claim is that if our object of enquiry is the discourse itself, both as a representation of and tool for learning, then we ought to be interested in the ways that discourse is used to create 'common knowledge'. For that use, approaches which address a single row of the table will not be adequate. As the table highlights, work is developing in this area, and the parallel discussions on the pragmatic web may also offer insight. Fundamentally, as a recent CSCW paper claims, "Unlike a document, which offers a snapshot of a relatively stable distribution of topics, a conversation is an unfolding process in which the definition of a topic is continually renegotiated. The lack of a relatively well-structured, intentionally designed document limits the utility of the machine-learning approaches that drive modern topic-tracking algorithms." [7:341]. In that paper they thus "take as our unit of analysis a sequence of replies, seek[ing] to understand how clusters of words in these reply sequences change, merge, and split" with a particular interest in modelling the statistical properties of the co-occurrence of words over time, as opposed to modelling probabilities based on dictionary entries or other corpora. Of course, in the educational context we are not only interested in the current learning of individuals in isolation, but also in how language dynamically resources future learning.

By gathering such temporally located data we:

1. identify the apt-concepts for any particular discourse (row 1)
2. understand the network of intercolutors and their contributions – both in terms of their use of concepts, and their contributions to building shared knowledge, to engaging in interthinking and collaboration (rows 2 and 3)
3. and understanding how these discourses are related across time and location (row 4)

Thus while current approaches to sociocultural DCLA go some way to identifying learning talk, further work should explore the ways that such patterns of talk are used in dynamic highly localised contexts for building common knowledge.

[3] We are aware that under one reading of the table it might appear that functions map the content of utterances to their focus. We leave it to the reader to decide if this is indeed the case.

Talk is a crucial part of learning, and new developments in DCLA afford opportunity for new analysis of this dynamic human interaction, including the potential for novel new formative and summative assessments. However, talk is a multi-faceted, co-constructed, and dynamic tool. The DLCA community must tread cautiously in its deployment of tools, and in developing tools with limited – but perhaps unstated – views on the nature of language use for learning.

4. REFERENCES

1. Buckingham Shum, S., Lind, M., and Weigand, H., eds. *Proceedings ICPW'07: 2nd International Conference on the Pragmatic Web, 22-23 Oct. 2007, Tilburg: NL*. 2007.
2. Edwards, D. and Mercer, N. *Common knowledge: the development of understanding in the classroom*. Routledge, London, UK, 1987.
3. Ferguson, R. and Buckingham Shum, S. Learning analytics to identify exploratory dialogue within synchronous text chat. *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, (2011), 99–103.
4. Ferguson, R., Wei, Z., He, Y., and Buckingham Shum, S. An Evaluation of Learning Analytics to Identify Exploratory Dialogue in Online Discussions. ACM Press: New York (2013).
5. Halliday, M.A.K., Hasan, R., and Christie, F. *Language, Context, and Text: Aspects of Language in a Social-Semiotic Perspective*. Oxford University Press, USA, 1989.
6. Haythornthwaite, C. and De Laat, M. Social Networks and Learning Networks: Using social network perspectives to understand social learning. *7th International Conference on Networked Learning*, (2010).
7. Introne, J.E. and Drescher, M. Analyzing the flow of knowledge in computer mediated teams. *Proceedings of the 2013 conference on Computer supported cooperative work*, ACM (2013), 341–356.
8. Iwanska, L. and Zadrozny, W. Introduction to the Special Issue on Context in Natural Language Processing. *Computational Intelligence 13*, 3 (1997), 301–308.
9. Kibble, R. Reasoning About Propositional Commitments in Dialogue. *Research on Language and Computation 4*, 2-3 (2006), 179–202.
10. Knight, S., Buckingham Shum, S., and Littleton, K. Epistemology, Pedagogy, Assessment and Learning Analytics. ACM Press (2013).
11. Knight, S. Creating a supportive environment for classroom dialogue. In S. Hennessy, P. Warwick, L. Brown, D. Rawlins and C. Neale, eds., *Developing interactive teaching and learning using the IWB*. Open University Press, 2013.
12. De Liddo, A., Buckingham Shum, S., Quinto, I., Bachler, M., and Cannavacciuolo, L. Discourse-centric learning analytics. *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, ACM (2011), 23–33.
13. Littleton, K. and Howe, C. *Educational dialogues: understanding and promoting productive interaction*. Routledge, Abingdon, Oxon, 2010.
14. Littleton, K. and Mercer, N. *Interthinking: putting talk to work*. Routledge, London, 2012.
15. Mercer, N., Dawes, L., Wegerif, R., and Sams, C. Reasoning as a scientist: ways of helping children to use language to learn science. *British Educational Research Journal 30*, 3 (2004), 359–377.
16. Mercer, N. and Littleton, K. *Dialogue and the Development of Children's Thinking: A Sociocultural Approach*. Routledge, 2007.
17. Mercer, N. and Sams, C. Teaching children how to use language to solve maths problems. *Language and Education 20*, 6 (2006), 507–528.
18. Mercer, N., Wegerif, R., and Dawes, L. Children's Talk and the Development of Reasoning in the Classroom. *British Educational Research Journal 25*, 1 (1999), 95–111.
19. Mercer, N. *Words & Minds: How we use language to think together*. Routledge, Oxon, 2000.
20. Michaels, S., O'Connor, M.C., Hall, M.W., and Resnick, L. Accountable talk: classroom conversation that works. *Pittsburg: University of Pittsburgh*, (2002).
21. Oshima, J., Oshima, R., and Matsuzawa, Y. Knowledge Building Discourse Explorer: a social network analysis application for knowledge building discourse. *Educational Technology Research and Development 60*, 5 (2012), 903–921.
22. Peregrin, J. The nature of meaning: Brandom versus Chomsky. *Pragmatics & Cognition 13*, 1 (2005), 39–57.
23. Piwek, P. Perspectives on Dialogue: Introduction to this Special Issue. *Research on Language and Computation 4*, 2-3 (2006), 143–152.
24. Resnick, L.B. Making America smarter: The real goal of school reform. *Developing minds: A resource book for teaching thinking*, (2001), 3–6.
25. Rojas-Drummond, S., Littleton, K., Hernández, F., and Zúñiga, M. Dialogical interactions among peers in collaborative writing contexts. In K. Littleton and C. Howe, eds., *Educational dialogues: Understanding and promoting productive interaction*. Routledge, Abingdon, Oxon, 2010, 128–148.
26. Rosé, C., Wang, Y.-C., Cui, Y., et al. Analyzing collaborative learning processes automatically: Exploiting the advances of computational linguistics in computer-supported collaborative learning. *International Journal of Computer-Supported Collaborative Learning 3*, 3 (2008), 237–271.
27. Sándor, Á. Modeling metadiscourse conveying the author's rhetorical strategy in biomedical research abstracts. *Revue fran\ccaise de linguistique appliquée 12*, 2 (2007), 97–108.
28. Schoop, M., Moor, A. de, and Dietz, J.L.G. The pragmatic web: a manifesto. *Commun. ACM 49*, 5 (2006), 75–76.
29. Sie, R.L.L., Ullmann, T.D., Rajagopal, K., Cela, K., Bitter–Rijpkema, M., and Sloep, P.B. Social network analysis for technology–enhanced learning: review and future directions. *International Journal of Technology Enhanced Learning 4*, 3 (2012), 172–190.
30. Tablan, V., Roberts, I., Cunningham, H., and Bontcheva, K. GATECloud.net: a platform for large-scale, open-source text processing on the cloud. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences 371*, 1983 (2013), 20120071.
31. Wegerif, R., Littleton, K., Dawes, L., Mercer, N., and Rowe, D. Widening access to educational opportunities through teaching children how to reason together. *Westminster Studies in Education 27*, 2 (2004), 143.
32. Wells, G. *Dialogic inquiry: towards a sociocultural practice and theory of education*. Cambridge University Press, 1999.
33. Epistemic Network Analysis: A Prototype for 21st-Century Assessment of Learning. *International Journal of Learning and Media 1*, 2 (2009), 33–53.