M. Cherubini, J. van der Pol, and P. Dillenbourg. Grounding is not shared understanding: Distinguishing grounding at an utterance and knowledge level. In *CONTEXT'05, the Fifth International and Interdisciplinary Conference on Modeling and Using Context*, Paris, France, July 5-8 2005.

Grounding is not Shared Understanding: Distinguishing Grounding at an Utterance and Knowledge Level*

Mauro Cherubini[1], Jakko van der Pol[2], Pierre Dillenbourg[1]

[1] CRAFT, Swiss Federal Institute of Technology, Lausanne (EPFL),
Bât. CE1.631, Station 1
CH-1015 Lausanne, Switzerland
{Mauro.Cherubini, Pierre.Dillenbourg}@epfl.ch
http://people.epfl.ch/Mauro.Cherubini
[2] IVLOS, University of Utrecht, Heidelberglaan 8
kamer 384 (Bestuurgebouw)
NL-3508 Utrecht TC, The Netherlaands
j.vanderpol@ivlos.uu.nl

Abstract. This paper argues that for the study and facilitation of collaborative learning, existing theories of grounding such as that of Clark and Shaefer [5] cannot be applied without adjustments. When comparing collaborative learning and conversation, four dimensions can be identified where grounding at a knowledge level differs from the grounding at an utterance level. Firstly, the indirect access and the existence of a range of manifest meanings, poses the need for a notion of 'groundedness'. Secondly, we propose providing evidence in 'co-referenced actions' to be an important process as well as an additional marker to assess grounding. Thirdly, instead of simply repairing misunderstandings after they arise, 'perspective taking' becomes a more prominent mechanism. Fourthly, effort into grounding is turned from needing to be minimised, into needing to be 'optimised'.

1 Introduction

Many studies of Computer Supported Collaborative Learning (CSCL) that identify grounding as an important process, analyse it using the theory of (or models based on) Clark and Shaefer [5]. However, the application of their theory within the field of CSCL holds some problems. As a linguistic theory, it analyses conversation on a micro or 'utterance' level and is not developed to describe the macro or 'knowledge' level, which is associated with learning. While the micro level focuses on the dialogue interchange occurring between two or more interlocutors, the macro level refers to the shared understanding that is constructed as a consequence of that exchange [10]. We

* Collaboration for this paper was partly made possible by funding from the Dutch Organisation for Scientific Research (NWO).

argue that the observable presentation and acceptation of utterances, as described in Clark and Shaefer's contribution theory, cannot automatically be translated into the sharing of knowledge. As Koschmann's [14] example of a learning conversation between surgeon and student in an operation room shows, even repeated presentation and acceptation phases of a concrete referent in a shared environment, can result in different personal representations at a knowledge level.

Since language is not a direct translation of a speaker or writer's knowledge, the interaction between knowledge and language that we find within CSCL, is a complex one [1]. While everyday human interaction has developed to be very efficient in the recognition of mutual intentions, communicating about knowledge (or 'semantic grounding' [2]) cannot automatically rely on the same unproblematic and self-regulating character of 'grounding-for-conversation'. Our reason for stressing this, is while we believe in the great potential of communication to produce learning, we want to caution that not all communication will automatically do so. When analysing or designing for collaborative learning, we need to take into account the idea that successful conversation is not necessarily the same as successful knowledge sharing.

To explore the complex interaction between conversation and learning, this article will investigate the (subtle) differences of the characteristics, evidence, principles and mechanisms of grounding at the micro and macro level. To give this a practical context, we will present two examples: the use of mobile messaging for a spatial collaboration task 1.1, and the use of asynchronous electronic discussion for collaborative text processing 1.2.

1.1 Example A

In this first example[1], we will illustrate the limited information that acknowledgements give us about grounding at a knowledge level with an instance of human-to-human IT mediated communication, where two agents are coordinating for a meeting in an urban environment. The two peers exchange messages using an SMS (Short Messaging System, a system used to exchange short text content on mobile phones) chat, with the aim to guiding themselves in the actual space, towards the goal to reach a physical co-presence. Below, we will report the exchange transcript (Table 1) and the reference to the city map (Fig. 1).

[1] We recognise that this first example does not pertain to formal learning situations. However, we believe that informal learning as the coordination for a meeting can be assimilated to the mechanisms of knowledge construction related to conventional learning

Table 1: Transcript of the example conversation. In the third column we coded the transcript using the formalisation proposed by Traum [24]

Agent	Msg. #	Contrib./Act	Message Content	[Actual Action]	Map.#
A	1	initiate / initiate^{I}(1)	Can we meet at St. Francis church at 9?	[Standing in "x"]	1
B	2	ReqRepair^{R}(1)	Ok. Where is it? I am at the St. Paul 's station.	[Standing in "y"]	1
A	3	repair^{I}(1)	Go to the central plaza. Take left and the first right. Then the first left. See ya	[Standing in "x"]	1
B	4	ack^{R}(1)	Ok. I am on my way.	[Walking towards "z"]	2
B	5	ReqRepair^{R}(2)	I am lost. No way on left. I took right at the first junction but there were two streets. I took right again.	[Walking back towards "j"]	3
A	6	repair^{I}(2)ReqAck^{I}(2)	No, sorry. There you must stay on the main road. You should see me.	[Walking towards "k"]	3
B	7	ack^{R}(2)	Ok.	[Walking towards "k"]	4
A	8	initiate^{I}(3)	I am waiting at the red cross on the left hand side of the st.	[Standing in "k"]	4
B	9	ack^{R}(3)ReqAck^{R}(3)	Found the red cross office. Where are you?	[Standing in "w"]	5
A	10	cancel^{I}(3)initiate^{I}(4)	No, sorry it was another cross :-) Keep going for another two blocks.	[Standing in "k"]	5
B	11	ack^{R}(4)	Ok.	[Walking towards "k"]	5

Fig. 1: Map references used in the transcript at Table 1

If we try to model the described situation using Clark and Shaefer's Contribution Model$_{1989}$ [5], or Traum's Grounding Acts Model$_{1999}$ [24], we reach the conclusion that A and B have grounded their conversation at each acknowledgment. More precisely: once both presentation and acceptance phases have been completed, the peers will have grounded a certain contribution (at utterance n. 4, 7, 9, 11). There is the tendency in CSCL to correlate the rate of acknowledgments with the level of shared understanding on the assumption that the provision of evidences of reception is enough to infer the understanding of the signal and the corresponding incorporation in the contributor's beliefs. Additionally, when using these models, it is difficult to operationalise a lack of understanding, as in the example provided when B leads to point "z", because B provided clearly evidence of acceptance as per message 4, on Table 1.

Our claim is that in order to take into account the complexity of this kind of interaction we need to look at the situation from a knowledge construction point of view. From there, we argue, new descriptors of grounding are needed. Therefore, to stay with our example, we can say that the respondent B had an illusion of grounding between point "y" and point "j", until s/he realised that multiple solutions were possi-

tween point "y" and point "j", until s/he realised that multiple solutions were possible and s/he did not have enough information to solve the ambiguity.

1.2 Example B

The second example was collected in a study of using asynchronous discussion groups for the collaborative processing of academic texts [25]. This example serves mainly to show the fuzzy, ambiguous nature of communication when talking about abstract concepts, and the more nuanced levels of grounding that have to be distinguished. To fully illustrate this point, we would have to give an account of a whole discussion thread with 10 or 20 messages where students try to make sense of their subject matter. Then we would see a conversation that 'in form' would behave like the contribution theory describes, while it still would be very hard to determine how the different messages relate to one another on a knowledge level. Often, while all messages roughly concern the same topic and all consist of slightly different points of view, it will remain unclear whether anyone understood someone else's message, or whether any new knowledge was 'build' (see Fig. 2).

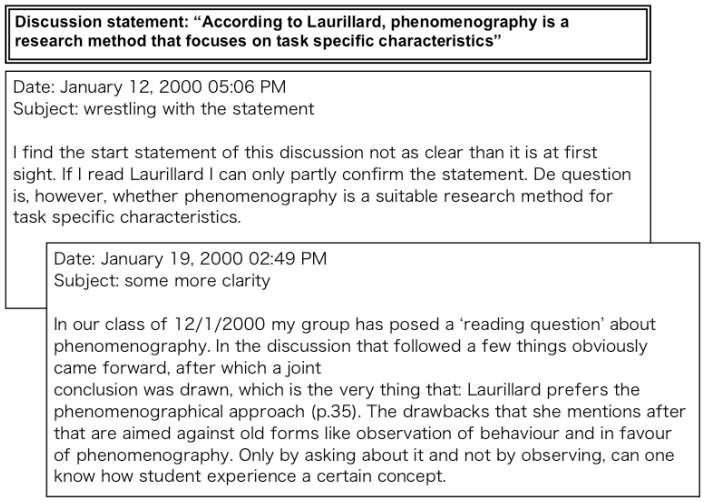

Fig. 2: Sample from an electronic discussion between university students. Grammatical errors were copied, to give a correct impression of the original ambiguities

The second message, which aims to provide 'some more clarity' on the abstract subject of Laurillard's account of 'phenomenography', can be considered as an attempt for grounding at a knowledge level. The most important aspect of this follow up message, is that it present a certain gradual 'topic drift': the question "what is phenomenography" is answered by a message about the question "how does Laurillard evaluate phenomenography". This topic drift results in the fact that existing views are

being exchanged, but because of a lack of relevance of one message to another, little new knowledge can be build.

2 Four Dimensions of Grounding at a Micro and Macro Levels

Using the presented examples, we will now elaborate on the difference between the micro and macro level, in four interrelated dimensions (see Fig. 3). Firstly, our examples show that the broad range of possible meanings on a knowledge level makes grounding more difficult, and is more likely to result in partial understanding then at a conversation level. Secondly, when it comes to measuring successful grounding, we propose to look at levels of commitment and co-referenced action, which might demonstrate (degrees of) shared knowledge better then acknowledgements. Thirdly, we will look at the underlying principles and see that because grounding is essentially efficiency-driven, the notion of 'effort' plays a central, but different, role at both levels. Finally, we will investigate where this effort is or should be directed and identify of perspective taking [13] as a primal grounding mechanisms on the knowledge level.

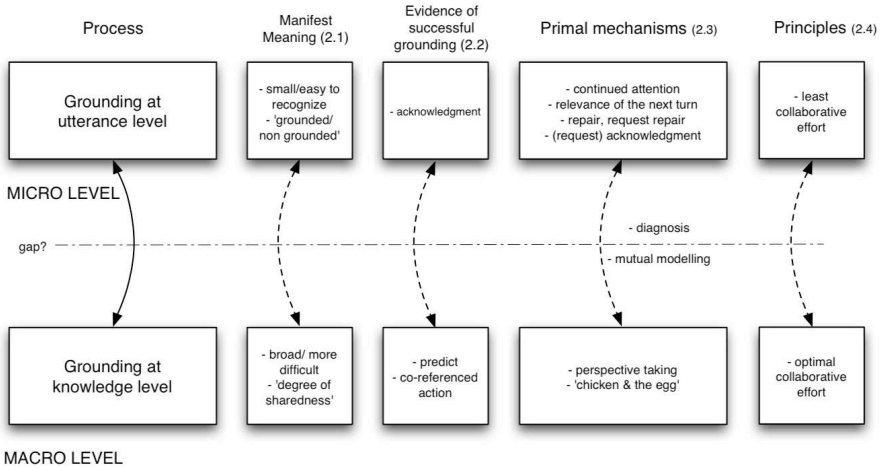

Fig. 3: A four-component model of grounding at utterance and knowledge levels

2.1 Manifest Meaning

Knowledge can never be accessed directly. As Laurillard [16] states, we have to *infer* conceptual information ('descriptions of the world') from the physical or communicative interactions we make in this world, thus making abstract learning, or communicating about knowledge, an essentially *mediated* phenomenon. Since this mediation is never perfect, and 'common ground' can never be reached completely ([11] referring

to Wittgenstein), we will use the notion of 'mutual cognitive environment' instead [26]. Sperber and Wilson define a cognitive environment as *the set of facts that are 'manifest' at a certain moment to a person: the facts that he or she is capable of representing and accepting as true or probably true*. Or, in our words, what is manifest for a certain person is the range of possible meaning that is evoked or triggered by the presented evidence, in a certain context. This collection of meanings that are associated to a certain action, concept or statement can even be so broad that it includes contradictory points of view [3]. The difference with Clarks description of common ground, is that to say two people share a cognitive environment does not imply they make the same assumptions; merely that they are *capable* of doing so.

While Clark's experiments started from the idea that a piece information `x` is either known or unknown to person `A` or `B`, the notion of 'manifestness' shows that there are also many degrees in between, and many different ways of 'knowing piece of information `x'`. We can say that the bigger the overlap is between the manifest meanings of different conversation partners, the more successful their grounding. When looking at the two levels we distinguished, we can state that the need for a notion of 'groundedness', which can account for subtle differences in interpretation, is even greater at a knowledge level than it is at an utterance level. Or, as Andriessen and Alargamot ([1], p. 8) put it: "*semantic understanding is something gradual*". Also, the smaller and more focused a range of manifest meanings is, the better the chances for successful grounding. This depends on what one is grounding: an intention or speech act, a literal meaning, a statement, or a certain point of view. The more elaborate and complex the grounding object, the more difficult grounding. Because the range of possible interpretations will usually be broader at the knowledge level than at the utterance level, grounding will also be more difficult at that level, and "a communicative intention can be fulfilled without the corresponding informative intention being fulfilled" [26]. We would like to see the distinction between the micro and the macro level not strictly as a dichotomy, but rather as a range, for instance going from recognizing simple intentions, to recognizing literal meanings, more elaborated statements and finally complex points of view.

2.2 Evidence of Successful Grounding

In concordance with Sperber and Wilson's account of the *evidence* that messages provide to guide their interpretation, the same can be said about analysing grounding. The more evidence we have, the more we know about the levels of shared understanding (though it may never be conclusive). As we have stated in the introduction, we do not think acknowledgements are always a valid measure of shared understanding. Apart from different goals at the two levels (see 2.4), Ross, Green and House [21] have shown that an (partial) 'illusion of shared knowledge' is not only possible, but also even likely to occur (called the *false consensus effect*). Therefore we propose to look at verbal and physical *actions* as well. Bereiter's term 'knowledgeability' [3], or 'being able to take intelligent action', indicates that (verbal or physical) actions intrin-

sically contain knowledge. In our first Example 1.1 the 'information bearing actions' one can identify are the coordination of tuning attempts with the agreed plan. If the pair agree to a certain strategy and then implement it coherently, we can infer that the pair successfully grounded to a high degree. Or, to state it more generally, if someone '*commits to a previous statement, and subsequently does something directly related to it in the forthcoming action or statement*' (we use this notion of commitment in accord with [8]). Since this relatedness between communicative actions requires a large overlap in the cognitive context and shared referents, we will label them as *co-referenced actions*.

In the asynchronous discussion groups (Example 1.2) we can look at the alignment of questions and answers. An answer that follows a question seems like a legitimate and useful speech act (utterance level), but only if the relevance of the content is established, we can deduce if it is also a successful knowledge-building act. On a knowledge level, for an action to be 'co-referenced', it is required that it refers to a shared piece of knowledge and needs to be relevant from someone else's view. According to Sperber & Wilson: "*something is relevant to an individual when it connects with background information he has available to yield conclusions that matter to him: say, by answering a question he had in mind, improving his knowledge on a certain topic, settling a doubt, confirming a suspicion, or correcting a mistaken impression*." ([27], p. 608). Our examples show that, while at an utterance level, both recognising a certain speech act, (such as identify a question by its question mark) and providing a relevant response (giving an answer) is pretty straightforward, on knowledge level the requirements for action to be relevant or co-referenced are much higher.

2.3 Grounding Mechanisms

At an utterance level, human communication is very efficient by investing minimal effort in elaborate message design or conscientious interpretation, but rather by jumping to (subjective) conclusions and repairing a possible misunderstanding *after* it arises. At a knowledge level however, we have seen that because of the mediated nature of grounding and the more complex collections of associated (manifest) meanings, this presents more problems. Miscommunication can be both harder to detect (thus cannot be relied upon to reveal itself) and to repair. Therefore, the grounding mechanisms at the knowledge level might present the most important shift from the utterance level. To understand what nuanced meaning other people attribute to certain statements, one must 'put oneself in the other's shoes' and try to identify which meaning will be relevant for that person [26]. In order to infer someone else's cognitive environment or 'frame of reference', both for reading and writing messages (audience design), we rely on strategies like *perspective taking* [13] and *mutual modeling* (for a definition see [19]).

While at an utterance level repair mechanisms are know to be self-regulating (the less shared understanding, the more grounding will take place, see for example [10]), this is less evident for knowledge level perspective taking. It seems that at this level, the 'chicken & the egg' relation between grounding and common ground ("*It is hard*

to find some if you don't have some already and you don't have any unless you find it", [15], p. 4) is even more prevalent than it is at the utterance level. This shows that at a macro level, knowledge of other's perspectives plays a role as a prerequisite as well as an outcome and the same goes for one's knowledge of the subject matter. Because identifying another's frame of reference is easier if one has knowledge of the different possible frames of reference that exist, perspective taking is also tied to existing knowledge. This underlines the reciprocal relationship between individual and collective processes in collaborative learning [23]: it is not only so that individual learning results from collaborative processes, but individual knowledge also influences the success of collaboration.

2.4 Grounding Principles

First of all, grounding is functional and driven by mechanisms of efficiency, as Clark & Wilkes-Gibbs [6] demonstrate with their 'principle of least collaborative effort' and Wilson and Sperber [26] in their 'relevance-theoretic comprehension procedure'. The fact that in grounding: no more effort will be invested that what is 'sufficient', can explain the lack of co-referenced actions in our examples. For students the costs (relative to the goals) may simply be too high, especially because high-level learning goals are usually translated into practical tasks, with which students deal in a pragmatic way. Taking the perspective of someone else may take more effort than staying within one's own perspective, and what is 'sufficient to continue the conversation' might not be 'sufficient for learning' [2]. That is why, for learning, instead of trying to 'minimize the collaborative effort', we strive for an: *optimal collaborative effort* [9].

The effect of effort into perspective taking and co-referenced actions is twofold: not only does relevant feedback enhance collaborative knowledge building, but the effort after shared meaning itself is also strongly associated with learning [22], especially if the effort is directed at the knowledge level (or 'semantic grounding', see [2]). Spending effort into trying to understand another perspective *is* learning: it is leaving one's preconceptions and trying to new information and insights. The is also true for reading, since *perspective taking* for comprehending messages is closely related to the comprehending process when studying scientific texts.

3 Conclusions

Context is inextricably present when we grasp meanings and when we infer knowledge. The pragmatic tradition of relevance highlights the action-oriented nature of intelligence, where the term 'action' is to be understood in a broad sense that includes reasoning behaviour, or communicative acts [7].

When looking at the relevance of communicative actions in collaborative learning, we have described that providing evidence and acting in a co-referenced way is crucial for developing a shared understanding. The more evidence is presented, the easier

it becomes to take another's perspective, act in a co-referenced way and enhance the degree of shared understanding.

We suggest the implications for design and research on grounding in collaborative learning might involve an effort to facilitate grounding at a knowledge level. For instance, communication tools could be developed that provide more (focused and detailed) contextual information which serves to limit the range of manifest meanings of the concepts that are being used, and thus to increase the chances of shared understanding. Also, since the use of acknowledgements as markers of shared understanding is problematic, we propose to create markers that can give an account of the relevance of communicative actions in regards to the reasoning process. As an example of this, we think 'operationalising' the 'co-referencedness' of actions on a knowledge level, as measure of shared understanding, would be a valuable effort. As a final remark, we want to conclude with a question for discussion:

- How can we 'operationalise' the concept of Perspective Taking?

The concept of co-referenced action, as defined in this paper, presents several limitations. It is sometimes difficult to have a multi-modality of communication that may make visible incongruences between intentions and actuations. We need other markers. Relevance theory states that an input is relevant when, and only when, its processing yields a positive cognitive effect [26], meaning that the receiver of the input will generate and act an expectation of a particular cognitive effect to be achieved by the incoming input [18]. These 'acts' of expectation may be observed and accounted as yielding back to the reasoning and inference process of the learner and on the particular perspective s/he is taking in relation to the processed input.

References

1. Alamargot, D. Andriessen, J.: The "power" of text production activity in collaborative modelling: nine recommendations to make a computer supported situation work. In M. Baker, P. Brna, K. Stenning, A. Tiberghien (eds): The Role of Communication in Learning to Model. Lawrence Erlbaum Associates. NJ (in press)
2. Baker, M., Hansen, T., Joiner, R., Traum, D.: Collaborative Learning: Cognitive and Computational Approaches, chapter The role of grounding in collaborative learning tasks, Pergamon / Elsevier Science, Amsterdam (1999) 31–63
3. Bereiter, C.: Education and mind in the Knowledge Age. Lawrence Erlbaum Associates, Inc. (2002)
4. Billig, M.: Ideological dillema's; a social psychology of everyday thinking. SAGE publications, London (1988)
5. Clark H. H., Shaeffer, E. F.: Contributing to discourse. Cognitive Science 13 (1989) 259-294
6. Clark, H.H., Wilkes Gibbs, D.: Referring as a Collaborative Process, Cognition, 22 (1986) 1-39
7. Ekbia, H. R., Maguitman, A. G.: Modeling and Using Context (Proceedings of the conference CONTEXT 2001, Dundee, UK, July), volume 2116 of Lecture Notes in Artificial Intelligence, chapter Context and Relevance: A Pragmatic Approach. Springer, Berlin, (2001) 156–169

8. DiEugenio, B., Jordan, P. W., Thomason, R. H., Moore, J. D.: The agreement process: an empirical investigation of human-human computer-mediated collaborative dialogues. International Journal of Human Computer Studies. (2000) 1–57
9. Dillenbourg, P., Traum, D., and Schneider, D.: Grounding in multi-modal task-oriented collaboration. In Proceedings of the European Conference on AI in Education, pages, Portugal (2000) 415–425
10. Dillenbourg, P. and Traum, D.: The complementarity of a whiteboard and chat in building a shared solution. Journal of the Learning Sciences, 26 (in press)
11. Draper, S. W., Anderson, A.: The significance of dialogue in learning and observing learning. Computers Educ., 17(1) (1991) 93-107
12. Gergle, D., Kraut, R. E., Fussel, S. R.: Language Efficiency and Visual Technology: Minimizing Collaborative Effort with Visual Information. Journal of Language and Social Psychology, 23(4) (2004) 491-517.
13. Järvelä, S., Häkkinen, P.: Web based cases in teaching and learning - reciprocal understanding and perspective taking in conversation. CAL99 (Computer Assisted Learning), London (1999)
14. Koschmann, T.: Reconsidering Common Ground: Examining Clark's Contribution Theory in the OR. ECSCW 2003: the Eighth European Conference on Computer-Supported Cooperative Work, Kluwer Academic Publishing (2003)
15. Krauss, R. M., Fussell, S. R.: Mutual knowledge and communicative effectiveness. In J. Galegher, R. E. Kraut, C. Egido (eds.): Intellectual teamwork: Social and technological foundations of cooperative work. Hillsdale, NJ: Erlbaum (1990)
16. Laurillard. D.: Rethinking university teaching. London, Routeledge (1993)
17. Mäkitalo, K., Häkkinen, P.: Building and maintaining common ground in web-based interaction. In Proceedings of the 9th Conference for Research on Learning and Instruction, Fribourg, Switzerland (2001)
18. Matsui, T.: Modeling and Using Context (Proceedings of the conference CONTEXT 2001, Dundee, UK, July), volume 2116 of Lecture Notes in Artificial Intelligence, chapter Experimental Pragmatics: Towards Testing Relevance-Based Predictions about Anaphoric Bridging Inferences, Springer, Berlin (2001) 248–258
19. Nova, N., Wehrle, T., Goslin, J., Bourquin, Y., Dillenbourg, P.: The impacts of awareness tools on mutual modelling in a collaborative video-game. In Favela, J. and Decouchant, D., editors, Proceedings of the 9th International Workshop on Groupware, Autrans, France (2003) 99–108
20. Pickering, M. J., Garrod S.: Toward a mechanistic psychology of dialogue. BBS (under review)
21. Ross, L., Greene, D., House, P.: The false consensus phenomenon: an attributional bias in self-perception and social perception processes. Journal of Experimental Social Psychology, 13 (1977) 279-301
22. Schwartz, D. L., & Lin, X. D.: Computers, productive agency, and the effort after shared meaning. Journal of Computing in Higher Education, 12 (2) (2000) 3-33
23. Stahl, G.: A model of collaborative knowledge-building. In Proceedings of the Fourth International Conferences of the Learning Sciences (ICLS), MI, USA. Ann Arbor (2000)
24. Traum, D. R.: Computational models of grounding in collaborative systems. In working notes of AAAI Fall Symposium on Psychological Models of Communication (1999) 124–131
25. Van der Pol, J.: Identifying and modelling variables in complex CSCL-situations. Case study: the use of asynchronous electronic discussions. CSCL2002 workshop Designing Computational Models of Collaborative Learning Interaction, Boulder, Colorado (2001)
26. Wilson, D., Sperber, D.: Relevance: communication & cognition. Blackwell publishing, Oxford (1995)

27. Wilson, D., Sperber, D.: Relevance Theory. In G. Ward and L. Horn (eds.) Handbook of Pragmatics. Oxford: Blackwell (2004) 607-632