Chapter to appear in P. Dillenbourg (Ed),*Collaborative Learning: Cognitive and Computational Approches*. Elsevier Science / Pergamon

Chapter 3

THE ROLE OF GROUNDING IN COLLABORATIVE LEARNING TASKS

Michael Baker
GRIC-COAST, CNRS & Université Lumière Lyon 2 (F)
mbaker@univ-lyon2.fr

Tia Hansen
Institute of Psychology, University of Århus (DK)
tia@psy.au.dk

Richard Joiner
Faculty of Education, The Open University (UK)
R.W.Joiner@open.ac.uk

David Traum
UMIACS, University of Maryland (USA)
traum@cs.umd.edu

ABSTRACT

Collaborative learning tasks involve interaction between multiple participants, who thus need to maintain some degree of mutual understanding. The process by which this is accomplished is termed *grounding*. The way in which collaboration, grounding and learning take place is largely determined by the task, the situation and the tools available. This paper discusses relations between grounding, collaboration and learning, drawing on research from two main areas: the Language Sciences and Cultural-Historical Activity Theory ("CHAT"). We build a unifying perspective of mutual understanding mediated by material and semiotic tools that can be used for analysis as well as for design of collaborative learning tasks, especially those that are carried out via computer-mediated communication. We illustrate the perspective with reference to a particular computer-mediated collaborative learning situation in the domain of physics.

1. Introduction

Collaborative learning is a complex phenomenon that can be analysed at many different levels. One of the most basic dichotomies is between individual and social levels of analysis. From one perspective, all action and learning can be seen as individual — it is individuals who act and learn — and notions of group learning and action must be derived from these basic building blocks. Yet, collective action must be seen as more than just a sum of isolated individual acts. Moreover, as described in numerous studies (see e.g. Plotzner *et al.*, *this volume*; Littleton *et al.*, *this volume*), collaborative learning can be both quantitatively and qualitatively different from individual learning.

From another perspective, society or culture is more basic than the individual, having a continuing historical presence, gradually evolving as a consequence of the accumulated thinking and action of the individuals that have comprised it. The main challenge here is to understand the specific nature of human learning that makes this accumulation of efforts possible. The answer suggested here is that human beings are capable of externalising (or "objectivising") thinking into tools (material tools, such as pens and computers, and semiotic tools, such as sign-systems). Such tools can then be appropriated by others (including new members of the culture, especially children), who can in turn further use and refine them. *Appropriation* — the special kind of learning that concerns such cultural objects — takes sign-mediated assistance from other members of the culture, who scaffold children's first attempts with the cultural object in such a way that they gradually move from being able to use tools under guidance to being able to use them on their own, and in their own way (Leontjev, 1981; core points also quoted in Tolman, 1988). In Vygotsky's words: "Every function in the child's cultural development appears twice: [...] first, *between* people (*interpsychological*), and then *inside* the child (*intrapsychological*)." (Vygotsky, 1978, p. 57).

Clearly both of these perspectives are crucial to a comprehensive account of collaborative learning: to explain both how individual action can fruitfully combine to yield collaboration and learning, and how the cultural tools of a society can be appropriated and used by individuals through interaction within the society. The main aim of this chapter is therefore to explore the extent to which a theoretical perspective on collaborative learning can be constructed, that unifies the two previously mentioned ones: "individual to society" and "society to individual". Given the ambitious nature of such an endeavour, we have chosen to focus on a single phenomenon that is central to the interplay between individuals and between individuals and society: the phenomenon of *grounding* (Clark & Wilkes-Gibbs, 1986).

Grounding is the name given to the interactive processes by which *common ground* (or *mutual understanding*) between individuals is constructed and maintained. Some mutual understanding between individuals will already exist at the start of any interaction, having been attained through integration of the individuals into a common culture. However, as an important part of the communication process, this common ground will need to be augmented with new information related to different facets of the activity, such as the individuals themselves, the tools, the goals and setting of the activity. Thus, during the interaction, as a result of grounding, learning may take place, in virtue of appropriation of refined tools. Our challenge is therefore to understand how these processes — grounding and appropriation — operating on quite different timescales, lead to collaborative learning. In order to respond to this challenge we need a deep understanding of the role that tools play in learning within cultures, together with micro-level analyses of how grounding actually takes place in the carrying out of concrete collaborative tasks. A unified perspective on the role of grounding in collaborative learning therefore needs to take into account both the roles of culture and of inter-individual interaction.

We draw on contributions from two main areas: *Language Sciences* research on grounding and related phenomena (e.g. Moeschler, 1985; Clark & Wilkes-Gibbs, 1986; Clark and Schaefer, 1989; Allwood *et al.*, 1991; Roulet, 1992) and *Cultural-Historical Activity Theory* ("CHAT") (e.g. Leontjev, 1981; Vygotsky, 1978, 1986; Cole & Engeström, 1993; Wertsch, 1994). Language Sciences research provides fine-grained cognitive models of the grounding process, collaboration, and how the two relate, within the short timescale of isolated verbal interactions. CHAT, on the other hand, provides us with the notion of *learning as appropriation of tools*, and enables us to understand the role of language (the 'tool of tools' according to Vygotsky) in collaborative learning. These two areas of research are thus complementary given that grounding is a process of language interaction. CHAT also enables us to situate the role of verbal interaction, and grounding processes in particular, within a cultural-historical timescale.

Whilst each area of research has a separate contribution to make to our project, some pre-existing theoretical connections between the two are readily apparent. For example, the terms "intersubjectivity" (e.g. Forman, 1992), "negotiation of meaning" (e.g. Lave, 1991) and the "interpsychological plane of functioning" (e.g. Wertsch, 1994), referred to in CHAT-inspired areas of psychology, appear to refer to phenomena that closely relate to grounding, as described in language sciences[1]. Nevertheless, although these terms appear to designate similar phenomenon, we should not neglect the fact that the theories that underlie them are radically different: a deeper theoretical comparison is required if we are to benefit from both of these areas of research in understanding the role of grounding in collaborative learning.

In this chapter, we have concretised our theoretical perspective in the form of a simple framework (or model) for analysing collaborative learning tasks that highlights the different types of mutual understanding that need to be achieved for learning from collaboration, as mediated by tools. The framework is intended to be useful for designing collaborative learning tasks, particularly those that are computer-based or mediated, since even the tools for interaction are, to some extent, under the designer’s control. Throughout the chapter, we use a specific computer-mediated collaborative learning environment ("C-CHENE" — c.f., Baker & Lund, 1997) in order to illustrate our theoretical perspective and to instantiate our analytical framework.

We begin by discussing the elements of our object of study (grounding, collaboration, learning) and the possible relations between them from the two different theoretical perspectives: Language Sciences and CHAT. Section 2 gives a brief overview of work in languages sciences on grounding. Section 3 then considers how learning takes place, from a more macro-level, involving appropriation of tools. Next we consider how to combine these two areas of work to achieve a perspective on the role of grounding in collaborative learning (section 4). We then present the analytical framework, which draws on each of these areas (in Section 5) and describe an instantiation of it with respect to the C-CHENE computer-supported collaborative learning environment, in section 6. We conclude by synthesising and developing our main claims, and with suggestions for future research on the role of grounding in computer-supported collaborative learning.

2. Grounding

A *common ground* of mutual understanding, knowledge, beliefs, assumptions, presuppositions, and so on, has been claimed to be necessary for many aspects of communication and collaboration. *Grounding* is the process by which agents augment and maintain such a common ground. Although agents who interact will usually already possess some such common ground, perhaps in virtue of their common membership of a particular culture or social group, their physical co-presence or even due to their previous interactions, this common ground will also need to be augmented and maintained during the interaction itself, in order to take into account new aspects of the common situation or task.

[1] For other recent attempts to integrate discourse analyses and CHAT, see, e.g, Wells (1996).

2.1 Basic principles

Clark and Marshall (1981) pointed out that mere accessibility or presentation of information is not sufficient for it to be added to the common ground. Depending on the situation, other assumptions may be required, such as community co-membership, rationality, locatability, attention, and rationality. Clark and Schaefer (1989) went beyond this, claiming that feedback of some sort was needed to actually ground material in conversation, and that this grounding process was collaborative, requiring effort by both partners to achieve common ground. They point out that it is not necessary to fully ground every aspect of the interaction, merely that the participants attain the following "*grounding criterion*":

> "The contributor and the partners mutually believe that the partners have understood what the contributor meant to a criterion sufficient for the current purpose." (Clark & Schaefer, 1989, p. 262)

What actually counts as a "sufficient" criterion of mutual understanding depends of course on the reasons for needing this information in the common ground, and can vary with the type of information and the collaborators' local and overall goals. As we discuss later in the chapter, the criterion may need to be particularly stringent if the collaborators' activity is to lead to *learning*.

The grounding process involves, in addition to the mentioning of facts and proposals in the presence of another, processes of diagnosis (to monitor the state of the other collaborator) and feedback. When things are going smoothly, this feedback can be just simple acknowledgements (or even implicit, by performing an action which is contingent on understanding). However, feedback can also take the form of *repairs* when understanding seems to deviate from commonality. Collaborators do not always explicitly attempt to check that everything is mutually understood. Rather, they often make *assumptions* as to what the common ground is and will become, on the basis that this will be more economical. Clark and Wilkes-Gibbs (1986) introduced the principle of *least collaborative effort*, claiming that: "In conversation the participants try to minimise their collaborative effort — the work that both do from the initiation of each contribution to its mutual acceptance." Thus a common form of feedback is a simple signal that the addressee is listening and comprehending rather than a display of precisely how the addressee is understanding the speaker. When it becomes manifest that an assumption of shared understanding has turned out to be unwarranted, there will usually be some attempt to repair and firmly establish the common ground — though only where this is deemed to be worth the effort involved.

Many aspects of interaction may be grounded, not just the meanings of utterances. Allwood *et al.* (1991) describe four basic communicative functions, which correspond to 'levels' at which problems for maintaining common ground may arise (c.f. also Clark, 1994):

- *contact* (whether the interlocutor is willing and able to continue the interaction);
- *perception* (whether the interlocutor is willing and able to perceive the message);
- *understanding* (whether the interlocutor is willing and able to understand the message);
- *attitudinal reaction* (whether the interlocutor is willing and able to react and adequately respond to the message, specifically whether s/he accepts or rejects it).

Clearly, there is an ordering between functions: one can not genuinely establish an attitudinal reaction unless the message is understood, which requires perception and, in turn, contact. In fact Clark and Schaefer (1987) show that explicitly grounding on one of these level reveals that grounding at a higher level is not achieved. The process of proposal, diagnosis and feedback which are central to grounding can also be viewed as a kind of *negotiation*, where the different levels represent different objects of negotiation (Moeschler, 1985; Roulet, 1992; Baker, 1994). In fact, nearly all aspects of management of the interaction (see Bunt, 1995) may be negotiable (e.g. turn-taking, opening and closing the dialogue, focusing, time management).

Whether researchers speak of grounding, negotiation or intersubjectivity (see §1 above), there is nevertheless common agreement that various forms of linguistic and non-linguistic feedback constitute the basic mechanisms by which the common ground is achieved and maintained. However, the precise form that these mechanisms take also depends on the *media* of communication concerned (Clark & Brennan, 1991) — a point that we return in the subsequent discussion of the role of tools in learning (§ 4.1).

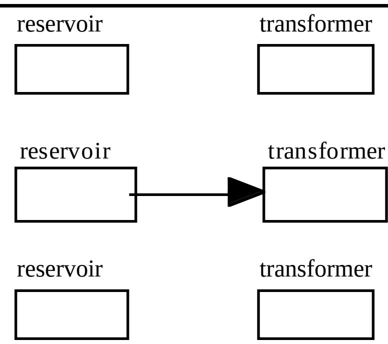
Different media (e.g. face-to-face, telephone, video-teleconference, terminal teleconference, and email) bring different *resources* to, and *constraints* on grounding. For example, to what extent does the medium support *copresence* (can see the same things), *cotemporality* (messages received at the same time as sent), *simultaneity* (can both parties send messages at the same time or do they have to take turns ?) or *sequentiality* (can the turns get out of sequence)? These constraints lead to different associated 'costs', in terms of the effort required of interactional participants. For example, how difficult is it to change speaker, to repair misunderstanding, to tell what is being responded to or to formulate what one wants to say?

Since different media have different combinations of these constraints and costs, one would therefore expect the principle of least collaborative effort (*re* the grounding criterion) to predict different styles of grounding for use in different media. Designers of Computer-Mediated Communication systems, used for collaborative learning tasks, need to pay particular attention to these factors (see also Hansen *et al.*, *this volume*).

2.2 An example of grounding in computer-mediated communication

The short interaction sequence reproduced in Table 1 below gives a concrete illustration of some of the points made above. It is taken from a corpus of computer-mediated interactions produced by two students using the C-CHENE system (Baker & Lund, 1997).

Table 1. Extract from computer-mediated interaction using C-CHENE[†]

N	T (s)	S	Dialogue	Trace of successive interface states /actions
1	208	S2		 *(link from transformer to reservoir created then deleted)*
2	451	S1	I’ll do the transfer	
3	482	S2	what about me	
4	527	S1	you the second reservoir	
5	564	S2	OK, go ahead	
			<…>	

† N = line number ; T(s) = time in seconds ; Dialogue = typewritten synchronous messages (full screen-sharing). Only the dialogue history and the current state of the graphical interface is visible to both students.

Extended sequences from the corpus will be analysed in more detail later in the chapter (Table 3). For our present purposes all that is necessary for understanding the example is that the students are attempting to draw a diagram (called an "energy chain") together on a shared computer screen, to represent energy in a simple experimental situation (a battery linked to a bulb by two wires). The diagram consists of reservoirs and transformers of energy, with transfers of energy that link them together.

Firstly, with respect to grounding, the question arises as to how S2 could understand what S1 meant by "the transfer" (line 2)? The answer is: because S1 and S2 assume that they have a common ground. This common ground consists of a common memory for the problem-solving actions that have already been performed, which is partly externalised on the computer screen, and must be partly remembered (they assume common knowledge of the fact that S2 just drew an arrow from the transformer to the reservoir, then deleted it). However, the students must also assume common understanding that the arrow drawn then deleted was intended to represent 'a transfer'. Furthermore, S2 needs these common assumptions in order to understand that S1 is talking about a transfer between the reservoir and the transformer. Just as S2 needs to access these assumptions in order to understand what S1 meant, S1 has in fact *formulated* and *presented the contribution* of line 2 on the basis of the same assumptions. S1 can thus produce his utterance with greater economy (*production costs* with this typewritten medium are high). Note that although these assumptions appear minimally sufficient for S1 to convey an understanding to S2 of what S1 is intending to do (draw an arrow from the reservoir to the transformer, to represent a transfer), and thus for the interaction to continue, this does not guarantee individual or mutual understanding in a 'deeper' sense of what "transfer (of energy)" really is, in terms of the physics task (a point with which students often have difficulties).

Secondly, how does S1 know whether, or to what extent, S2 has understood the utterance of line 2, i.e. the extent to which the utterance is *grounded*? The case is quite complex, since no explicit feedback (e.g. "right", "OK, …") is given. At this point in the interaction the students are discussing 'who will do what'. S2 therefore demonstrates understanding of S1's utterance (line 2) in line 3 *by continuing relevantly* with respect to this common goal, i.e. S2's asking what s/he will do, as part of a negotiation of tasks (S2 will accept S1’s claimed task, as long as S2 has an acceptable assignment).

A similar analysis can be given for lines 4 and 5. What is significant about line 5 is that S2 explicitly grounds the proposed division of responsibilities ("OK, …") on the level of *agreement* thereby grounding *understanding* as well (agreement with x presupposes understanding of x). At the end of this extract, the proposals for who will do what, with what, and when, can be assumed to be grounded for the participants in the interaction to a sufficient degree.

From this example we can thus see: (i) some of the range of possible objects that can be grounded (referents such as "the transformer", meanings of words such as "transformer", responsibilities for and co-ordination of problem-solving), (ii) communicative levels at which grounding needs to take place (understanding, agreement), and (iii) the role of the medium in the way that utterances are produced and grounded.

2.3 Computational work on grounding

Grounding is becoming increasingly important in computer dialogue systems. This is particularly true for systems with spoken language interfaces because current speech recognition technology produces far more errors than a human listener would make. In order for the system to correctly interpret communications by the user, the system must sometimes try to verify or attempt to repair its first hypothesis. As with human-human communications, there are a number of possible grounding strategies, ranging from explicit requests for verification to demonstrating the current understanding (perhaps with a visual display), to implicitly revealing it by performing a contextually relevant action. Notions of the grounding criterion and principle of least collaborative effort apply for human-computer interaction, as well: often simple feedback from the user can be more efficient and accurate than computation. On the other hand, frequent requests for verification can be very annoying to the user, and slow down the interaction. Much current work (e.g., Smith & Hipp, 1994; Danielli & Gerbino, 1995) is devoted to investigations of the best styles, i.e. what kind and when to perform verifications.

Traum (1994) presented a computational model of grounding, extending previous work by Clark and Schaefer (1989). In this model, utterances are seen as the performance of particular kinds of speech acts (such as initiate, continue, repair, and acknowledge) which change the state of groundedness of some information. The model allows one to form a precise theory (which can still turn out to be incorrect) of what is grounded and what actions need to be performed to achieve grounding at any point in the conversation. Moreover the model included a relation of these acts to mental states of the agents in such a way that it could be implemented and used as a component in a human computer dialogue system (Allen *et al.*, 1994).

As well as providing the resource for a computational agent to engage in more flexible and natural grounding behaviour, such a model could also be used within a computer-mediated communication system, to help make explicit to the users how their grounding of content is proceeding. Dillenbourg *et al.* (1997) describe preliminary implementations of "observer" agents which provide visual views of several measures of collaboration, including degree of grounding.

Computational work on grounding may thus be important in educational settings, for the design of computer-mediated communication in distance learning and for design of computer dialogue systems in computer-based learning systems.

2.4 The relations between grounding, collaboration and collaborative learning

Having described grounding, we now discuss how it relates to collaboration and to collaborative learning. We argue that, under specific operational definitions of the terms "collaboration" and "collaborative learning", where the latter is interpreted as the learning that occurs in virtue of collaborative activity *per se*, grounding is an integral part of both.

As illustrated in the introductory chapter of this book (Dillenbourg, *this volume*), there are many notions of collaborative learning, and collaboration itself. Whilst various proposed definitions differ widely as to the specific necessary and sufficient ingredients of collaboration, most do converge on some relation to co-ordination and mutuality of the mental states of the collaborators. We will therefore use one recent and widely accepted definition of collaborative activity, proposed by Roschelle and Teasley (1995):

> "[collaboration is] a coordinated, synchronous activity that is the result of a continued attempt to construct and maintain a shared conception of a problem." (Roschelle & Teasley, 1995, p. 70).

Roschelle & Teasley's (1995) definition of collaboration, relying as it does on the attempt to construct and maintain a "shared conception", appears to be almost identical with the definitions of grounding and the common ground presented here. If this were so, then we could conclude that students are collaborating, quite simply, when they are engaged in grounding. However, we would claim that although grounding is a form of collaboration, and collaboration requires grounding, the two terms are not co-extensive. This is because grounding and collaboration can take place with respect to different objects and different communicative functions.

We claimed above (§2.1) that grounding can take place with respect to four communicative functions — contact, perception, understanding and agreement — and with respect to different 'objects' — meanings, propositions, rights, obligations, images, etc. When grounding and collaboration take place with respect to a specific level/object, then the two co-incide. However, grounding and interaction may take place with respect to different levels/objects *simultaneously* in a given situation. In this case, there can be both collaboration and non-collaboration, but with respect to different levels/objects. Grounding and collaboration do not, therefore, necessarily coincide, although they may do so in certain cases.

Consider the case of two students collaborating, as in the example from Table 1. They will already possess a certain common ground as members of the same class, having worked together before, and so on. At one point a disagreement over an intermediary solution to the problem could arise, be mutually recognised, leading to an argumentative discussion. The students may be grounding and thus collaborating to the extent that each is trying to understand the other, and to address the different points raised for or against the proposed solution. However, they may be competing rather than collaborating with respect to the communicative function of agreement on domain goals and propositions. More subtly, the domain-level disagreement may be accompanied by another disagreement, working on the level of the self-images that they attempt to impose in the interaction. For example, one may attempt to impose the image of being "more knowledgeable", or "more reasonable" than the other. In this example, there is thus co-existence of collaboration and its absence, but with respect to different levels/objects.

In summary, once we adopt the restricted notion of collaboration proposed by Roschelle and Teasley (op. cit.), this being a definition that has become widely accepted, then students are collaborating when they are attempting to ground in some way, at some level. If we then adopt the premise that a useful concept of collaborative learning is the learning that occurs in virtue of the collaboration *per se*, then learning-from-grounding is "collaborative learning" in this sense[2].

3. Learning as appropriation of tools

Grounding, as described previously, is an important aspect of collaborating, and thus collaborative learning. Another crucial component to investigating the role of grounding in collaborative learning is, of course, learning itself. We now extend our theoretical perspective on collaborative learning to incorporate elements of CHAT, based on the work of Vygotsky (1978, 1986), Leontjev (1981) and Luria (1979). This enables us to take into account the role of tools and to extend the timescale of activity considered. Fundamentally, CHAT views learning and development as a sociocultural activity where the child appropriates culturally accumulated knowledge with the assistance of other people and tools available in the setting.

3.1 Tool-mediated activity

Within CHAT the basic structure of human cognition is characterised as a triangle with the subject (s) and object (o) on two sides of its base and the medium or artefacts at its

---

[2] One may thus wonder whether a student learning with the help of a more advanced tutor would be considered 'collaborative learning'. Clearly the two can be pursuing the same goal — that the student learns — and will engage in grounding behaviour. While this might be termed a collaboration, we would not consider it to be collaborative learning.

apex (m/a) (see Figure 1). In this model, the subject refers to an active, conscious agent, and the object represents the object of his/her cognition and action, which is the thing or phenomenon that s/he is thinking about and acting upon.

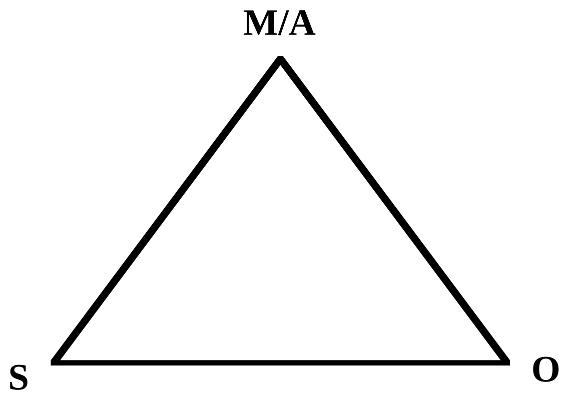


Figure 1. The mediational triangle (subject object and medium/artefact)

Natural or unmediated action is represented by the line between S and O, and refers to those capacities that the subject's own organism supplies, i.e., what the senses reveal, and the instincts that the subject has about what to do with an object of this kind. For example, a small child faced with a pot of boiling water senses an object with certain physical qualities (spatial qualities, colour, humidity, heat), follows his/her instinct to explore unknown objects and experiences new sensations of heat and pain. This is unmediated action. By unmediated action and cognition the child can learn that this object provides unpleasant sensations when touched. S/he may also be able to learn to distinguish between objects that have this quality and objects that do not. However, it takes mediated action to learn to understand why. The latter is a theoretical issue, which presupposes ability to go beyond the immediacy of sensory experience.

Mediated action is depicted as the interaction of the subject with the object, mediated by auxiliary means (this is indicated by the lines s-m/a and m/a-o). Tool mediated action refers not only to material tools (e.g., thermometers, computers, etc.), but also to semiotic ones (concepts, signs, and models). Language opens up a whole range of possibilities. Without language people could only deal with those things they could perceive and manipulate directly. With language they can deal with things that cannot be perceived directly and with products of past generations' experience.

When a subject's actions are mediated, the task itself is fundamentally changed, relative to natural or unmediated action, and so is the cognition that goes with it. The task becomes a more complex structure of actions and cognitive processing since subjects must incorporate the three elements of the triangle and the possible relations between them into a unified idea, in order to plan their actions and accomplish the task.

An example of action mediated by material tools would be when a child uses a thermometer to measure the temperature of a liquid. But even before s/he has learnt how to do that, s/he already has learnt to employ semiotic mediation in her/his quest for understanding. For example s/he has begun to understand what is meant by 'boiling kettle'. S/he has learnt words such as 'boiling' and 'temperature' and has a rudimentary understanding of what those concepts mean. Similarly, the example discussed in section 2 above shows the use of two types of semiotic tools — the energy chain graphical 'language', and of course typewritten language — that are embedded in specific material tools (computer interfaces).

Children are said to *appropriate* cultural objects (material and semiotic tools), when they learn from other members of the culture how those cultural objects are used, and what they are used to accomplish. Appropriation is not a process of rote-learning, in which the individual simply adopts the facts and assumptions of the culture. Children appropriate these objects by participating in their use with more expert members of that culture. Learning results from the child's own experiences and practice with the object under the guidance of an expert. Thus, it is not a matter of information transmission from the expert to the novice, but of the novice "making this tool his own" (Leontjev, 1981).

3.2 Zone of proximal development

Within the viewpoint of CHAT, the appropriation of semiotic tools must occur within the "zone of proximal development" ("ZPD") (Vygotsky, 1978,1986). This is defined as "the difference between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" (Vygotsky, 1978, p. 86). Children interacting in the zone of proximal development are participating in more advanced tool mediated activity than they are capable of independently, and in doing so they appropriate these cognitive tools, which advances their own independent problem solving.

Within the ZPD the novice will not perceive the meaning of the object as fully as the expert user, and the novice and expert may have vastly different interpretations. An object such as a diagram or a representation on a computer interface may be understood very differently by the participants. Likewise, an utterance may be understood quite differently. This is not necessarily a problem (in reality, all utterances are always understood slightly different by different interpreters, c.f. Rommetveit, 1979, 1985). In some instances it is not essential that the participants strive for sharedness of interpretations, as long as they agree enough to carry on the interaction (e.g., if they can point to the same place on the screen, they do not need the same interpretation of what is there for their attention to stay focused on that bit of the interface). This ability to talk about objects which do not have a completely shared meaning is not a problem, but in fact an advantage. It allows for learning to occur when people with different perspectives interact. In order to untangle differences in understanding, participants are forced to make their assumptions explicit, to argue, reason about and exemplify them. In order to reach agreement they need to construct a conception which allows for apparently insightful assumptions from both parties to be integrated, which will often require a conceptualisation that is more complex than the individuals' original ones.

3.3 Intersubjectivity

Thus an important process in the social origins of learning is joint construction of mutual understanding or intersubjectivity (Forman 1992; Rommetveit, 1979, 1985). Wertsch (1985) identifies four levels of intersubjectivity in the transition from inter-psychological functioning to intra-psychological functioning in an interaction between a mother and a child. The task was to copy a puzzle (a truck) in accordance with a model. Wertsch characterised the core goal directed action in the adults situation definition as a series of three steps:

i) Consult the model to determine the identity and location of the next piece;
ii) Select the piece identified in step 1 from the pile of pieces;
iii) Add the piece selected in step 2 to the copy.

The first level of intersubjectivity is characterised by the fact that the child's situation-definition is so different from the adult's that this makes combination very difficult. The adult may try to direct the child, but the child's understanding of the setting is so limited that the child does not interpret the adult's utterances appropriately. For example, one child shifted his attention to the real windows in the room in response to a mother's comment to look at the truck window.

At the second level, the child has a limited understanding of the setting. Sharing the adult's basic understanding of the objects in the setting (i.e. they depict pieces of a truck). However the child does not fully understand the relationship that has been constructed between the truck and the model, and will often fail to respond appropriately to adult directions. For example in response to one mother asking the child "to put the wheel on next" the child picked up the wheel and placed it in the wrong place, not taking in account of where the wheel was on the model.

At the third level, the child can respond to the adult's directions. (S)he is making the appropriate inferences to the adult's statements, which indicates that inter-psychological functioning is beginning to account for much of the child's performance. The adult does not have to specify all the steps, because the child has a fairly complete situation definition. At the fourth level the child takes over complete responsibility for carrying out the task.

Wertsch (1985) characterises learning as the transition from inter- to intra-psychological functioning. This process involves negotiating an intersubjective situation definition with the tutor. Initially this will be performed on the basis of a primitive intersubjectivity with the tutor and subsequently by going through a number of situation re-definitions until finally the child acquires a mature culturally appropriate situation-definition which provides the basis for self regulation. Children working with adults can interact at a more advanced level of intersubjective understanding than they would be capable of independently.

In summary, from the vantage point of CHAT, learning can be seen as the appropriation of cultural tools. Grounding (on some level, with respect to some object) is thus necessary but clearly not sufficient for collaborative learning (as we define it). Firstly, if students are to learn from grounding, the common ground, or 'background' of material and semiotic tools provided by culture must be already appropriated by students to some degree. This is precisely what is usually left out by approaches that focus exclusively on the internal dynamics of interpersonal interactions. Secondly, collaboration provides a vehicle for learning if the individual participates in solving a task that is beyond his or her present autonomous ability. However, there are constraints on how far collaboration can promote cognitive abilities at any given point: grounding, or 'intersubjectivity' must be situated within the ZPD if it is to lead to learning.

4. Towards synthesis: grounding and collaborative learning as the appropriation of tools

In the previous two sections we presented some basic material necessary for understanding grounding as an interactive process and the CHAT theory of learning as the appropriation of tools. We now discuss some points of convergence and divergence between these two areas of research, with a view to understanding how grounding could lead to collaborative learning.

4.1 Language as a tool and an object of grounding in communication

The most obvious convergence between the grounding model and the notion of learning as appropriation of tools obtains via the notions of *language* and *communication*. Grounding is a process whose primary function is to ensure effective communication in the sense of shared understanding of utterances in interaction. Although non-verbal signals (e.g. gestures) and situational factors (such as co-presence) play an important role in this process, the primary objects that are themselves to be grounded are in fact linguistic productions (utterances). The extent to which interactions are linguistically or else non-verbally based is a function of how much common ground is already established: when participants do not know each other very well, they will need to speak; when they have interacted together for fifty years, a slight raising of an eyebrow may be sufficient for communication. So, we can conclude that grounding in the collaborative learning situations that we consider here is primarily language-based.

From the point of view of the Activity Theorist, language is a (semiotic) tool, a mediator of activity and thinking. During their development, children learn to use language (in speech) as a tool in different ways: to enlist help from others by communicating with them, to shape their ongoing activity, to plan their future activity, to structure their reasoning, and so on (Vygotsky, 1978). They may also refine the meanings of words in order to use them as more elaborated tools in new situations — for example, they may learn to elaborate the meaning of the word "heat" from its use in everyday life to its use in the science classroom. Language is thus a multiform tool: it enables interpersonal communication, it is instrumental in individual and collective activity and thinking.

How does this theory of language as a tool transform our way of seeing learners who are engaged in grounded interaction ? Three main points of convergence can be identified.

Firstly, the instrumental view of language as a tool (for guiding action, exchanging information, planning, and so on) leads us to concentrate on the interactional context, the *discourse genre*, in which grounding takes place and to ask the following question: *why* are the learners grounding, what are their main *goals* ? This gives us a way of understanding the grounding criterion mentioned above (§2.1), where it is stated that participants attempt to attain mutual understanding to a degree that is deemed sufficient *for the current purpose*. An understanding of learners' purposes (goals, motivations) is therefore crucial to understanding their grounding activity. In the next section, we discuss how these goals may be local to the problem-solving task, but also largely determined by the subjects' interpretations of the *situation* in which they find themselves.

Secondly, language — a tool — may itself be an *object* of mutual understanding. For example, learners may be led to attempt to understand a language itself (e.g. foreign language learning) or to inquire about the 'deeper' meaning of a particular term in their common language. This view thus leads us to ask the question, in addition to the above: *what* are the students grounding ? The answer from activity theory in this case (language) is: semiotic tools. These include symbolic, iconic and natural languages. The type of learning especially associated with grounding will therefore be appropriation of the very medium with which grounding takes place. We discuss such a formal language, whose appropriation constitutes learning in a specific domain of physics, in section 6 of this chapter.

Thirdly, communication in language may sometimes involve the use of material (communication) tools. In spoken face-to-face interactions this is less evident, but is more so in the case of computer-mediated communication. In terms of the grounding model, such tools are considered as the communication "channel". However, such a metaphor leads could lead one to think of the channel as a passive medium, however much "noise" is introduced. There is thus an important convergence between the notion of material tools, constituting the physical communication medium, as *mediators* that shape communication and the results concerning the role of media in grounding mentioned earlier (§2.1). Thus communication "costs" and "constraints" within the grounding model can be understood as difficulties in appropriating material tools, that influence the way in which semiotic tools are used and appropriated in grounding.

Viewing learners engaged in grounding as appropriating tools leads us to focus on understanding learners' goals within specific types of discourse, on what is being grounded and on how the material tools used for communication itself shape the grounding process.

4.2 Grounding, intersubjectivity and learning as appropriation of tools

What is the type of learning associated with grounding — at least from a theoretical point of view — and what are the necessary conditions for its occurrence? We mentioned above that CHAT leads us to pose the important question of learners' goals underlying their grounding activity, and the focus of it. To deal with the second question first — what learners are attempting to ground — from the point of view of the grounding model (and see, e.g. Schwartz, 1995) we make the following conjecture: *collaborative learning will be associated with a relatively high grounding criterion, with a high degree of cognitive-interactional effort.*

However, for learning to take place in a particular knowledge domain, this effort needs to be directed towards the right object. In simple terms, we can say that grounding can take place on pragmatic and semantic levels[3]. At the pragmatic level, grounding involves, for example, interactional participants understanding each others' communicative intentions, what the other is 'trying to tell them'. At the semantic level, it involves searching for common understanding of referents and meanings of terms. Pragmatic level grounding is part of *learning to collaborate*: one learns to generally understand what the other will be trying to tell us. Semantic level grounding however, relates to attaining mutual understanding of what is meant by certain terms and expressions; it thus relates more closely to learning in a specific knowledge domain by means of interpersonal interaction. Learning to understand one another and learning to understand semiotic tools are thus two important aspects of collaborative learning as we define it. It seems plausible that students will need to have learned how to collaborate in order to learn together in a task domain. We can thus view learning by grounding as a gradual transition from pragmatic to semantic grounding.

Finally, from our above discussion, it should be clear that what has been termed semantic grounding can also be viewed as appropriation of a semiotic tool. Our previously mentioned conjecture can thus be reformulated as follows: *collaborative learning will be associated with a gradual transition from the use of language as a medium for grounding communication (pragmatic) to grounding on the level of the medium itself (semantic), leading to appropriation of the medium.*

Why should students expend extra effort in semantic grounding, rather than contenting themselves with a level of pragmatic grounding that enables the interaction to continue in a minimal sense ? In this theoretical discussion, we can only point to the necessity for detailed analysis of learners' goals, their motivations, as mentioned above. Some tasks may be intrinsically motivating for some students; with other students, extrinsic situational goals may come to the forefront. It also appears plausible that learners' must have some understanding of what they are trying to do, or 'supposed' to do in order to be willing to expend effort on trying to understand the task: what researcher or teacher has not had the experience of discovering subsequently that the subjects/students had no clear understanding of what they were doing or had been asked to do ? A more concrete question is: what are the conditions that would favour such grounding activity ? Some answers may be found in the research on intersubjectivity, referred to earlier (§3.3). On one hand, semantic (or 'deep') grounding can not lead to learning if the concepts dealt with are simply beyond the students' zone of proximal development. On the other, it seems that an optimal level of difference between learners' conceptual viewpoints can stimulate the desire to gain deeper understanding.

4.3 Recapitulation

Grounding is a process that is directed primarily towards attaining mutual understanding of linguistic productions (utterances). Language is thus the primary

[3] Note that the semantics/pragmatics distinction is a subject of much controversy in language sciences — see e.g. Allwood (1976) and Levinson (1983) for reviews of these debates.

medium of grounding and (in virtue of its role as a semiotic tool) may itself be the object of attempted mutual understanding. Learning from grounding — collaborative learning — can thus be viewed as appropriation of semiotic tools, mediated by those very tools. Interactants' goals in specific situations determine the extent to which they will be willing to expend effort in achieving mutual understanding. The learning associated with grounding will be associated with a high degree of expended cognitive-interactional effort to gain mutual understanding. Appropriation of cognitive tools can be viewed as a gradual transition from pragmatic level grounding — learning to understand each other, to collaborate — towards learning to understand the semiotic tools in a specific domain (that may also be languages themselves). Apart from intrinsic and extrinsic (situational) motivations, learning from grounding can only occur within the zone of proximal development and may be stimulated by an optimal level of differences between individual perspectives. Gaining deep conceptual understanding takes an amount of time that extends well beyond interactions themselves and which draws on interactants' understandings of tools inherited from culture.

We conclude that *understanding the role of grounding in collaborative learning requires a unit of cognitive analysis that includes agents, tools and goals in situation, together with relations of understanding between them.* In the next section we present an analytical framework that encapsulates this synthesis .

5. A framework for analysing the role of grounding in collaborative learning tasks

We present a simple framework for analysing the role of grounding in collaborative learning tasks. The framework highlights the different forms of mutual understanding that will need to be achieved and gives a central place to tools as both mediators of understanding and objects of it. It may thus be used as an aid to designers of collaborative learning tasks, and in particular those that are performed via computer-mediated communication. In this case, designers have to actually decide how to engineer communicational and interactional possibilities for potential learners, and so must take into account the costs and opportunities that these channels provide for achieving different types of mutual understanding.

According to our framework, a collaborative learning task is defined by the pursuit of certain goals by a group of agents via the use of specific tools in a specific situation. This goal-oriented activity requires achieving different types of understanding and mutuality.

The basic elements of the analytical framework are shown below in Figure 2.

tools themselves (material tools such as communicative media for producing messages and manipulating the environment, and semiotic tools such as language), and *mediators* of agents' understanding of each other, the task and the goals-situation (the curved lines that pass through the circle around τ). The agents' goals (γ) and the situation (Σ) in which they are pursued have been represented as a unique yet complex entity since the two may be viewed as inextricably linked (see the discussion below).

The symbols for the entities in the diagram of Figure 2 should be viewed as variables whose specific instantiations can give rise to representations of a large variety of collaborative learning tasks. *Learning* is not represented explicitly in the diagram, since neither is *time*. *Learning may be achieved when the entities and relations change throughout time.*

5.1 Agents (α_1, α_2)

These are the distinct entities that collaborate. At a very general level they could be quite heterogeneous kinds of entities, such as individual people, groups, or even corporate entities. Some computer programs can also be viewed as agents (although this view is not without its critics[5]). The basic criterion is that each agent can meaningfully be conceived as having its own mental state and ability to act and interact with the other entities represented in the model. Each agent should be able to communicate with the other, and represent the mental state of the other (at least enough to engage in a grounding process). In addition, agents should be able to conceptualise and act within the situation towards achievement of the goals and use tools. Agents should also be able to learn about these aspects and be held responsible (to the other agent) for specific actions, including communicative ones.

Even with respect to prototypical agents — single human beings — there is a large variety of types of agents, particularly when one considers their relationship to each other. Agents could be adults or children, could have equal or unequal abilities, could be novices or accomplished at the tasks and using the tools, could be familiar or unfamiliar to each other, in an equal or hierarchical social relation to each other, and so on.

5.2 Goals and Situation (γ, Σ)

What social-interactional situation are the agents embedded in, and what are they trying to do? What are the constraints on what they can do? Several distinctions need to be made.

The first is between *external* and *internal* (or 'subjective') goals. Clearly, the goals of a collaborative learning task, as determined by a teacher or an experimenter, are not necessarily the same as those perceived and pursued by the learner agents. Within agents' internal goals the possibility of conflict between them is an important factor to be taken into account.

Secondly, agents' goals can be considered as relating to the concrete activity to be performed (for example, acquiring understanding of a concept or a fact, solving a problem) or as situational (for example, the goal of pleasing the teacher, or getting a good mark in an educational institution). Such situational goals are not necessarily 'inferior' to those that directly relate to the task to be achieved. In fact, the situational goals may be exploited in school settings as vehicles for initiating children in activities which, when adequately supported, may lead to concentration on the task goals themselves (Leontjev, 1981)

---

[5] See e.g. Agre (1995) for a review of research on artificial agents in artificial intelligence and Winograd & Flores (1986) for a well-known critique of this AI view.

Social-institutional situations may be viewed as imposing a set of constraints, as may concrete activities or problems. Both need to be taken into account, but for the purpose of this chapter it is not useful to make a sharp division between the two; thus we have treated goals and situations as a single yet composite entity in the framework.

Thirdly, any specific instantiation of the model presupposes an understanding of what is to count as a 'situation'. Thus, from one point of view, a problem solved during a psychology experiment is not a 'situation' to the extent that the social relationship between the experimenter and subject, and the latter's understanding of the external goals (see above) are not taken into account. From another point of view (see e.g., Perret-Clermont, Perret & Bell, 1991) an experimental situation is a social situation *per se*, given the fact that the experimenter will necessarily represent some social role or position from the point of view of the subject.

In general, situations will be multifaceted, and there can be several goals for each agent or group of them, along the distinctions outlined above.

5.3 Tools (τ)

As mentioned above, tools are both *objects* of (mutual) understanding and *mediators* of it.

There are few *a priori* limits to what can actually be a tool: what makes something a tool or not is the role that it plays in activity itself. On the other hand there is a crucial distinction in types of tools, in that sometimes an object becomes a tool in a particular situation, in an *ad hoc* manner, whereas other objects are *created* to be tools for *particular kinds of activities.* It is, of course, the latter type of tools which are complex cultural achievements, and which accordingly require instruction in order to be appropriated.

We consider two main types of tools: *material* tools (such as hammers and computers and *semiotic* tools (such as natural or formal languages). The role of each is well illustrated by the case of computer-mediated collaborative work. Here, the technology is on one hand, of course, a material tool: it allows concrete tasks to be achieved, including the task of communicating. On the other hand, the technology is also the support for semiotic tools such as natural language and other formal languages that may be incorporated in the task environment (see later §6 of this chapter).

Finally, it is to be noted that, within our model, tools (or any other entity) are not to be considered in isolation from other entities: they are not mere 'add-ons' to humans; the unit of analysis is person(s)-acting-with-mediational-means, viewed as an indissociable whole (Mammen, 1993).

5.4 Relations in the framework: processes of understanding and learning

A single type of relation between entities in the model is taken into account: (tool-mediated) *understanding*. There are in fact two forms of it: "agent α's understanding of x", and "agents α_1 and α_2's mutual understanding of x". In both cases, "x" can also be another agent or a group of them. We do not give here a formal definition of *understanding*, relying for our present purposes on common-sense intuitions. Of course, "understanding" will be different for different types of agents. While each of the following sentences is meaningful in an appropriate context, the word "understanding" does not capture precisely the same relationship between agent and understood concept: "John understands the energy chain model.", "The computer understands what John meant in drawing a box.", "Volkswagen understands how to sell cars.".

The relations of understanding between agents α_1/α_2 and the goals-in-situation pass *exclusively* via mediating tools. This means that, in our model, we exclude consideration of understanding that is *not* mediated by tools.

We also exclude the case of a single solitary agent understanding by perception and action (even though this is also arguably mediated by semiotic tools), since our model is restricted to situations where several agents collaborate. In this case, it is clear that nearly all understanding will be mediated by tools, especially when communication is involved. Given our own definition of collaborative learning described above, all understanding is mediated by semiotic tools, and sometimes by material tools. In fact, the grounding process, as modelled by Clark & Schaefer (1989) and Traum (1994), whose constituents are speech acts such as acknowledgements, verifications, and repairs, is itself a prime semiotic tool for achieving mutual understanding.

As pointed out above, *learning* may take place when the entities and relations of the model undergo *change* through *time*. Usually we will be concerned with changes that are viewed as positive from some normative point of view (what is usually understood as learning — Johnson-Laird, 1988, p. 129). Theoretically, therefore, the following types of change and learning are possible in collaborative learning tasks (see Table 2).

Table 2. Types of change (and learning) in the framework for analysing collaborative learning tasks[†]

	Change (model syntax)	Explanation / example
Entities	$\Delta\ \alpha_i$	cognitive change in an agent (learning, un-learning, acquiring 'correct' or 'incorrect' knowledge or beliefs)
	$\Delta\ \tau$	change in material or semiotic tools (e.g. introduction of a new computer system, language development)
	$\Delta\ \gamma,\Sigma$	change in agents' goals-in-situation (e.g. pedagogical goals may change in educational institutions)
Relations	$\Delta\ \alpha_1 \leftrightarrow \alpha_2$	change in mutual understanding between agents
	$\Delta\ \alpha_1 \rightarrow \alpha_2\ /\ \alpha_2 \rightarrow \alpha_1$	change in an agent's understanding of another agent
	$\Delta\ \alpha_i \rightarrow \tau$	change in agents' understanding of tools
	$\Delta\ \alpha_i \rightarrow \gamma,\Sigma$	changes in agents' understanding of goals and situation

[†] " $\rightarrow$", " $\leftrightarrow$" represent relations of tool-mediated understanding and mutual understanding respectively; Δ x represents change in x.

In most collaborative learning situations, it is assumed that the tools, goals and situation do not undergo change, often because the *timescale* considered is usually little more than an hour. In our framework we assume a *complex dynamic system*, in which all entities and relations between them can change. In the case of material tools or social situations, for example, this will usually involve considering a much longer timescale (e.g. for a computer system or for educational practices to change), in accordance with CHAT (see Cole & Engeström, 1993). Thus, the necessity to include consideration of a wider timescale radically changes the types of understanding that need to be taken into account in the collaborative learning situation under study. Once we consider a broader timescale, we need to consider changes in goals of institutions and of the tools that are elaborated and appropriated. On a micro level, grounding is a prime mechanism by which these global changes take place.

Changes in one entity or a relation can thus effect changes in another, and this may be bi-directional. The specific causal relation with which we are principally concerned here is of course collaborative learning, a cognitive change in an agent produced as a result of a change in their mutual understanding during interaction: $\Delta\ \alpha_1 \leftrightarrow \alpha_2$ *=causes=>* $\Delta\ \alpha_i$. The causal arrow may, however, be reversed since changes in agents effected by other means may trigger the necessity to attain new mutual understanding. There may be several cycles of this as agents oscillate between understanding and non-understanding

(Miyake, 1986). Changes in mutual understanding may also concern the tools, goals and situation. Similarly, when tasks and goals change over longer periods of time (either as represented by agents or as a result of changes in a social situation, such as within a school) then this may require and cause changes in agents understanding of each other.

Mapping out the relevant entities and understanding involved in collaborative learning tasks, and studying the possible range of interdependencies between changes, thus enables us to consider a wider range of types of learning than is usually taken into account. In addition, it provides a way of modelling collaborative learning situations within which grounding and collaboration processes are integrated within use of material and semiotic tools.

6. An illustrative analysis: a Computer-Supported Collaborative Learning environment for modelling energy in physics

We now illustrate the use of our framework in analysing use of a specific computer-mediated collaborative learning ("CSCL") environment — C-CHENE[6] (Baker & Lund, 1997). The environment was designed for learning about modelling in physics, and the theory/model of energy in particular (Tiberghien, 1994, 1996; Devi, Tiberghien, Baker & Brna, 1996). We have chosen to analyse this environment since the task it supports provides a good illustration of the complexity of material and semiotic tools and because it has been experimented with using two different computer-mediated communication interfaces. The latter allows us to compare the different grounding mechanisms provided by the interfaces and to suggest how they could be improved in further research.

6.1 The energy modelling task

In the part of the energy modelling teaching sequence discussed here, students are asked to produce qualitative models (diagrams) called "energy chains" in order to represent energy storage, transfer and transformation in simple experimental situations (the task will be discussed in more detail throughout the ensuing analysis). Figure 3 shows an example (correct) energy chain for an experimental situation where a battery is linked to a bulb by two wires (see later Table 3 for an example of a student energy chain for this experiment).

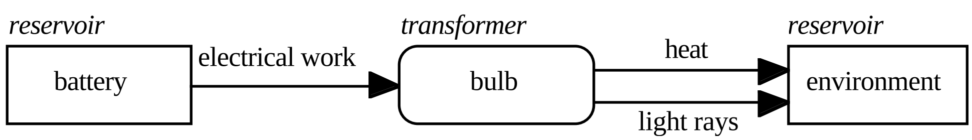


Figure 3. Energy chain (correct) for experiment where a bulb is connected to a battery by two wires.

The didactic rationale of this task is as follows: by attempting to instantiate the simple qualitative model for energy in concrete cases, the students will be led to co-construct an understanding of the concepts of "energy" and "model".

---

[6] "CHENE" = "CHaîne ENErgétique" = "Energy Chain". "C-CHENE" = Collaborative CHENE.

In the studies referred to here, the students worked in pairs, and at a distance via a network, each on their separate computers but sharing and working on the same virtual space. Energy chains were drawn using a customised graphical interface and all (typewritten) communication during the task execution passed via specially designed communication interfaces.

6.2 The agents (α_1 and α_2)

The students who participated in the studies were 16-17 years old, and all came from the same class of a secondary school (lycée) in Lyon. Their teacher asked for volunteers who had some facility with typing on a computer keyboard. The students grouped themselves into "friendship pairs".

6.3 The goals and the situation (γ, Σ)

In any learning situation, we clearly can not assume that the goals that designers of the situation intend learners to have will genuinely be perceived and adopted by them. We therefore distinguish: the overall research goals of the task designer, the designers' goals with respect to what the students are intended to learn, and the goals that the students appeared to be following (from the researchers' analysis).

The overall research goals of the studies were to understand the cognitive processes of modelling used by students and to understand how students collaborate in performing such tasks, focusing on the role of computer-mediated communication. The designer's learning goals for students were as follows. As a result of creating energy chains for the three experiments, students should gain a better understanding of: (a) the concept of energy in physics (e.g. energy can neither be created nor destroyed, it does not correspond to a 'tangible' entity like a physical object such as a battery, nor to one of its electrical properties, such as resistance); and (b) the nature of modelling in science (e.g. there is no necessary one-to-one matching between model and reality, models select certain phenomena but not others, they have an internal syntax, and so on).

The task was not invented solely as a research tool, but also with the educational goals that it should be used in actual classrooms for learning about energy (it is currently being incorporated in the French physics teaching National Curriculum). The educational goal of the study was therefore also to test the feasibility of a part of a complex teaching sequence in a computer-mediated communication setting. In technological terms, we also had the goal of testing the usability of a relatively new version of the C-CHENE software.

From analysis of students' interactions with C-CHENE (Baker & Lund, 1997), the students' goals appear to have been the following: (a) to manage using a piece of complex software, with which they were previously unfamiliar; (b) to draw an energy chain that satisfied the energy chain model rules; (c) to draw an energy chain that they both agreed upon; and (d) to have a friendly discussion.

Students were not always concerned with getting the 'right answer' (although, rarely, some did obtain it), nor with understanding energy modelling. Rather, they often aimed to produce a solution that was minimally satisfactory to each, and to get through the task with minimum effort, especially with minimum (typewritten) communication. The most important researcher goals — that the students should learn what energy and modelling are — was rarely explicitly recognised (i.e. discussed or adopted). The fact that the task was carried out outside school hours, in a research institution without teacher presence or assessment is undoubtedly important in explaining the students' activity.

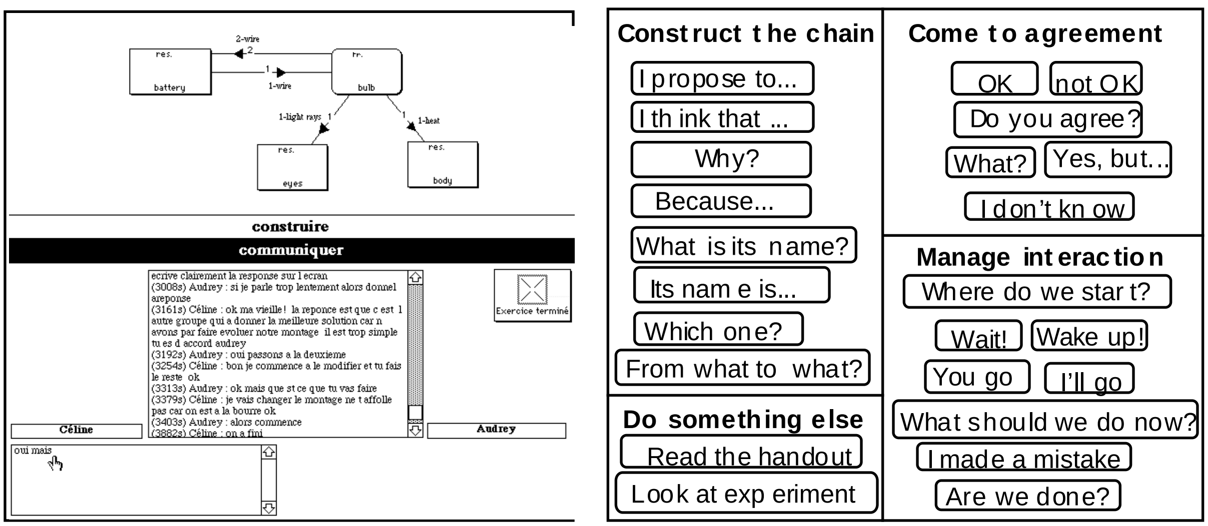
6.4 The tools (τ)

The principal tools provided and used for the task are both material (computer software and hardware) and semiotic. Into the latter category, as well as natural language (typewritten), the energy chain graphical language is of particular interest, as is the fact that the semiotic tools are embedded in specific material tools.

6.4.1 Material tools (technology).

Collaborative problem-solving interactions with two different communication interfaces were compared (see Figure 4 below): (1) a “chat-box” interface and (2) a “dedicated" interface[7].

In the first case, students type their messages in their own boxes, then ‘send’ them (tabkey), which causes the sender's box to close and that of the other student to open. The message sent is added to the end of a continually visible interaction history. The upper part of the screen is where energy chains are drawn, with specially adapted menus; change between construction and communication modes is done with the two central buttons that go across the screen. The idea of such a customised design was to visually reinforce the co-ordination constraints here: communication and construction can not be done simultaneously, and only one student can communicate at a time.



“Chat-box” interface *“Dedicated” interface*

Figure 4. Two communication interfaces of the C-CHENE CSCL environment: left, "chat-box"; right, "dedicated".

The dedicated interface provides buttons for specific speech acts required for performing this task (see Baker & Lund, 1997, for design considerations). Again, the interaction history is visible. The principle behind this interface was that the speech act buttons would alleviate the typing burden, freeing up effort for the energy chain task itself, and that the presence of buttons such as “because” and “why?” would encourage students to reflect on their solutions, during discussion, instead of simply drawing their solution. Specific buttons are also provided for facilitating *grounding:* “OK”, “not OK”, “I don’t know”, “do you agree?”.

---

[7] Only the communication buttons are shown for the dedicated interface in Figure 4, redrawn and translated from the original French. The energy chain construction interface remains the same as with the chat-box interface.

6.4.2 Semiotic tools (typewritten natural language and the energy chain language).

In addition to the natural language used by the students to communicate, the energy chain model is also a semiotic tool. It was not a pre-existing cultural object since it was created as a didactic simplification, whose use by students in the solving of specific modelling problems would lead them to co-construct understanding of the concepts of energy and of modelling itself.

The ‘lexicon’ of the model is a set of boxes, arrows and labels with determinate meanings within the model itself. It has a simple set of syntactic constraints (e.g. “start and end with a reservoir”, “no more than three arrows between two boxes”). The didactic aim is then that students will co-construct the *semantics* of the language — the intermediary concepts necessary for understanding the nature of the relations between model and reality (Tiberghien, 1994).

The specific way in which this semiotic tool is embedded in the material tools may have important effects on students' cognitive processes. For example, when drawing energy chains on paper, students could decide that “the bulb gives out light”, and so draw an arrow out from the bulb without specifying where it goes to. This is not possible on the graphical interface: the arrow can not go in mid air, it must be linked to another box. So in this case, students must draw a box (unlabelled at present), deciding already whether it is a transformer or a reservoir. They are therefore obliged to make a more global matching of a subpart of an energy chain. Although such global matching represents a more complex modelling process, it is not clear whether imposing it from the start will facilitate learning. The important point to be recognised is that such constraints need to be considered explicitly in designing educational tools.

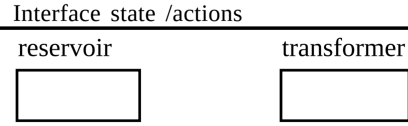
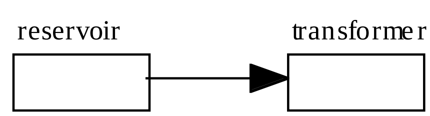
6.5 Relations in the framework: grounding and learning

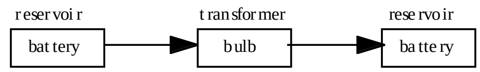
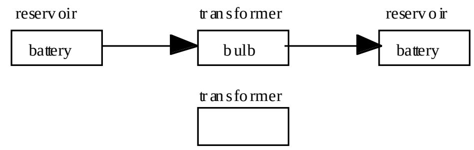
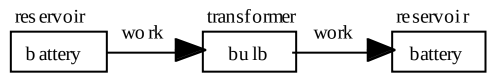
We discuss relations in our analytical framework — what needs to be grounded and how it is grounded — with respect to two example interactions, produced by pairs of students with the chat-box and dedicated interfaces, respectively, of C-CHENE.

6.5.1 Example interaction sequences

Table 3 shows a complete interaction of 30 minutes duration (see also §2.2 above) produced by two female students for the battery-wires-bulb experiment, using the chat-box interface of C-CHENE.

Table 3. Interaction with chat-box interface of C-CHENE, students S1 and S2, battery-bulb experiment†

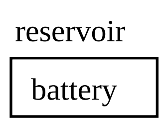
N	T (s)	S	Dialogue	Interface state /actions
1	208	S2		 *(link transformer -> reservoir created then deleted)*
2	451	S1	I’ll do the transfer	
3	482	S2	what about me	
4	527	S1	you the second reservoir	
5	564	S2	OK, go ahead	
6	591	S1		
7	617	S1	over to you	

8	707	S2	but you have to start with a reservoir and finish with a reservoir	
9	803	S1	I began with a reservoir then a transformer now I have to put a reservoir OK	
10	828	S2	OK	
11	868	S1		
12	1173	S1	do you think it's right, what I did with our diagram	
13	1259	S2	it's right but there has to be a generator or something like that	
14	1309	S1	the battery is the generator! banana brain!	
15	1369	S2		
16	1408	S1	what are you doing	
17	1530	S1		
18	1715	S1	I've named the transfers because that's what's asked for on the sheet about transfers! look!	
19	1760	S2	you and your errors so give me a break all the time's going, we have to finish	
20	1830	S1	it's finished now OK let's call the old lady OK my dear!	
			*** end of interaction ***	

† The interaction has been translated from the original French. N = line number; T(s) = time from start of session in seconds; S = speaker (S1 or S2).

The first eleven minutes of the interaction shown in Table 3 (lines 1-11), during which a reservoir for the battery and a transformer for the bulb were created, can be compared with the following extract from the beginning of a dedicated-interface interaction for the same task (Table 4 below), of ten minutes duration. Only the part up to where the students create a reservoir for the battery and a transformer for the bulb is shown, for sake of brevity.

Table 4. Dedicated-interface interaction (extract) for battery-bulb experiment. †

N	T(s)	S	Dialogue	Interface state / actions
1	56	S1	<Read the sheet>	
2	95	S1	<Where do we begin?>	
3	121	S2	<I propose to …><create a reservoir>	
4	144	S1	<Which one?>	
5	179	S1	<Hello there!>	
6	200	S2	<I don't know>	
7	213	S1	<I think that …> the first reservoir is called battery	
8	262	S1	<Do you agree?>	
9	287	S2	<Agreed>	
10	312	S1		
11	357	S1	<What shall we do?>	
12	376	S2	<I propose to …><create a reservoir>	
13	397	S1	<Not agreed>	
14	412	S1	<I propose to …><create a transformer>	
15	432	S1	<Do you agree?>	
16	457	S2	<Why?> <14>	

17	472	S1	<I think that …> the bulb is a transformer of electrical energy
18	450	S2	<Agreed>
19	554	S1	<I think that …> I'm going to create the transformer called bulb
20	604	S2	<Agreed>
21	645	S1	
		S2	<…>

†Text between "<…>" corresponds to an automatic tracing of the speech act buttons that are clicked. All other text is typed by the students themselves. Numbers such as "<4>" refer to line numbers of the automatic transcription, as referred to by the students (e.g. line 16 above). Again, the example is translated from the original French.

6.5.2 Tool-mediated grounding of the students' interaction (α_1—[τ]—α_2)

The two interactions shown above differ markedly in the form of the students' collaboration (Baker, 1995) — who will act (graphically), who will communicate, when and with respect to what — and consequently, in the way that grounding takes place.

In the case of the chat-box interaction (Table 3), the students' interaction appears to be *graphical-action-based*, communication being reduced to a minimum (e.g. the interaction even begins by S1 drawing two rectangles without previous discussion), this being one way of reducing co-ordination problems with the interface. Instead of attempting to co-construct the solution, on the basis of a more reflective discussion, the students adopt collaboration pattern that reduces to: "you draw, I will tell you what I think". Within this restricted form of collaboration, co-ordination is also negotiated explicitly from the outset of the interaction (lines 2-5). *Grounding* therefore often takes place via a combination of perception of graphical action and communication, but which is not always successful in maintaining mutual *understanding* (e.g. line 16 — "what are you doing?"). The students do however seem concerned to attain mutual *agreement:* they ask whether the other agrees (line 12) and when they have performed a graphical action without previously negotiating who will act, they explain their actions (line 18).

The form of collaboration and the way in which grounding takes place with the dedicated interface interaction (Table 4) is, by contrast, much more *discussion-based*. The students begin by discussing a common strategy, by proposing a solution element which will not be drawn before common agreement is reached (see e.g. lines 3-10 in Table 4). The form of collaboration is "discuss possible solution, reach agreement, draw solution". As well as explicit grounding acts on the *agreement* level (lines 8,9,13,15,18,20), the students also ground the interaction along the levels of *contact* (line 5) and *understanding* (lines 16-17)[8].

With both interfaces, however, although grounding can take place via co-perception of graphical actions, students also had to perform some activities individually — they each had a paper copy of the energy model and their own experimental apparatus — in which case there was no common perception. This obliged the students to make explicit what they were doing, and what aspect of the text they were talking about, thus adding to communication costs. In future designs of the system, these activities therefore need to be co-perceivable. On the other hand, the presence of a commonly perceivable interaction history may have helped in focusing the interaction, this being one of the possible advantages of computer-mediated communication over face-to-face audio communication.

8 See discussion of Allwood *et al.* (1991), in previous §2.1 of this chapter.

6.5.3 Students' understanding of the tools (α_i—τ)

The forms of collaboration and grounding that were produced with each interface appear to be linked to the extent to which students are able to *appropriate (learn) the material tools* — the computer-mediated communication interfaces. Given that their communication was typewritten, producing linguistic feedback in grounding involves a higher production cost with the chat-box interface than with the dedicated one (clicking on a button). This may in turn explain why a more discussion-based form of collaboration was adopted, accordingly, with the dedicated interface.

The students were given a 30 minute training session prior to the study, in which they were supposed to learn to use the physical tools (computer interfaces). Clearly this was insufficient, since students experienced co-ordination problems during the study. In fact, it is difficult to see why students should invest much effort in learning the tools, since they knew that they were research prototypes which they would be unlikely to encounter subsequently.

The specific semiotic tool used here — the energy chain model — presented few learning difficulties, at least on the surface level. On a deeper, semantic, level, learning about this tool constituted one of the major goals of the task, from the researchers' point of view, i.e. the desired learning here may be viewed as *appropriation of a semiotic tool*. Interestingly, students did not necessarily view this tool as 'fixed', even though it was provided from an authoritative source (the researcher): they sometimes elaborated its meaning[9].

6.5.4 Students' understanding of goals-in-situation (α_1—γ, Σ—α_2)

As mentioned above, the students' main goals with the chat-box interface interaction rarely appeared to match with those of the researchers: acquiring a deeper understanding of energy and modelling. The students' main goals appeared to be: *learn how to use the tools*, *draw a diagram* (solve the explicitly given task) and *maintain social interaction*. With the dedicated interface interaction, however, students did engage in explanation of the nature of energy (see e.g. lines 16-17 of Table 4). Later on in that specific interaction, more extended explanation sequences were produced that more closely approached the researchers' goals.

There are two main lines of explanation for these facts. The first is that the cognitive load associated with learning the material tools was sufficiently high so as to exclude the kind of reflection that would be associated with gaining deeper understanding of the domain. Thus the students tried to minimise co-ordination problems associated with switching between communication and construction. The second relates to the novelty of the situation in which the students found themselves. Given that the task took place outside a school environment during free time, without teacher assessment, this meant that little was actually at stake for the students. They therefore expended little effort with the task beyond learning the material tools themselves and minimally drawing a diagram.

All these factors contribute to explaining the students' apparent goals as a function of learning to understand each other and the tools. The interest of such an *a posteriori* analysis is therefore that it highlights how the situation, the goals and the tools require re-designing so as to bring researcher and student goals closer together.

---

[9] An example is described in Baker (1996); one student implicitly elaborated the meaning of the energy chain rule "A complete energy chain must start and end with a reservoir" to include "… and the first and last reservoirs can not correspond to the same object or event in the experimental situation".

6.6 Reflections: how does analysis with the model inform re-design?

Above all, the analysis reveals the importance of *tools* — material and semiotic — in mediating understanding and learning within a computer-mediated communication situation.

Firstly, the physics task can be viewed as one of appropriation of a semiotic tool — the energy chain model. In future versions of C-CHENE, we therefore need to carefully reconsider the way in which this cognitive tool is embodied in material tools so that the constraints that are imposed are those that are most likely to foster modelling processes in students' problem-solving that are at an appropriate level of complexity for them at a given point in their problem solving.

Secondly, the communication (material) tools provided have a marked effect on grounding. Although some progress was made in facilitating grounding in language interaction with the second dedicated interface, other aspects of the task (energy model rules, experimental apparatus) need to be made mutually perceivable during communication to reduce grounding costs.

Thirdly, the difficulties involved in learning to use the communication tools may be important in explaining the mismatch between researcher and student goals. Clearly, if we want students to learn energy and modelling concepts, we have to reduce the cognitive-communicational load required in learning the tools, and in learning about the task.

Finally, we need to directly address what the students' understanding of the task situation is so that, even outside school constraints, there is something at stake for them in problem-solving.

7. Concluding remarks

Our main aim in this chapter was to explore synergies between two areas — language sciences research on grounding, and CHAT research on learning — as a means of furthering our theoretical understanding of collaborative learning. This exploration gave rise to a simple framework for analysing and designing collaborative learning situations (especially those that involve computer-mediated communication) that highlights the different forms of grounding that need to be taken into account.

From the point of view of understanding the cognitive effects of communicative interactions produced in teaching-learning situations, the CHAT view of learning as appropriation of tools appears particularly appropriate, given its emphasis on the role of language, the 'tool of tools', in interpersonal interactions and in thinking. From the point of view of CHAT, micro-analyses of the interactive processes at work in grounding, as well as the different planes on which it can operate, may be of potential use in understanding precisely how the interpersonal plane of functioning within the ZPD leads to appropriation of tools.

The overall picture of collaborative learning that emerges from this research may be summarised as follows: collaborative learning is associated with the increased cognitive-interactional effort involved in the transition from learning to understand each other to learning to understand the meanings of the semiotic tools that constitute the mediators of interpersonal interaction. Understanding how and when grounding leads to collaborative learning therefore requires detailed analysis of learners' *goals* in specific *situations* that motivate them to go further in their attempt to gain mutual understanding, of the extent to which the *material and semiotic tools* that they use can be appropriated, and, especially, the different forms of mutual understanding that will need to be achieved by the learners with respect to these 'objects'. It is precisely these aspects of collaborative learning situations that our analytical framework is designed to highlight.

We have proposed a first step in a possible longer term research project that aims to elaborate a new unifying theoretical framework for understanding collaborative learning. Although a much more far-reaching theoretical integration remains to be achieved, we nevertheless hope that this work will provide a basis for further research from the point of view of each of the contributing disciplines. As with any such framework, an important test of its usefulness will be the extent to which it can give rise to the design of effective collaborative learning situations in the future, particularly those that involve computer-mediated communication.

Acknowledgements

This research was carried out within the framework of the European Science Foundation research programme "Learning in Humans and Machines", to which we are grateful for financial support. David Traum was partially supported while writing this paper by the Swiss National Science Foundation under Research contract 1114-040711.94, the U.S. Army Research Office under contract/grant number DAAH 04 95 10628 and the U.S. National Science Foundation under grant IRI-9311988. Tia Hansen was supported by the Danish Council for Research in the Humanities under grant no. 9502778. We would like to thank members of the GRIC-COAST research team who participated in the studies analysed in this chapter and Bob Lewis, Pierre Dillenbourg, Karen Littleton, Ulrich Hoppe, Daniel Schwartz and Yrö Engeström for comments on earlier versions of this work.

References

Agre, P. (1995). Computational research on interaction and agency. In P. Agre & S. Rosenschein (Eds.) *Computational Theories of Interaction and Agency* (Special Issue). *Artificial Intelligence*, 72, 1-52.

Allen, J. F., Schubert, L. K., Ferguson, G., Heeman, P., Hwang, C. H., Kato, T., Light, M., Martin, N., Miller, B., Poesio, M., & Traum, D. R. (1994). The TRAINS project: a case study in building a conversational planning agent. *Journal of Experimental and Theoretical Artificial Intelligence*, 7, 7-48.

Allwood, J. (1976). *Linguistic Communication in Action and Co-operation: A Study in Pragmatics*. Gothenburg Monographs in Linguistics 2. University of Gothenburg, Department of Linguistics, Sweden.

Allwood, J., Nivre, J. & Ahlsén, E. (1991). *On the Semantics and Pragmatics of Linguistic Feedback.* Gothenburg Papers in Theoretical Linguistics No. 64. University of Gothenburg, Department of Linguistics, Sweden.

Baker, M. J. & Lund, K. (1997). Promoting reflective interactions in a CSCL environment. *Journal of Computer Assisted Learning*, 13, 175-193.

Baker, M. J. (1994). A model for negotiation in teaching-learning dialogues. *Journal of Artificial Intelligence in Education*, 5 (2), 299-254.

Baker, M. J. (1995). Negotiation in Collaborative Problem-Solving Dialogues. In R.-J Beun, M.J. Baker & M. Reiner (Eds.) *Dialogue and Instruction*, pp. 39-55. Berlin: Springer-Verlag.

Baker, M. J. (1996). Argumentation et co-construction des connaissances. *Interaction et Cognitions* 1(2-3), 157-191. Paris: L'Harmattan.

Bunt, H.C. (1995). Dialogue Control Functions and Interaction Design. In R.J. Beun, M.J. Baker & M. Reiner (Eds.) *Dialogue and Instruction,* pp. 197-214. Berlin: Springer-Verlag.

Clark, H. H & Brennan, S. (1991). Grounding in communication. In L.B. Resnick, J.M. Levine & S.D. Teasley (Eds.) *Perspectives on Socially Shared Cognition*, pp. 127-149. Washington DC: American Psychological Association.

Clark, H. H. & Schaefer, E. F. (1987). Collaborating on contributions to conversation. *Language and Cognitive Processes*, 2, 1-23.

Clark, H.H. & Marshall, C.R. (1981). Definite reference and mutual knowledge. In A.K. Joshi, B.L. Webber & I.A. Sag (Eds.) *Elements of Discourse Understanding* , pp. 10-63. Cambridge: Cambridge University Press.

Clark, H.H. & Schaefer, E.F. (1989). Contributing to Discourse. *Cognitive Science* 13, 259-294.

Clark, H.H. & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22, 1-39.

Clark, H.H. (1994). Managing problems in speaking. *Speech Communication*, 15, 243-250.

Cole, M. & Engeström, Y. (1993). A cultural-historical approach to distributed cognition. In G. Salomon (Ed.) *Distributed Cognitions: Psychological and educational considerations*, pp. 1-46. Cambridge, Mass.: Cambridge University press.

Danieli, M. & Gerbino, E. (1995). Metrics for evaluating dialogue strategies in a spoken language system. In *Proceedings of the 1995 AAAI Spring Symposium on Empirical Methods in Discourse Interpretation and Generation*, pp. 34-39.

Devi, R., Tiberghien, A, Baker, M., & Brna, P. (1996). Modelling students' construction of energy models in physics. *Instructional Science*, 24, 259-293.

Dillenbourg, P., Jermann, D., Schneider, D., Traum, D & Buiu, C. (1997). The design of MOO agents: Implications from a study on multi-modal collaborative problem solving. In *Proceedings, 8th World Conference on Artificial Intelligence in Education* (AI-ED 97), Kobé (Japan), pp. 15 - 22. Amsterdam: IOS Press.

Forman, E. A. (1992). Discourse, Intersubjectivity, and the Development of Peer Collaboration: A Vygotskian Approach. In L. T. Winegar & J. Valsiner (Eds.) *Children's Development within Social Context*: *Vol 2. metatheoritical, theoretical and methodological issues*, pp. 143-159. Hillsdale, NJ: Lawrence Erlbaum Associates.

Johnson-Laird, P.N. (1988). *The Computer and the Mind: An Introduction to Cognitive Science*. London: Fontana.

Lave, J. (1991). Situating Learning in Communities of Practice. In L.B. Resnick, J.M. Levine & S.D. Teasley (Eds.) *Perspectives on Socially Shared Cognition*, pp. 63-84. Washington D.C.: American Psychological Association.

Leontjev, A. N (1981). The problem of activity in psychology. In J. V. Werstch (Ed.) *The concept of activity in soviet psychology*, pp. 37-71. Armonk, NY: Sharp.

Levinson, S.C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.

Luria, A.R. (1979). *The Making of Mind.* Cambridge, Mass.: Harvard University Press.

Mammen, J. (1993). The elements of psychology. In N. Engelsted, M. Hedegaard, B. Karpatschof & A. Mortensen (Eds.) *The Societal Subject*, pp. 29-44. Århus: Århus University Press.

Miyake, N. (1986). Constructive Interaction and the Iterative Process of Understanding. *Cognitive Science*, 10 (2), 151-177.

Moeschler, J. (1985). *Argumentation et Conversation.* Paris: Crédif-Hatier.

Perret-Clermont, A-N., Perret, J-F. & Bell, N. (1991). The Social Construction of Meaning and Cognitive Activity in Elementary School Children. In L.B. Resnick, J.M. Levine & S.D. Teasley (Eds.) *Perspectives on Socially Shared Cognition*, pp. 41-62. Washington D.C.: American Psychological Association.

Rommetveit, R. (1979). On the architecture of intersubjectivity. In R. Rommetveit & R. M. Blaker (Eds.) *Studies of language, thought, and verbal communication*,. pp. 93-108. London: Academic Press.

Rommetveit, R. (1985). Language acquisition as increasing linguistic structuring of experience and symbolic behaviour control. In J. Wertsch (Ed.) *Culture communication and cognition: vygotskian perspectives*, pp. 183-204. Cambridge: Cambridge University Press.

Roschelle, J. & Teasley, S.D. (1995). The Construction of Shared Knowledge in Collaborative Problem Solving. In C. O'Malley (Ed.)*Computer Supported Collaborative Learning*,. pp. 69-100. Berlin: Springer-Verlag.

Roulet, E. (1992). On the Structure of Conversation as Negotiation. In H. Parret & J. Verschueren (Eds.) *(On) Searle on Conversation*, pp. 91-100. Amsterdam: John Benjamins.

Schwartz D. L. (1995) The emergence of abstract representations in dyadic problem solving. *Journal of the Learning Sciences*, 4, 321-354.

Smith, R.W. & Hipp, D.R. (1994). *Spoken Natural Language Dialog Systems: A Practical Approach*. Oxford: Oxford University Press.

Tiberghien, A. (1994). Modeling as a basis for analyzing teaching-learning situations. *Learning and Instruction* 4, 71-87.

Tiberghien, A. (1996). Construction of prototypical situations in teaching the concept of energy. In G. Welford, J. Osborne & P. Scott (Eds.) *Research in Science Education in Europe*, pp. 100-114. London: Falmer Press.

Tolman, C.W. (1988). The Basic Vocabulary of Activity Theory. *Multiple Newsletter for Activity Theory*, 1, 14-20.

Traum, D. R. (1994). *A Computational Theory of Grounding in Natural Language Conversation*. PhD thesis, Department of Computer Science, University of Rochester. Also available as TR 545, Department of Computer Science, University of Rochester.

Vygotsky L. (1978). *Mind in Society: The development of higher psychological processes.* M. Cole, V. John-Steiner, S. Scribner & E. Souberman (Eds.). Cambridge: Cambridge University Press.

Vygotsky L. (1986). *Thought and Language.* A. Kozulin (Ed.) Cambridge, MA: MIT Press.

Wells, G. (1996) Using the tool-kit of discourse in the activity of learning and teaching. *Mind, Culture and Activity*, 3 *(2)*, 74-101.

Werstch, J. V. (1985). *Vygotsky and the social formation of mind*. Cambridge, Mass.: Harvard University Press.

Wertsch, J. W. (1994). The Primacy of Mediated Action in Sociocultural Studies. *Mind, Culture, and Activity*, 1 (4), 202-208.

Winograd, T. & Flores, F. (1986). *Understanding Computers and Cognition*. Reading, Mass.: Addison-Wesley.