Learning in contexts. Four cases in higher education.

Bannink, E.A.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
1. RELEVANT RESEARCH AREAS AND THEORETICAL APPROACHES
1. RELEVANT RESEARCH AREAS AND THEORETICAL APPROACHES

1.1 The social analysis of context

It does not seem possible to come up with a single, precise definition of the notion of context, since it means different things within alternative research paradigms (Goodwin & Duranti 1992). In this section a brief, non-exhaustive survey will be presented of a variety of different perspectives on the analysis of context to date.

As early as 1923 the anthropologist Malinowski proposed that linguistic analysis should be supplemented by ethnographic analysis of the situations in which speech occurs. He introduced the notion of a context of situation. He also argued that language should be conceptualized as a mode of practical action and stated that language is ‘an indispensable element of concerted human action’ (Malinowski 1923: 316 cited in Duranti & Goodwin 1992: 15). Wittgenstein (1958) corroborated this analysis in his *Philosophical Investigations*. He emphasized the embedding of language within human activities and suggested that understanding a language, and by implication, being able to grasp the meaning of utterances, involves knowing the nature of the activity in which the utterances play a role. Wittgenstein therefore proposed that context should be used as a point of departure for uncovering the multifaceted variety of thought and action made available by the different language games human beings engage in.

Austin (1962) built on Malinowski’s and Wittgenstein’s stance that language is used to perform action. He presented an apparatus to talk about how utterances become social acts and distinguished three types of speech acts we simultaneously perform when we speak: a locutionary act – the act of saying something; an illocutionary act – the act a speaker can perform in saying something by means of the conventional force of the locutionary act; a perlocutionary act – the effect
produced by uttering a particular locutionary act. In other words, in saying something, we are always doing something. Austin argued that any act of speaking takes place within a particular context and is evaluated with respect to such a context. He saw context as a set of recognizable conventions which provide the infrastructure through which an utterance gains its force as a particular type of action. It prompted him to introduce the notion of felicity conditions: conditions which have to be fulfilled for a speech act to be successful. Austin’s ideas were made known to a broader public by John Searle (1969) in his speech act theory. But this theory has its problems. It gives analytic primacy to the isolated sentence and seeks in the first place to establish the ‘literal’ meaning of the utterance. It then goes on to account for variations in meaning according to the context. This entails that the context is seen by speech act theory as some sort of overlay: it fails to see that utterances are in the first place contextually produced and understood. Besides, the assumptions speech act theory draws on do not take cultural variability into account: ‘action’ is not a universal dimension of human existence and does need further analysis (Duranti 1997). Moreover, speech act theory takes the conscious intentions of the speaker as the originator of the meaning making process and sees the hearer as a projection of the speaker’s wants and attitude: the hearer is not really seen as an active participant in the interaction. Finally, it does not take into account that many situations exist in which actions emerge not from the speaker alone in a single turn, but rather are collaboratively defined through a process of interaction in which recipients play a very active role (cf. Kendon 1992).

The notion of speech event presented a solution to the problems mentioned. Hymes was the first to propose a non-linguistic unit, the ‘event’ as the frame of reference for interpreting speech (1972). He defined speech events as: ‘activities, or aspects of activities, that are directly governed by rules or norms for the use of speech’ (1972: 56), which meant attention to units of discourse larger than the sentence. Examples of speech events include: a service encounter, a business meeting, a telephone conversation, a lesson, etc. For each speech event Hymes defined a set of eight conventions as particular constraints on setting, participants, roles, genre, the ‘business’ at hand and the procedures. These conventions should be seen as normative expectations that provide a grid for the organization of experience rather than fixed rules for appropriate and non-appropriate behavior.

Sociology has made several important contributions to the analysis of context. The work of interactional sociologist Ervin Goffman on face-to-face interaction provides an analytical framework for the description and interaction of multi-party interaction (1967, 1974), and the ‘frames’ that can be invoked by a single speaker within talk itself (1974, 1979). Ethnomethodology (Garfinkel 1967) investigates methods used by social actors in interpreting their everyday life: how ‘members’ negotiate or achieve a common context. It argues that context is accomplished
through ‘ongoing, moment-by-moment social and cognitive work; participants display to each other their understanding of the events they are engaged in as part of the process through which these very same events are performed and constituted as social activities’ (Goodwin & Duranti 1992:28). Within this sociological framework conversation analysis emerged, mainly through the work of Sacks, Schegloff and Jefferson (especially Sacks, Schegloff and Jefferson 1974; Schegloff & Sacks 1973; Schegloff, Jefferson & Sacks 1977). Conversation analysis accords primacy to the analysis of talk as a body of situated practices (Atkinson & Heritage 1984). Its central goal is ‘the description and explication of the competences that ordinary speakers use and rely on in participating in intelligible, socially organised interaction’ (Heritage & Atkinson 1984: 1). A basic assumption throughout is Garfinkel’s proposal that producing behavior and understanding and dealing with it are ‘accomplished as the accountable products of common sets of procedures’ (1967:1). Conversation analysis uses as its data naturally occurring interactions and focusses on sequential analysis: whatever is said will be said in some sequential context and its illucutionary force will be determined by reference to what has been said before. According to Heritage (1984) the production of talk is doubly contextual: an utterance not only relies upon the existing context for its production and interpretation, it is also an event in its own right that shapes a new context for the action that will follow it. In this sense, every utterance is both context-shaped and context renewing: context-shaped in the ways that it is designed and understood by reference to the environment of actions in which it participates; context renewing because it contributes to the environing sequence of actions.

Whereas conversation analysis focuses on the structural properties of naturally occurring interaction with less attention paid to what speakers have to say once they have got the floor, discourse analysis investigates the much broader area of the form and function of what is said or written. Like pragmatics and sociolinguistics, discourse analysis has its roots not only in linguistics but also in the social sciences and in philosophy. Discourse analysists argue that discourse is a practical, social and cultural phenomenon. Language users engaging in discourse accomplish social acts and participate in social interaction. Such interaction is in turn embedded in various social and cultural contexts, such as informal gatherings with friends, or professional, institutional encounters such as meetings or lessons. The analysis of discourse as ongoing social action focusses on order and organization: discourse analysis sets as its goal the development of a comprehensive theory of the relations between language use and sociocultural contexts.
Learning in Contexts

It may be concluded that work in a number of fields has shown that context is a socially constituted, interactively sustained, dynamic, and time-bound phenomenon. Humans are capable of reshaping the context that provides organization for their actions within the interaction itself. Context and talk stand in a mutually reflexive relationship to each other: talk shapes context as context shapes talk, or, in other words: people create the world as they claim it to be. What is ‘the’ context is continually under negotiation: it is co-constructed from moment-to moment by the participants in the event themselves.


A non-exhaustive survey of the findings in the research areas and approaches of importance to this study will be presented below.

1.2 Linguistic Anthropology

Linguistic anthropology studies linguistic practices and addresses the interface between language and culture. It considers language a social tool and speaking a cultural practice. Linguistic anthropologists work at producing ethnographically grounded accounts of linguistic structures as used by real people in real time and real space. They examine language through the lenses of anthropological concerns. These include the transmission and reproduction of culture, the relationship between cultural systems and different forms of social organization, and the role of the material conditions of existence in a people’s understanding of the world. Linguistic anthropology is interested in speakers as social actors, in language as both a resource for and a product of social interaction, and in speech communities as simultaneously real and imaginary entities whose boundaries are constantly being reshaped and negotiated through countless acts of speaking. It applies traditional ethnographic methods to the study of everyday speaking and draws on work done in ethnomethodology and sociolinguistics (ethnography of communication; Gumperz & Hymes 1964). In such an approach the unit of analysis cannot be purely linguistic, it has to be a social unit, the communicative event,
Learning in Contexts

which is based on or includes speech.

1.2.1 Participation frameworks

It is my belief that the language that students have drawn on for talking about speaking and hearing is not well-adapted to its purpose ... it takes global folk categories (like speaker and hearer) for granted instead of decomposing them into smaller, analytically coherent elements.

(Goffman 1981: 129)

Taking the communicative event as unit of analysis means the introduction of the category ‘participant’ in the analysis. The notion of participation stresses the inherently socially collective and distributed quality of any act of speaking. Participation as an analytical concept replaces old dichotomies like speaker-hearer or sender-receiver and takes more subtle distinctions into consideration. Philips (1972) introduced the term participant structure in her work on the school performance of American Indian children to refer to the ways in which teachers arranged verbal interactions with their students. Goffman expanded both Philips’ ideas and Hymes’ approach to the notion of ‘participant’ in speech events in his paper ‘Footing’ (1979). By footing Goffman refers to the position or alignment an individual takes in uttering a given linguistic expression. Speakers let the hearer know how an utterance should be taken, the illocutionary force they mean to give it, the scene in which it should be placed, the character it is being said by, to, or on behalf of. Footing refers to the subtle ways in which we assume different types of statuses and stances vis-à-vis our own as well as others’ words. For the total configuration of such statuses at any given time Goffman uses the term participation framework. In place of the ‘speaker’ he posits what he calls the production format. Goffman argues that it is too simple to identify the pronoun ‘I’ with the category ‘speaker’ or ‘writer’. In his model ‘I’ refers to three distinct roles: animator, principal and author. The animator is the entity who is actually speaking, the sounding-box; the principal is the person who is ultimately responsible for what is being said; and the author is the one who selects the words being spoken. These three roles often coincide, but in face-to-face talk as well as in structured interviews and public appearances they may be split apart. A similar reasoning applies to the ‘hearer’. Goffman points out that in a certain situation there may be all kinds of people who ‘hear’ what is being said, but they are not all entitled to be part of the communicative event. Those who are part of the event Goffman named ratified participants, those who are not, unrated participants. In multi-party interactions, moreover, there may be addressed and unaddressed recipients among the ratified participants. Unrated participants may be bystanders: those people who have some kind of access – aural, visual – to the
Learning in Contexts

encounter. They can be overhearers or eavesdroppers: depending on the context they may be expected to pretend they are not present or might make their presence and understanding of the ongoing event quite clear. Goffman goes on to typify talk according to the kinds of participants involved: dominating communication – the main engagement between ratified participants; and subordinate communication – ‘talk that is manned, timed and pitched to constitute a perceivedly limited inference to what might be said to be the dominating communication in its vicinity’ (Goffman 1981: 133). It might be argued that the most important aspect of Goffman’s model has been that it pushes beyond dyadic talk and opens up the possibility of a differentiated approach to multi-party talk.

Goffman acknowledged that there are many forms of talk in which speaker-hearer roles would have to be mapped onto several persons. In order to solve this problem, Levinson (1988) attempted to decompose the concepts of speaker and hearer even further into a set of underlying constituent concepts (‘sets of roles’) and came to some seventeen participant roles, hierarchically organized in a scheme of feature matrices. Irvine (1996) argues that the number of participant roles arrived at by the decompositional approach may prove an endless and unconstrained process, and that the participant roles in Levinson’s model are as decontextualized as in the classic speaker-hearer model. She states that:

The problem of participant roles is part of a larger set of issues: an utterance’s conversational ‘reach’, backward and forward; the interpretative frameworks on which participants draw; the social personae whose voices are echoed, commented upon, or responded to; whether participants acknowledge that they are engaged in a joint conversational activity at all; and so on.

(Irvine 1996: 135)

For these reasons she proposes a simple set of primary participant roles: speaker, addressee and third parties present and absent, while deriving the more subtle types from a notion of intersecting frames and dialogic relations. For this she draws partly on the Bakhtinian notion of multi-vocality which focuses on the forms of discourse that cannot be attributed simply to the act of an individual speaker.

In an effort to go beyond face-to-face interaction and in order to emphasize the emergent nature of participation frameworks, Hanks (1996a) suggests that the notion of participation framework should partly be replaced by community. He proposes Lave & Wenger’s notion of community of practice (1991), which is defined as ‘an aggregate of people coming together around mutual engagement in an endeavor’ (Eckert & McConnell-Ginet 1992: 464, cited in Hanks 1996: 221) as a basis for the study of larger collectivities. Hanks claims that this approach provides a framework in which to describe complex mediated communication and
Learning in Contexts

to define modes of participation relative to the processes through which they are constituted. He states that no single meta-language for participant roles has proved to be adequate at all levels of description and argues that this will be an extremely complicated problem to solve. All that can be done is ‘to cast our descriptions of face-to-face participation and larger-scale discursive formations in such a way that they intersect – or if not, that the points of divergence are made visible’ (1996a: 223).

1.2.2 Indexicality

Deixis in the traditional linguistic sense refers to the fact that certain linguistic forms have a direct pragmatic interpretation depending on the parameters of the speech situation. Deixis and the linguistic forms that subserve it play a central role in language use. Levinson (1983: 54) argued that it is the ‘single most obvious way in which the relationship between language and context is reflected in the structures of languages themselves…’. Indexes are the basic tools that help participants negotiate this issue. If we accept that the repertoire of roles in verbal communication is much greater than the traditional dyad of speaker and hearer, work needs to be done by participants in an interaction to know ‘what’s going on’. Hanks points out that deixis can be seen as a framework for organizing the actor’s access to the context of speech at the moment of utterance: ‘Deictic reference organizes the field of interaction into a foreground upon a background, as figure and ground organize the visual field’ (1996a:61). A basic property of the indexical context of interaction is that it is dynamic: the indexical framework of references changes during an interaction, as participants exchange information, shift topics, establish common grounds, etc. Speaking is a continuous process of contextualization. According to Gumperz this interactional work is performed through a vast range of contextualization cues (see also ‘footing’; Goffman 1979, section 1.2.1; ‘POP- and PUSH markers’; Polanyi & Scha 1983, section 1.4). They help:

... speakers signal and listeners interpret what the activity is, how semantic content is to be understood and how each sentence relates to what precedes or follows ... These features are habitually used and perceived but rarely consciously noted and almost never talked about directly.

(Gumperz 1992: 131)

So contextualization cues are a subclass of indexical signs which let people know what is going on in any given situation and how interaction is expected to proceed. When they are misinterpreted or missed the communication is in trouble. Contextualization cues can operate at various levels of speech production:
syntactic, phonological, morphological, semantic, but may also occur at the levels of prosody, paralinguistic signs, markers of speech tempo, overlaps, laughter, and code-choice (cf. Erickson & Schultz 1981). They can be any verbal or non-verbal sign that helps speakers to hint at, and clarify, and listeners to make inferences about the situation. Contextualization cues enter into the process of inference at several degrees of generality. First, they may operate at the perceptual level, where they serve to provide information on matters like possible turn construction units, foregrounding or backgrounding of items of information, separating shared or known items from new information, etc. The second level is that of sequencing. Inferences at this level yield situated interpretations of communicative intent. They may also be used on the more global level of *framing* (Goffman 1974). On this level they may raise expectations about what is to come at some point beyond the immediate sequence to yield predictions about possible outcomes of an exchange, about suitable topics, and about interpersonal relations. In everyday interactions these levels merge. Participants themselves are concerned with the situated interpretations of what they hear, with the process of arriving at a socially active notion of context (Gumperz 1992).

1.3 Linguistic pragmatics: Politeness theory

Goffman (1967) argued that the fragile and precarious nature of social order is mainly due to the tension between how people present themselves in public, and how others suspect or accept that representation, between ‘official’ identities and hidden selves, between the ‘front stage’ and the ‘back stage’. Individuals, in interactions, are concerned with presenting and maintaining a public image of themselves. Goffman defined this image as ‘face’: ‘the positive social value a person effectively claims for himself by the line others assume he has taken during a particular contact.’ (1967: 5). Following Goffman, Brown & Levinson ([1978]1987), in a seminal study, proposed that maintaining face is a basic motivation in human interaction and has two dimensions. One is *positive face*: a person’s concern that he or she be thought well of by others. At the same time, though, people wish to preserve a certain degree of autonomy and the right not to be imposed on: this is their *negative face*. In Brown & Levinson’s view social interaction is based on a person balancing the satisfaction of their own positive and negative face needs with the face needs of others. The need to balance face needs, according to Brown & Levinson, derives from the fact that most communicative acts are inherently imposing or face threatening. *Politeness strategies* are the means by which interactants fend off and redress these risks to face. The choice of strategy is determined by the configuration of three contextual variables: social distance, relative power and the ‘weight’ of the imposition. Politeness, then, as seen by Brown & Levinson, is ‘the intentional, strategic behavior of an individual
Learning in Contexts

meant to satisfy self and other face wants in case of threat, enacted via positive and negative styles of redress’ (Blum-Kulka 1997).

If we say something that represents a threat to others’ expectations regarding their self-image, we perform a face threatening act (FTA). Alternatively, if we, as speaker in this situation, can say something to lessen the possible threat, it is called a face saving act. When we attempt to save another person’s face, we can pay attention to their negative face (i.e. their need to be independent, to have freedom of action and not to be imposed on by others) or their positive face (i.e. their need to be accepted or liked by others, to be connected to the group). The tendency to use positive politeness forms, emphasizing closeness between speaker and hearer, can be seen as a solidarity strategy; negative politeness forms, emphasizing the hearer’s right to freedom, can be seen as deference strategies. The perceived seriousness of an FTA will determine the selection of redressive strategy. To do an FTA baldly, without redress, implies overwhelming urgency, minimal threat, or vastly superior power on the part of the speaker. The more serious the FTA, the more negative politeness forms will be selected, because the effectiveness of positive politeness depends on vulnerable assumptions about the speaker-hearer relationship. According to Brown & Levinson interactants have a choice of five politeness strategies: 

- **bald on-record** strategies – represent the most direct option (e.g. imperatives: ‘open your books …’), often followed by mitigating devices or hedges like ‘please’ or ‘would you’; 
- **positive politeness** strategies – enhance the positive face needs of the interlocutor (e.g. in-group identity markers, jokes); 
- **negative politeness** strategies – geared to satisfy the hearer’s negative face wants (e.g. giving options); 
- **off-the-record** strategies – leave maximum options for deniability (e.g. non-conventional indirect requests: ‘it’s a bit chilly in here’); 
- **opting out** – saying nothing at all. Brown & Levinson claim that this list represents a scale of politeness: the more indirect the utterance the more polite it is considered to be.

But face systems are not context-free: a speaker’s estimate of power, distance and imposition interacts with other factors in determining the choice of politeness strategies, such as communicative goals, the medium of the interaction and the relation between the interactants. Scollon and Wong-Scollon (1995) extensively discuss the dilemmas arising from the fact that persons belong to a number of voluntary (professional, institutional) and involuntary (cultural, gender, generational) discourse systems at the same time. They argue, for instance, that ‘[as a professional communicator like a teacher] one is simultaneously a member of the discourse system of one’s professional or occupational group and of one’s institutional, organizational, or corporate discourse system’ (1995: 203). These systems may make competing demands for membership and identity, which results in complex and multiple identities. Van Dam van Isselt (1993) contends that politeness
theory needs ‘to be repaired to accommodate multi-person interactions and task-oriented settings, where different notions of what constitutes a face-threatening act may obtain’ (23). Finally, cross-cultural research on politeness seriously challenges the universal nature of politeness models (Kasper 1990). Cultures differ widely in their interactional styles, so that whereas directness is the accepted behavior in a given situation in one culture, indirectness may be the norm in the same situation in another (cf. Scollon & Wong-Scollon 1995; Goddard & Wierzbicka 1997).

1.4 Discourse models

Discourse units have an internal structure that is accessible to study. A number of researchers have been concerned with the development of ‘discourse grammars’. Polanyi & Scha (1983; Polanyi 1988) have proposed a model of discourse structure that allows for the recursive embedding of discourse units. Polanyi & Scha argue that all types of talk, whether it is ‘telling a joke’, ‘ordering a meal’, ‘having a conversation’, etc. have characteristic features and constraints. The assignment of utterances to the category ‘same sort of talk’ as preceding utterances of ‘different sort of talk’ brings Polanyi and Scha to argue that the actual flow of talk should be seen as separable in various discourse units to which individual utterances are seen to belong. Such units, strung together, make up the actual discourses which are produced. These units range from one word ‘echoes’, response cries (Goffman 1979) and non-verbal acts, through minimal clauses and sentences to very large discourse structures. In their model Polanyi & Scha distinguish three main types of discourse units. Type I are small, local structures involving turn-taking, i.e. adjacency pairs (greetings, question-answer). These have traditionally been studied within the framework of conversation analysis. Type II structures are coherent semantic structures built up through sustained talk, such as stories, reports, arguments, etc. Type III structures are socially defined, recognizable occasions for talk, often of a highly focused sort, such as service encounters, meetings, religious rituals, lessons, etc. (‘speech events’; Hymes 1964). These discourse units may be recursively embedded. This means that stories may be embedded in stories, but also that a speech event (e.g. a service encounter) may be embedded in a story (if a person pauses in the middle of a story to buy something). An embedded episode is signalled by so-called PUSH and POP-markers: these may be linguistic, paralinguistic, prosodic or non-verbal. They signal the beginning and end of an embedded constituent to the embedding structure (see also chapter 3).

Van Dam (van Isselt) has shown how Polanyi & Scha’s proposals can profitably be used for the description and interpretation of classroom multi-party situations the ‘lesson proper’ is embedded in (the ‘classroom machine’; 1993, 1995). Van Dam (van Isselt) uses the metaphor of the ‘stack’, consisting of units under construction.
At the bottom of the stack, at the deepest level, there are symbols like ‘speech event’, which define the unmarked context for the interpretation of the events. At the highest level, at the top, there is the unit which is currently being processed. In between, all sorts of units-under-construction (ongoing discourse units) may still be accessible. These units may be oriented to, resumed or overruled. Participants, however, have to recognize these shifts in the discourse to be able to know what is going on. When some symbol is put on top of the stack (‘PUSH’) a new context is created for the interpretation rules of new input: a parameter reset has been effectuated which creates an embedded domain. When the embedding is over we return to the embedding context by removing the symbol on top of the stack (‘POP’). Units that are put on top of the stack inherit their higher order contextual parameters and values. So what is on top of the stack defines current context, but more deeply embedding contexts that are not overruled are still valid. In traditional teacher-fronted lessons, for instance, students do not have the right to self-select. A spontaneous contribution of a student then may be constructed as a ‘non-event’ in the context of such a lesson. A unit, however, may also be processed in an embedding cultural domain in terms of a different type of grammar, e.g. ‘conversation’. In that case particular classroom constraints are suspended: at that moment a more informal ‘current’ situation locally obtains. This analysis entails that the speech event ‘LESSON’ does not contextualize monolithic discourse structures with invariant speaker and hearer roles and ritual business only. Frames can be embedded in or superimposed upon one another. Interactional arrangements, current roles and identities may be shifted into and out of at will.

The main distinction between this approach, based on the Polanyi & Scha proposal, and the numerous other systems for the analysis of classroom interactions which have been developed in the past (see Dunkin & Biddle 1974; Good & Brophy 1984; Cohen & Manion 1985; Chaudron 1988) is that it enables us to analyse classroom interaction with the same tools as interactions in other contexts. Within the dominant context ‘lesson’ different types of conventions (conversational, other) may be temporarily imported, in embeddings or interruptions. Unconventional behavior, unexpected events simply cannot be coded in the traditional models, let alone be analysed and explained: they do not fit into a system that has been specifically designed for classroom interaction, since presuppositions about the conduct which might be expected in this context have been built into the system in the form of pre-coded expectations. Therefore these kinds of observation instruments are unable to describe or analyze these unexpected actions and events – even though they are sometimes of great importance for the course of the lesson —, in the same way that it is impossible for some grammars to trace the meaning of sentences that that grammar has classified as ungrammatical.
1.5 A social perspective on L2 classroom learning

On the whole, SLA theory traditionally has concentrated on:

... trying to elucidate the linguistic 'route' of development followed by the learner and to account for it in terms of the workings of internal psychological mechanisms, whether language specific or more general.


As a result of this narrow, psycholinguistically orientated, construction of language and learning, little research on L2 learning has actually been carried out in the larger social context of real communicative interaction (Van Lier 1988, Nunan 1991). A growing body of studies, however, take a theoretical position which views the language learning process as essentially social, in which interaction itself constitutes the learning process (for surveys of recent investigations, see Hall & Verplaetse 2000, Lantolf 2000). This is not a new view. As early as 1978 Hatch recommended a discourse approach for the analysis of L2 data. However, the introduction of Vygotsky's *sociocultural learning theory* into the domain of SLA in the 1990s has given extra impetus to the view that language learning should be seen in essentially social terms. Socio-cultural theory is predicated upon understanding the importance of interaction between people for the formation of mental activities (Vygotsky 1978). Sociocultural theorists argue that learners are active constructors of their own learning environment, which they shape through their choice of goals and operations (Lantolf 2000). From this perspective on language and learning, classrooms are considered important sites of development.

Sullivan (2000a) adopts the sociocultural perspective in her investigation of the communicative language teaching paradigm and argues that a number of its underlying assumptions (e.g. language learning is done through an exchange of information; the tasks and materials should be 'authentic') and practices (e.g. the prominence of group- and pairwork) essentially reflect Western, Anglo-Saxon cultural values. She argues in favour of a broadening of the definition of communicative language teaching to 'one that can include teacher-led, playful oral narrative styles' (2000a:131), which is more in tune with non-Western culture and value systems. Sullivan claims that this would also benefit students in Western language classes since 'the real-life language of native speakers includes form-focussed and playful language as well as task-based language' (2000b; see also Van Lier 1988: 227).

Van Lier (2000) advocates an *ecological* way of researching language learning (see also Kramsch, in press). This approach links cognitive processes with social
processes: it entails that language and learning should be seen as relationships among learners and between learners and their environment. According to Van Lier an ecological approach to (language) learning:

... shifts the emphasis from scientific reductionism to the notion of emergence. Instead of assuming that every phenomenon can be explained in terms of simpler phenomena or components, it says that at every level of development properties emerge that cannot be reduced to those of prior levels. Second, ecology says that not all of cognition and learning can be explained in terms of processes that go on inside the head. Finally, an ecological approach asserts that the perceptual and social activity of the learner, and particularly the verbal and nonverbal interaction in which the learner engages, are central to an understanding of learning.

(Van Lier 2000: 246)

It may be concluded that, although most SLA research still ‘abstracts away from the ‘noise’ of real life’ (Leather & Van Dam, in preparation), a number of approaches to SLA aim at more empirically valid theoretical models of language learning. To achieve this they address the socially constructed environment that constitutes the learning — the (classroom) setting and the way students interact within it as well as the historical and cultural context of the world outside — and thereby re-assess a number of basic assumptions in traditional SLA research (cf. Leather & Van Dam, in preparation).

1.6 Ethnography of education

In their book *Interpretive ethnography of education* Spindler & Spindler argue that since a few decades anthropologists have ‘come home’:

Today they do research in all places where people gather. Within any social setting, and in any social scene within a setting, whether great or small, social actors are carrying on a culturally constructed dialogue. This dialogue is expressed in behavior, words, symbols, and in the application of cultural knowledge to make instrumental activities and social situations work for one.

(Spindler & Spindler 1987: 2)

One of the results of this shift in research interest is a scientific discipline called the *ethnography of education* or *anthropology of schooling*, which aims to analyse ‘the’ culture of a school or class as a collection of cultural moments. It claims that the educational context is a social context and that social identities are construed in situations. Therefore, it takes as its focal point social dimensions and roles.
Ethnographers of education go into the school and the classroom and let the data lead the way. These should consist of the experiences of insiders. Ethnography of schooling views the school and the classroom as authentic social settings (cf. Breen 1985) and the questions that are asked center on cultural objects. In this way, it is claimed, our cultural blindness will come to the fore: what we do not see or only perceive selectively. Ethnographies of education take many forms: they range from micro-studies of classrooms to studies of students and teachers and of communities and their schooling. They address both the interactive and wider social contexts: the cultural expectations which are structurally inherited in the situation (for an excellent overview of research issues, see Mehan 1998).

1.6.1 Discontinuities between home and school culture: macro-analysis

From the early seventies of the twentieth century onwards, ‘minority’ education has attracted great research interest in the field of the ethnography of education. It became increasingly clear that children from ethnic minority families often face problems because their home culture is radically different from the social mainstream. In the USA in particular, numerous research projects into these problems have been set up, which aim at describing and analyzing how children respond to differences into the socialization processes within their homes and community (Philips 1972; 1983; Heath 1982; Ochs 1982, 1988; Ogbu 1982, 1987; Delgado-Gaitan 1987; Macias 1987). They address discrepancies between home and school cultures minority children often are confronted with and which may result in discontinuities in the enculturation process: ‘... abrupt transition from one mode of being and behaving to another’ (Spindler 1974:308). Discontinuities are attributed to differences in culturally patterned socialization processes between home, community and classroom and sees differences in culture as the source of interaction conflicts.

Philips (1972) investigated the discontinuities between the home and school environments of Warm Spring Indian children and discovered that the participation framework children from this tribe were used to at home differed widely from those American mainstream children experienced in the same setting. Philips found that, at home, the native American children were accustomed to lateral networks of peers. These were much more important than hierarchical role-differentiated networks of adults and children. Philips, therefore, attributed the generally poor school performance of Indian children to the frequency of unfamiliar and threatening participant structures at school.

Ochs (1982, 1988) in her work on language socialization studied mothers with their children in Samoa and discovered that ‘baby talk’ is not universal: the way mothers
Learning in Contexts

communicate with their children relates to local theories and practices of child
trearing, including the social relationship between child and caretakers. This may
result in a conflict of cultural norms for using language and may cause children to
fail to successfully adapt to the educational practices they are immersed in once
they enter school.

Heath (1983) compared the way (Black and White) middle-income teachers talked
to their Black low-income elementary school students in a community she called
Trackton, with the way in which these teachers talked to their own children at
home. She found that the language they used at home with their own children was
very similar to the language they used at school: their children were taught to label
and name objects and to talk about them out of context. In school the teachers used
utterances that were interrogative in form but directive in pragmatic function,
known information questions, and questions that asked for information from books.
The (Black) Trackton parents would use statements or imperatives instead of
questions, and if they asked questions, they would not be the display questions
asked at school. These findings led Heath to the conclusion that Trackton children
were not prepared to cope with the major characteristics of the language used in
classrooms.

The work of John Ogbu addresses the disparities in school performance among
different ethnic groups (1982, 1987). He reframes the discontinuities issue and
raises the question why some groups of minority students in the USA are more
successful in school than others. To explain these differences, Ogbu has developed
a theory which distinguishes between different ethnic minorities on the basis of
their history — the initial terms by which a particular group was incorporated into
the society it now exists in — and the adaptive response the group has made to the
subsequent discriminatory treatment that it has experienced. Ogbu distinguishes
two major categories of minorities: involuntary minorities, who were brought into
their present society through slavery, conquest or colonization, and immigrant
minorities, who have moved into their present society for socio-economic or
political reasons. Ogbu’s central premise is that involuntary minorities experience
more difficulty in dealing with cultural, linguistic and structural barriers in
education than immigrant minorities. Ogbu’s framework has been very influential,
but has its problems. It fails to account for intragroup variability and overlooks
gender and generational influences. Moreover, the situation in the USA and in
other countries which traditionally have been accustomed to receiving immigrants
differs significantly from the situation in countries that have only recently come to
receive large numbers of immigrants (cf. Eldering 1997). Finally, his typology is
argued to be too simplistic because of its dichotomous nature and in danger of
contributing to stereotypical images of ethnic minorities (Gibson 1997, Hermans
1998).
Interactional analysis of classroom discourse coupled with home/school and cross-cultural comparisons of discourse have generated pedagogical strategies to make instruction culturally compatible. Delgado-Gaitán (1987) proposes in her study of the interaction at home and at school of Mexican children in California that students' home culture should be introduced in the classroom to promote a more comfortable and productive learning environment and to ensure the children's full participation in their schooling as 'they learn to accept their traditions while making new social and cultural transitions' (1987: 359). Macías (1987) reports a successful pre-school program in which children from the American Indian tribe of the Papago are taught by teachers from their own tribe. Papago teachers in the program developed their own 'hidden curriculum' geared to minimize the negative effects of cultural discontinuity: they linked unfamiliar classroom content and tasks to the children's home culture to ensure that the child's appreciation of its home culture would not be eroded.

The latter studies both address situations in which the school population is homogeneous: all students belong to the same ethnic group. The 'solutions' proposed – the introduction of home culture in the classroom and the introduction of same ethnic group teachers – are certainly not feasible in settings where the student population consists of more than one ethnic group. Spindler & Spindler suggest cultural therapy as an instrument to explore and deal with cultural differences in those settings. They define cultural therapy as:

... a process of bringing one's own culture in its manifold forms – assumptions, goals, values, beliefs and communicative modes – to a level of awareness that permits one to perceive it as a potential bias in social interaction and in the acquisition and transmission of skills and knowledge. One's culture as well as the 'other's' culture becomes a 'third presence', removed somewhat from the person, so that one's actions can be taken as involuntarily 'caused' by one's culture, rather than uniquely determined by one's personality.

(Spindler 1999: 466)

Cultural therapy may be seen as a 'sensitization' process: it makes the nature of conflicts explicit on cultural terms. Through cultural therapy, Spindler & Spindler claim, the nature of problems is seen in an objectified manner. This makes it possible to 'unpack' potential conflicts, misunderstandings and 'blind spots' and to reach a new perception and interpretation of behavior. Bannink & Van Dam (van Isselt) (1998) propose that, since classroom practices are always situated in particular cultural environments, 'cultural therapy' should play an important role in teacher education. They analyze 'fieldnotes' from a diary written by two students of Dutch secondary education to argue for a reconsideration of the relevant data in
teacher education courses and suggest that student teachers should be made aware of the cultural component of their 'personal theories' of teaching and learning in order to 'unlearn' their ethnocentric, default assumptions in the interpretation of non-verbal and verbal behavior.

1.6.2 Research into classroom culture: micro-analysis

Micro-analyses of classroom talk focus on various aspects of the organization of classroom interaction (early studies: Mc Dermott 1977; Mehan 1979, 1980; Erickson 1981, 1982). In these analyses classroom contexts are seen as authentic, unmarked social situations and as socially active entities constructed by teachers and students. The premise of the studies is that, if we are to understand the nature of teaching, socialization and classroom culture, we need 'a means of systematically describing the actively constructed social action process of the classroom' (Green & Wallat 1981: 193).

Dorr-Bremme (1990) investigated the social organization of a daily group meeting in an American primary school classroom. He found that transitions between class topics were vulnerable to disruptions: order would sometimes break down. A close analysis showed that these disruptions were not unpredictable: they were explainable with reference to subtle contextualization cues that the teacher, unconsciously and routinely, provides but occasionally omits at these boundaries. Students act on the presence or omission of these cues, which thus become 'a tacit, jointly constructed means of discourse regulation and social control' (1990: 379).

Poole (1992) analysed the interaction between teacher and students in two beginning ESL classes in the light of the language socialization perspective articulated by Ochs and Schieffelin (1983). They argue that in American middle class socialization contexts, asymmetrical interactions are consistently marked by accommodation of expert to novice through the use of simplified registers (baby-talk and teacher-talk). Also there is a tendency for caregivers and experts to suppress the display of power differences. Poole found in her data of teacher-student interactions that teachers exercise a good deal of control, but do this in a manner consistent with white middle class American norms, which brings her to state that the teacher role is culturally constrained and motivated: teachers fulfil a culturally prescribed role. Teacher-student interactions mirror cultural dimensions that necessarily affect both the teaching and learning processes.

Van Dam (van Isselt) (1993, 1995) argued that the social dimensions and constraints of multi-person classroom floors are underanalyzed in existing models for the interpretation of L2 classroom events. This results in recommendations
from SLA theory and teaching methodology for teacher and student behaviors that are premature and do not take into account the wider socio-cultural and local interactional contexts of their production in classroom situations.

1.7 Relevance to the study

As mentioned before, a multi-disciplinary framework will be employed in this dissertation to analyze the paradoxes, problems and dilemmas which are a result of the discourse complexity of the task-oriented, multi-person interactional processes that go on in the educational encounters under investigation. Insights from linguistics, anthropology, and sociology will be combined into a discourse framework that attempts to do justice to the interactional, institutional and socio-cultural levels of contexts of teaching and learning. The advantage of this discourse approach over traditional linguistics is that it enables the researcher to deal with multi-party interactions – which lessons by definition are – and makes it possible to include prosodic features of speech (intonation, pitch, volume etc.), laughter, silence and non-verbal behavior in the data.

Within the structural discourse framework employed, linguistic pragmatics offers important insights into the interactional phenomena related to politeness theory. Issues concerning ‘face’ are important in classroom situations, where asymmetrical power relations between teacher and students obtain and where the reason for the presence of the students is their relative incompetence with respect to the teacher. From the ethnography of schooling the notion of the situatedness of learning derives, which draws attention to the cultural and institutional dimensions of teaching and learning. From ethnomethodology the notion of the reflexivity of human activities has been adopted. People do not only re-enact the social structures that exist ‘out there’, but create and maintain the settings they are involved in. Interaction in educational encounters, therefore, is considered to be the creative product of the participants that both reflects and constitutes wider social structures. Work in face-to-face communication, especially Ervin Goffman’s writings on participation structures, contributes in an essential way to the understanding and analysis of multi-party interactions. Finally, discourse analysis itself has provided the conceptual framework for the identification of the different structural units and fields of interaction of the educational encounters under investigation and the world around them, and for the hierarchical ordering of the units within the encounters.