THE INTERSUBJECTIVITY OF STANCE TAKING IN TALK-IN-INTERACTION
Abstract

The following chapters compose a summary of five articles which investigate some of the practices by which interactants engage in stance taking in talk-in-interaction. This summary and the original articles suggest that expressions of stance in talk-in-interaction should be perceived as an intersubjective phenomena in which stances are taken and expressed first of all, on the basis of stances taken by co-participants in prior talk (backward-type intersubjectivity) and, second, on the basis of who the recipients / co-participants are (forward-type intersubjectivity). Second, it is suggested that stance taking is managed in complex ways by the use of various linguistic, interactional and embodied practices.

This work as a whole has been able to explain the use of some recurrent practices and practice combinations that are used as resources for stance taking. In brief, these are 1) the production of denials and claims for insufficient knowledge, and the linguistic NEG + POS pattern used to implement them, 2) the use of some gaze patterns in assessment sequences, 3) the use of category terms and associated category-bound activities in questions and answers in news interviews. In addition to this, this work shows that it is important to consider how stances are constructed on a turn level. The present work has discussed 4) the ‘positioning’ activity in news interviews, in which the interviewer’s turn as a whole builds up a preferred stance or sets up a difficult position for the interviewee and 5) the ‘alignment’ activity, i.e. the way in which the interviewee responds to and aligns with a difficult question.

This work provides a small contribution to a wide, rich and complex phenomenon. Further research is still required in order to understand how interlocutors in various interactional situations use language and their body to take stances.

Keywords: stance taking, talk-in-interaction, everyday conversation, news interviews, intersubjectivity, language, dialogic syntax
To Riitta, Aapo and Isak
Acknowledgements

Every now and then I doubted whether I was ever going to finish this dissertation. The original intention in late 2000 was to examine how people talk about other people in interaction. The aim was to investigate how people as part of their stance taking categorize other people according to their ethnic or gender backgrounds in television news interviews. This proved to be immensely difficult and eventually only one of articles that comprise the dissertation work deals with categorization or identities.

The most difficult time coincided with our return back to Finland from Santa Barbara in 2002. I abandoned my original research topic and began to analyze the video data and the transcripts with no clearly preformulated questions or hypotheses. However, it took a frustrating and a long year before anything interesting rose from the data that I had collected in the United States. During that year, my focus drifted to different directions, which made it very difficult to find any patterned and recurrent practices of stance taking—which of course was the primary aim of this work. Then suddenly—and at the last moment—in August 2003 things began to click. Around that time I made some interesting findings in my news interview data and began to get a general idea of the different ways in which ‘stance’ had been approached in the literature. This helped me to formulate the approach that was most suitable describing the phenomena that I had found by that time, and off I went. It has been constant analysis and writing ever since.

However, it is my firm belief that no one is able to accomplish anything alone. If it was not for my colleagues, my family, and my friends, I probably would not have been able to overcome the difficult times and finish this work. This dissertation work in its whole is an excellent example of collaboration on many different levels. I am deeply indebted to all those people who have knowingly and unknowingly influenced and helped me throughout this project.

It has been a privilege and a joy to work with the positive, encouraging and wonderful people in the Stance project in Oulu. My supervisor, Elise Kärkkäinen, who has simultaneously held the strings of the “Stance-project family” firmly in her hands, supervised several dissertations and been the lovely person she is, deserves a standing ovation for her help throughout this project. In addition to Elise, Tiina Keisanen and Mirka Rauniomaa, my “roommates,” and Maarit Niemelä have provided an inspiring work environment in the Stance project, which has been filled with insightful criticism, supportive words of encouragement, and joy and laughter. My other supervisor, Arja
Piirainen-Marsh, has read my articles and the dissertation manuscript with immense care and given extremely sharp-sighted comments and suggestions. There are no words to express my gratitude for the above people. Your encouragement, support, advice... You are simply the best!

I also want to express my sincere gratitude to the pre-examiners Joanna Thornborrow from the Cardiff University and Mathias Broth from the Stockholm University. Their criticism, suggestions for improvements and words of encouragement have been very important during the final stages of this work.

My interest in doing research gradually rose in the mid-1990s and not least because of Paul McIlvenny and the courses he taught at the University of Oulu at the time. His enthusiasm and eagerness to critically examine and observe the world around us was contagious. I will always remember his warnings about the competitiveness within the world of research. I now understand those warnings better than I did at the time. I am also indebted to Heikki Nyyssönen, who gave me the first possibility in 1998 to work in the Department of English in Oulu as a (post-)graduate student. It is quite clear that without Heikki, and the international research environment to which he introduced me, I would not be here.

The English Department at the University of Oulu has provided a stimulating working environment. I want to thank the whole staff for all these years. Each and every one of you has in one way or the other influenced this work. Especially I want to thank Liisa Bozkurt, Riitta Rajaniemi and Kirsti Rautio for their never-failing help regarding those everyday errands I tend to forget and so frequently do not understand. I also want to thank the people in the Humanities Faculty for their help. My special thanks go to the computer people at the Faculty, who have always helped me with the computer problems I have encountered. I also want to thank the staff in the Department of Finnish, Information Research and Logopedics for making me feel so very welcome in my new work.

The year at the University of California in Santa Barbara in 2001-2002 was a true learning experience. I am deeply grateful to John Du Bois for sharing his knowledge about language, culture and interaction with me. His influence is visible throughout this work. In addition to this, I want to express my sincere gratitude to Sandra Thompson and Gene Lerner. The above people gave me the possibility to stay and attend their courses at UCSB. They supported, helped, provided critical input, and gave the possibility to discuss various aspects of linguistics, interaction and stance. There is no doubt that without you this dissertation would have never been finished. However, I also want absolve them of any responsibility of the use I have made of their help. I also want to thank Lars Linton in the Department of Sociology for his help and support during my stay at UCSB, Dale Kunkel from the Department of Communication at UCSB for giving me the possibility to delve into the department's video archives and supplement my corpus and Erica Biely for copying the tapes for me.

I also want to thank the faculty and the students in Langnet – the Graduate School in Language Studies for providing an invaluable, encouraging, supportive, critical and cheerful academic environment. I especially want to mention Professors Marja-Leena Sorjonen, Auli Hakulinen and Jan-Ola Östman for their kind support, and Jaakko Leino for being an excellent coordinator. Finally, I want to thank the students of Langnet’s sub-
program “The Structure and Use of Language.” They are an amazing bunch of intelligent, funny and joyful people with whom you just like to spend your time and kärsiä ja hävetä.

This project would not have been possible without the funding from the following resources: Langnet and the Ministry of Education, which paid my salary for four years; Helvi Hakulinen-Sipilä and Eino Sipilä Foundation, which gave me a grant with which I was able to start my Ph.D. studies; Academy of Finland, which has funded the Stance project (research project 53671) and my year as a visiting scholar in Santa Barbara; and the Emil Aaltonen Foundation which has also funded the Stance project. I want to thank them for having confidence in this project.

I am very lucky to have several friends with whom I have been able to grow up, spend time, sport, talk about music, culture, literature, movies, sports, and do whatever you do with good friends. My special thanks go to my dear pals Tuomas, Esa, Mika H., Kimmo and Sami for enlivening my life with their presence. I also want to thank Minna and Kaitsu, the Wiren family, David and Sandy, Jonathan and his family; Sari, Steve and Fox, Jenny and Christian, Leelo and Miina, the Kärkkäinen family, Mika T., Mervi, Rea and Sampaa, and Kimmo and Johanna for making our stay in Santa Barbara such a joyous year.

My mother deserves a special mention, for she is responsible for my love for books and my eagerness to continuously ask questions and to seek for answers. Without that background I would not be here. I also want to thank my sister Maija and my brother Gerry for being what they are: dear, generous and sincere persons, and the rest of my family in Kempele and Oulu for providing a small but warm home base in my life. I also want to thank my mother-in-law for the generous help that she has provided whenever she has visited us, and the rest of my family-in-law for making me feel at home in their company and importantly a part of the family.

Finally, I am most grateful to my dear partner and best friend Riitta for her faithful and steadfast support, and her infallible sense of humor with which she has so often saved my days. Without you this would not have been possible. I love you. Aapo and Isak deserve to be nominated the best, most joyful and funny sons a father can dream of. It is your presence that makes all this so much more meaningful. I am so proud of you two. I love you, too. [Aapo ja Isak ansaitsevat nimityksen parhaimmiksi, iloisimmiksi ja hauskimmiksi pojiksi, joita kenelläkään isällä voi olla. Teidän läsnäolonne tekee tämän kaiken niin paljon merkittäväämääksi. Olette minulle rakkaita ja olen todella ylpeä teistä molemmista.]
List of original publications

This dissertation is based on the following original publications, which are referred to in the text by the Roman numerals I-V


Contents

Abstract
Acknowledgements
List of original publications
Contents
1 Introduction ...................................................................................................................15
  1.1 Intersubjectivity of language and social interaction ...............................................18
  1.2 The original papers .................................................................................................20
  1.3 Organization of the study........................................................................................23
2 Database ........................................................................................................................25
3 En route to dialogic syntax and stance taking: The benefits of using Discourse
   Transcription................................................................................................................28
4 Towards a functional-interactional approach to intersubjective stance taking ..........34
  4.1 Conversation analysis and intersubjective stance taking ........................................36
  4.2 Linguistic-interactional approaches to stance taking ................................................47
    4.2.1 The functional linguistic ‘theory of stance’ .....................................................47
    4.2.2 Dialogic syntax: Recycling linguistic elements in interaction ....................53
    4.2.3 Interactional linguistics ................................................................................60
  4.3 From single-speaker linguistic contributions to stance to intersubjective stance
      taking......................................................................................................................62
  4.4 Summary ................................................................................................................70
5 Intersubjective stance taking in news interviews...........................................................72
  5.1 News interviews as institutional interaction............................................................74
  5.2 News interview interaction ....................................................................................77
    5.2.1 The sequential organization of news interview interaction .......................77
    5.2.2 Turn design features of questions ....................................................................79
    5.2.3 Turn design features of answers–an intersubjective view to interviewee
        equivocation .........................................................................................................83
  5.3 Summary ................................................................................................................87
6 Embodiment and stance taking in interaction..............................................................88
  6.1 Gaze, facial expressions and intersubjectivity .........................................................91
  6.2 Relationship between talk and embodiment ........................................................92
1 Introduction

‘Stance’ is commonly understood as the expression of a speaker/writer’s attitude, perspective, point of view, standpoint, opinion or position. This understanding agrees with many dictionary definitions for example the one in the Concise Oxford Dictionary (eighth edition, 1990) that defines ‘stance’ as “a standpoint; an attitude of mind.” It has also passed on to much linguistic and interactional research, in which it has frequently been understood as a subjective attitude or position which rises from the speaker’s internal mental positions, aims and beliefs.

However, the present work approaches ‘stance’ from a different perspective. Whereas the term ‘stance’ is understood to be restricted to the speaker’s subjective stance, the notion stance taking is defined as an activity in which interlocutors display their attitudes, positions and points of view dialogically and intersubjectively in and through interaction. In other words, the focus of this work is on the dialogicality and intersubjectivity of stance taking. In order to make such an approach possible the present work relies on audiovisual recordings of real naturally-occurring talk and detailed transcriptions made of these data. The data used in the present work come from everyday interaction (original article IV) but primarily from news interviews recorded in the United States of America and Britain (original articles I, II, III, and V). As the data in this work show, speakers use a variety of linguistic, interactional and embodied practices to take stances. In addition to this, interlocutors take stances as part of the social actions that they produce in the unfolding interaction. In other words, stance taking is an element of the production of for example assessments, questions and answers, and is organized by these social actions.

The intersubjectivity of stance taking becomes evident in the data in the following four ways. First, speakers take stances on the basis of stances taken in the immediately prior discourse. Second, in news interview interaction in particular, interviewers design particular questions and stances encoded therein, in view of who the recipient is. Third, stance taking is frequently accomplished through linguistic and embodied ‘engagement’ between the interlocutors. Fourth, interlocutors do not necessarily simply agree or disagree with each others’ stances, but display careful convergent or divergent linguistic

---

1 A detailed overview of the database is provided in chapter 2.
‘alignment’ with them. The terms ‘engagement,’ ‘resonance’ and ‘alignment’ are used here to operationalize the notion of intersubjectivity. ‘Engagement’ refers to the way in which interlocutors interconnect in interaction. It can become visible in many ways. One form of engagement is the way in which interlocutors can co-construct a coherent course of action or orient to each other’s actions in relevant and appropriate ways (e.g. by giving an answer to a question). Another practical way for an interactant to engage with a co-interactant is to engage with the linguistic words and structures or even gestures used by her. Consider example (1) below, which has been analyzed by Du Bois (2001c; 2003b).

(1) Deadly Diseases, SBCSAE (11:43)

1 → JOANNE: yet he's still healthy.
2   He reminds me [of my brother].
3 → LENORE: [He's still walking] around,
4   I don't know how healthy he is.

Joanne and Lenore are talking about a third person. In line 3, by saying He’s still walking around, Lenore engages with the words and the syntactic structure of Joanne’s utterance he’s still healthy in line 1. Note that formally these two utterances share an element (he’s still), but are also different with regard to healthy and walking around. However, since these different phrases (adjectival and verbal) share the same syntactic position in an otherwise similar environment, they engage with each other. What is an important element of the linguistic engagement here is that the meaning of the phrase walking around could not be understood to refer to a state of health (i.e. less than healthy) without the immediate context and the fact that it is parallel with healthy. Consequently, what happens here is that the engagement and juxtaposition of healthy and walking around, which is provided for by the same syntactic position and the affinity between these phrases, activates and foregrounds a relationship between them. And since these two phrases engage, the participants (and us analysts) are able to understand the situated meaning of walking around as a state of health. This is what Du Bois calls ‘resonance’ (2003b). He claims that resonance is “a property of relations between elements in discourse, [which] as such cannot be attributed to any element in isolation.” In other words, the first meaning (healthy) has established a background or an environment that contributes to the meaning of walking around. In isolation, the phrase walking around would not receive this same situated meaning.

Whereas ‘engagement’ and ‘resonance’ are terms closely related to Dialogic Syntax (see section 4.2.2), the term ‘alignment’ is here used to refer to the relationships between the speakers’ stances. It does not refer to ‘agreement’ per se, but rather to the range of possible types of positions that speakers can take in relation to each other. For example, in example (1) above, Lenore does not just bluntly agree or disagree with Joanne’s stance (cf. I don’t know how healthy he is in line 4), but indeed suggests that the person they are talking about is less than “healthy” (see section 4.2.1 for a more detailed description of alignment). In sum, Lenore’s stance can be seen to build upon and engage with Joanne’s

---

stance, and align with it. Moreover, in doing so she also creates a situated and uncommon meaning for a particular phrase.

‘Engagement,’ ‘resonance’ and ‘alignment’ are thus specific groups of practices by which intersubjective stance taking can be accomplished. These notions are discussed in greater detail in the following chapters. Although stance has been addressed in various studies on language and talk-in-interaction, there is not much research that considers how stance taking relative to the above practices.

The examination of stance taking undertaken in this study represents a functional-interactional perspective, which combines the findings and methods from conversation analysis, the methods from interactional linguistics, and the discourse-functional work on ‘stance taking’ called the ‘theory of stance’.

In brief, conversation analysis examines the interactional organization of social actions and activities. Interactional linguistics uses the methods of conversation analysis in order to investigate how linguistic elements are used in interaction. Finally, the discourse-functional theory of stance (Du Bois 2002a; Du Bois 2002b; Du Bois 2003d; Du Bois 2004) provides a theoretical framework for understanding the intersubjectivity of stance taking. It is closely connected to an emerging syntactic theory, that of dialogic syntax (Du Bois 2001c), which is here used to make explicit the ways in which interlocutors engage and align with each other when they take stances. These approaches are discussed in detail in the following chapters.

It is naturally rather difficult to establish an unproblematic and coherent amalgamation of the above disciplines, because they have their own aims and agendas and different understandings of the relevant notions. Nevertheless, in spite of the apparent differences, they also share many interests and views that provide important starting points for a study of stance taking. For example, these disciplines are congruent as regards their view on empiricism. Each one of them bases its findings on real naturally-occurring interactions. They also share the view that “sociality” or “culture”—as elusive as these notions may be—are fundamentally enacted in human interaction. Previous studies have rarely considered stance from this perspective, but rather entertained a more decontextualized and speaker-internal view of stance. However, with the help of the afore-mentioned disciplines stance can be viewed as a dialogic and emergent activity in which interactants apply a wide range of interactional and linguistic practices to express their attitudes and perspectives. In other words, stance taking is a phenomenon that is accomplished through and affected by the interactional organization of a speech-exchange system, by the speaker actions, and the grammatical and lexical design of each utterance in its sequential context. The importance of looking at naturally-occurring interaction and the intersubjectivity of interaction are the common denominators for these disciplines. Therefore, they are not only important starting points to the study of stance taking but contribute greatly to the bringing of important new insights into understanding it.

3 Listing these three disciplines here does not mean that it is these only that can have a say about stance. On the contrary, there are various disciplinary perspectives, among those (linguistic) anthropology, philosophy, phenomenology, psychology, and neuroscience (Du Bois 2002a), that can shed light on the notion of stance. These fields are listed here, because they all share an interest in naturally-occurring spoken discourse.
With the help of this combined functional-interactional perspective, I identify linguistic, interactional and embodied practices that are used by interlocutors as resources for stance taking. The empirical analyses of the examples here and in the original papers provide evidence for the fact that such a functional-interactional perspective is indeed fruitful for examining how interlocutors take stances. This work is exploratory and does not aim to provide a clear definition of stance or stance taking. Rather, the aim of this work is to investigate linguistic, interactional and embodied practices, and combinations of them, by which stances are expressed and collaboratively negotiated and constructed in interaction. Next, in section 1.1 I provide some background for understanding the notion of intersubjectivity. In section 1.2 I give a brief outline of the research questions, aims and results of the original papers, and finally in section 1.3 I briefly describe the organization of the present summary.

1.1 Intersubjectivity of language and social interaction

As was mentioned above, a significant body of linguistic research has considered ‘stance’ more or less a subjective single-speaker phenomenon (e.g. Biber & Finegan 1988; Biber & Finegan 1989; Biber et al. 1999; Conrad & Biber 2000; Hunston & Thompson 2000; Martin 2000; Thompson & Hunston 2000). However, as will become apparent, stance taking is also, or even more so, an intersubjective activity and an interactional event that is established and requires the presence of two interlocutors or two subjectivities, in which they define their own subjectivities in relation to their co-participant’s subjectivity (Du Bois 2004). Intersubjectivity is central to all communication and language is the essential tool for accomplishing intersubjectivity. However, the fact that an interactional event, and stance taking more specifically, requires two interlocutors does not mean that the two just babble without considering and recognizing each other’s actions. On the contrary, as conversation analytic work has shown, social interaction follows certain patterned “rules” that interlocutors orient to. As has also been shown (Du Bois 2001c), and as is further supported here, interlocutors dialogically engage with each other’s linguistic structures and phrases, and also do this in order to take stances. Consequently, one specific focus here is on linguistic engagement and recycling through which interactants express and take stances in the sequentially unfolding interaction.

Benveniste (1971) was perhaps the first linguist to note that language is not only fundamentally subjective, but that in ‘discourse’ “every utterance assume[s] a speaker and a hearer, and in the speaker, the intention of influencing the other in some way” (1971, 209). Consequently, as Benveniste (1971, 230) suggested, the relationship between the speaker and hearer is intersubjective, a condition “which alone makes linguistic communication possible.” In addition to Benveniste, Harold Garfinkel’s work in ethnomethodology, which in turn was influenced by the work of Talcott Parsons, considered intersubjectivity as shared agreement and shared operation (Garfinkel 1984).
The notion of ‘dialogism’ discussed by Bakhtin (see Vološinov 1973) is in many respects analogous to ‘intersubjectivity’. The Bakhtinian view of dialogism was a critical response to the Saussurean notion of the unified and whole sign, which did not recognize the multiple and complex relationships that signs (e.g. words) can have with other signs. Indeed, for Bakhtin nothing is in itself a priori, but that existence is the event of co-being and a vast web of interconnections (Holquist 1990, 41; see also Vološinov 1973, 85). Even though dialogism for Bakhtin is very much a philosophical concept, a concept that pervades all aspects of existence, one of the core ideas of dialogism is that it is inseparable from language.

What is even more important for the present work is Bakhtin’s recognition that dialogicality penetrates language in two different ways. First, an utterance is a reflection of what has been said in the immediate social situation (or social milieu), in previous discourse (Vološinov 1973, 86). For Bakhtin the number of prior texts is basically infinite and the interconnections between utterances so vast that understanding or acknowledging all these connections is beyond the grasp of a single individual (Holquist 1990, 41). Second, a written or spoken ‘utterance’ (in Bakhtin’s terms) can be designed so as to anticipate a response from actual, potential or imagined recipients (Holquist 1990, 60; White 2003, 261; Vološinov 1973, 80). Bakhtin in fact says that “the immediate social situation and its immediate social participants determine the ‘occasional’ form and style of an utterance” (Vološinov 1973, 87). To put it differently, the word is oriented toward an addressee, and toward who that addressee might be (Vološinov 1973, 85). An important part of dialogicality is thus that a speaker can “tailor” an utterance specifically for a particular recipient, or, as Holquist (1990, 60) aptly points out, “an utterance is never in itself originary: an utterance is always an answer.” In other words, when a speaker produces an utterance she already takes into account possible responses and answers that the recipient is likely to produce.

Interestingly, this argument has also been discussed, albeit in different terms, both in cognitive approaches to human interaction and in conversation analytic work. Michael Tomasello (1999a; 1999b), who represents the cognitive approach, claims that when humans interact and communicate, they understand other persons as intentional agents and that there are “at least two actual perspectives on a situation” (Tomasello 1999b, 517). Bakhtin’s views also resonate with the central understanding of interaction maintained in conversation analysis, namely that the interlocutor’s talk can be context-shaped or context-renewing (Heritage 1984b, 242). This conversation analytic view is used for describing the sequential organization of interaction and the way in which the speakers’ actions are connected to each other. Conversation analysis has also examined how a turn can be designed so that it is sensitive to who the recipient is (see section 4.2.2) (Sacks et al. 1974, 727).5

The above ideas of intersubjectivity and dialogism are helpful for developing an understanding of stance taking. In simple terms, in the same way as every ‘utterance’ is

---

4 It is a much debated issue whether Mikhail Bakhtin and Valentin Nikolaevič Vološinov were in fact one and the same person. For more information see Dentith (1995) and Holquist (1990).

5 See for example Nuyts (2001), Schiffrin (1998), and Traugott and Dasher (2002) for other views of intersubjectivity.
linked with prior or future ‘texts’ or ‘actions’, a ‘stance’ uttered by a speaker can be seen to arise from a prior utterance or a prior stance. In the original papers I have called this aspect of intersubjectivity *backward-type intersubjectivity* (original article I). This term is here specifically used for describing how the stance “I now take” has been occasioned by a stance or stances taken in the immediately prior talk. In addition to this, a ‘stance’ can be understood to be specifically designed for a real or imagined recipient. This I have called *forward-type intersubjectivity* (original article I). Although these notions are closely connected to the ideas of intersubjectivity described above, they are here specifically used for describing the intersubjectivity of stance taking. A good example of the distinction between, for example, a context-shaping action and the forward-type intersubjectivity can be given from the ‘denial’ paper (original paper III). The interviewer’s question shapes the interactional context so that an answer is the next relevant action. However, as we can see in that paper, the way in which questions may set up a position for the interviewee to take a stance (forward-type intersubjectivity) impacts the design of the answer, so that the interviewee uses the denial + account action combination in order to invalidate the position built up in the question.

As relates to backward-type intersubjectivity there is an analytically significant dilemma of what might comprise a relevant “prior text” for stance taking. A single stance, similarly to Bakhtin’s dialogic utterance, can be seen to exist and materialize in many forms in interaction, and to basically originate from countless prior stances or, more broadly, socio-cultural beliefs. However, to start from an assumption that stance is ever-pervasive can be disconcerting when one intends to study how interactants engage in taking stances in interaction (see also original article I). Consequently, for analytical purposes, the boundary for what to include and what to exclude from analysis has to be set somewhere. As will be outlined in more detail in the following chapters, this study describes stance taking as it takes place in talk-in-interaction, and similarly to much conversation analytic and interactional linguistic research, considers the “prior text” to be those aspects that the interlocutors can be seen to orient to in situ. In other words, the evidence for the analytic claims made has to come from the data.

In the following section I briefly discuss the research questions, aims and the findings reported in the original papers.

### 1.2 The original papers

Throughout this summary, the previous work on stance is supplemented and cast in a new light on the basis of the findings reported in the five original articles (I-V) that comprise the essence of the whole dissertation work. The original articles concentrate on the following aspects of stance taking.

The ‘theory’ paper (I) considers the problems of understanding the notion of ‘stance’. It suggests that ‘stance’ is indeed a widely-used notion in linguistic, interactional and anthropological research, and it is approached and described in ways that are problematic for the present work. For example, the majority of the linguistic work has claimed that stance is the expression of the speaker’s subjective attitudes and beliefs. In view of this, it
is difficult to decide the level of detail on which to concentrate, because it often seems like “stance is everywhere” and “everything is stance.” In general, the ‘theory’ paper (I) suggests that one of the most fruitful and interesting ways to examine stance taking is to use the above-mentioned functional-interactional approach. This involves collecting and analyzing examples of routinized and recurrent linguistic and interactional resources that speakers use for stance taking. I further suggest that in order to look at how co-participants construct and display their stances, an analysis of the simultaneously deployed linguistic resources and sequential aspects of turn design and turn construction is required. Finally, I briefly consider an example of an intersubjective stance-taking activity in news interviews I have called *positioning* / *alignment*.

The ‘positioning’ paper (II) provides more detailed analyses of the positioning and alignment activities in news interviews. The paper suggests that although conversation analytic research generally assumes that interaction is intersubjective, its work on news interview data has concentrated primarily on actions and practices produced in one single turn or even one turn-constructional unit. These include practices by which interviewers design adversarial stances and maintain neutralistic stances and practices by which interviewees evade questions, shift the agendas of the questions or reformulate the question (e.g. Clayman & Heritage 2002). Therefore, conversation analytic research has given little attention to what the interviewers’ and interviewees’ turns together do in terms of stance taking, i.e. how the interviewee’s stance aligns with the stance in the interviewer’s question and how their turns engage with each other. The ‘positioning’ paper (II) examines in greater detail two stance-taking activities in news interviews, namely ‘positioning’ and ‘alignment’. In the paper I focus on how some linguistic practices (e.g. interrogative syntax), together with the ways in which the question turn is designed, encode and evoke preferred stances and presuppositions in the questions, i.e. how the questioning turn as a whole *sets up positions* for interviewees to take a stance. I also investigate how interviewees *align with* these positions by using the denial + account action combination.

The ‘denial’ paper (III) uses the methodology provided by the functional-interactional approach and concentrates in greater detail on two action combinations that act as resources for the interviewee’s stance taking in news interviews. The first action combination which is called ‘denial + account’ denies a presumption / presupposition in, or the adequacy of, the interviewer’s question and then provides an account for the denial. The second action combination, ‘claim for insufficient knowledge + account’, is used by the interviewee to first claim imperfect knowledge about the topical matter and then to provide an account. The ‘denial’ paper (III) concentrates on describing a recurrent linguistic practice through which these action combinations are produced by the interviewees. This is called the *NEG + POS* pattern, in which the interviewee uses epistemic stance markers and engages with the question by recycling a linguistic element from the interviewer’s turn. Previous research on news interviews has considered these types of interviewee answers evasive or otherwise violative of the projected trajectory of talk. However, the data show that denials and claims of insufficient knowledge engage strongly with the question. The engagement is especially manifest in the way in which the interviewee recycles a central linguistic element from the interviewer’s questioning turn. These two action combinations thus reflect the intersubjective relationship between the
question and the answer, and should not just be considered expressions of interviewee evasiveness. These frequent action combinations play an important role in organizing the interviewees’ stance taking in British and American news interviews.

The ‘gaze’ paper (IV) takes a somewhat different perspective compared to the first three papers. It considers how interactants in everyday interaction use different gaze patterns during the stance-taking activity. Although embodiment in general has received a fair amount of attention in interactional research, the relationship between embodied practices and stance taking has not been discussed in detail. The ‘gaze’ paper (IV) sheds light on this relationship by investigating the ways in which interactants use the concurrent organizations of assessments and particular gaze patterns as resources for stance taking in everyday conversation. It discusses stance taking in relation to three gaze patterns: congruent gaze point, mutual gaze and cut-off gaze. The paper provides further evidence of the intersubjectivity of stance taking and suggests that the above-mentioned gaze patterns play an important role in stance taking. The analyses show that although the interrelationship between gaze and assessments is manifold, certain gaze patterns are intertwined with the making of assessments, and therefore gaze and assessments can be seen together to function as resources for interactional and intersubjective stance taking.

In brief, the congruent gaze point pattern can be used for marking the Stance Object, the mutual gaze co-occurs with the production of a second agreeing assessment, and the cut-off gaze occurs at moments of divergent alignment and can act as a resource for the co-participants for tracing the meaning of a co-participant’s stance. However, it is not claimed that these gaze patterns have meanings in themselves or that they would implicate a speaker stance, but rather that gaze is an important element in the interactants’ joint and intersubjective constitution of stance.

Finally, in the ‘identity’ paper (V) I examine how co-participants in news interviews use membership categories as resources for constructing stances and negotiating identities. I draw on the above-mentioned functional-interactional approach and look at two examples from Crossfire (CNN). I show, on the one hand, how identity work becomes manifest in the interlocutors’ actions and turn-taking and, on the other, how ethnic identities for non-present parties are negotiated and constituted through linguistic engagement and alignment in this particular episode of Crossfire.

As becomes evident from the above the original papers have somewhat different aims and foci. This reflects the process of development in the overall research project. At the outset, my intention was to investigate how people talk about ethnic and racial minorities as an integral part of their stance-taking activity in news interviews. The ‘identity’ paper (V) reports a case study of such an instance. However, gathering a sufficient collection of such instances and examining the related recurrent linguistic practices would have required an enormous corpus. The original papers I, II and III, however, form a coherent body of work. They describe some practices by which interviewers and interviewees in news interviews engage and align with each other when they take stances. Finally, the ‘gaze’ paper (IV) stands out among the papers because it uses different data, everyday conversation. At the outset of the project the investigation of embodied practices of stance taking was considered very important. However, it proved to be difficult to examine embodiment in news interview data, because both or all the interlocutors are rarely visible in the television frame simultaneously. It became almost impossible to
examine how a particular gaze pattern or embodied practice is related to a co-participant’s talk or vice versa and, therefore, embodiment and stance taking were investigated in everyday talk.

Although the somewhat different foci and data in the original papers could be seen as a complication in terms of the coherence and depth of this work, they nevertheless are able to provide evidence of stance taking as emergent and intersubjective and not subjective activity. In addition to this, the papers also show that engagement and alignment are central elements of stance taking. Therefore, in this respect the original papers as a whole provide a good basis for future research. Chapter 6, which discusses embodiment and stance taking and chapter 7, which discusses categories and stance taking, especially suggest new directions for future research. In addition to this, chapter 7 takes a somewhat critical view to the ‘identity’ paper (V) and addresses some limitations of that paper. In this way, this work also demonstrates development in the understanding of stance taking and what the research on stance taking should concentrate on.

The following chapters provide a detailed and critical review of previous research on stance. Along the way the previous work is supplemented and cast in a new light with the help of the empirical findings in the original papers. The review of previous work describes traditions that have overlooked the intersubjectivity of stance, most importantly traditional linguistics, but also some work in social psychology. These investigations have mostly considered how single lexical items express speaker stances. However, despite the critical tone, the aim is to consider these approaches in a positive light, as studies that have taken important steps toward the understanding of stance and stance taking in discourse. The following chapters also give a detailed overview of those research traditions that provide important contributions to the understanding of stance taking as an intersubjective activity: conversation analysis, functional linguistics, interactional linguistics, and some work within embodiment research. This body of work provides the backbone for the present work both theoretically and methodologically.

In sum, the summary has the following aims. First, it aims to pull together the strings of previous work and provide an understanding of the resources, practices and actions by which interlocutors display stances and display an understanding of each other’s stances; i.e. how they manage their stance taking. Second, this summary discusses, compares, critically evaluates and connects the prior research on stance and stance-related phenomena with the findings in the original papers. Third, it casts a critical light on some findings in the original papers and thereby aims to show development in the way in which stance taking is understood. Fourth, this summary provides insights into and suggestions of new directions that stance research can pursue.

In the following section I briefly outline the way in which this summary is organized and how it helps to form an understanding of stance taking as an intersubjective activity.

### 1.3 Organization of the study

The following chapters are organized so that each chapter lays out a foundation for subsequent chapters, which then eventually lead up to a description and a summary of
how stance taking is understood and approached in this work. Consequently, already the database (chapter 2) and the transcription system (chapter 3) bear an important role in the analysis of stance taking. Chapter 4 first gives an outline of previous linguistic work on stance and claims that this body of work has considered stance as a subjective phenomenon. The latter part of chapter 4 then provides a brief outline of the work done in ethnomethodology, conversation analysis, functional linguistics and interactional linguistics, whose theoretical and methodological principles and findings have been used as a resource in this work. Chapter 4 also briefly discusses discursive social psychology and suggests that it provides an apt methodology and an excellent source of prior research for further research on stance taking. Chapter 5 gives an account of previous linguistic and interactional (mostly conversation analytic) work that has focused on the language and talk in news interviews. The aim of chapter 5 is to draw parallels between previous work on news interviews and the present study (cf. original articles I, II, III, and V). Then in chapter 6 I give a brief account of previous work on gestures and embodiment in interaction (cf. original article IV). Chapter 7 then takes a somewhat critical standpoint towards the ‘identity’ paper (V) and develops and casts the work in it in a new light. Then, in the conclusion in chapter 8, I pull the strings together and provide a general understanding of stance taking and discuss the role of the different approaches used in the present work. Throughout all chapters I connect previous work on stance or stance-related phenomena with the work reported in the original articles and provide evidence for the importance of considering stance taking as an intersubjective activity in which linguistic, interactional and embodied practices all play an important role.
2 Database

The data in this summary and in the original articles (I-IV) come from different sources. The ‘gaze’ paper (IV) stands out as an exception among the original articles, since it examines and describes phenomena in everyday interactions rather than news interviews. Two recordings were used to describe the phenomena. The first is called Game Night and it is owned by Cecilia Ford, who kindly gave permission to use the transcripts made of the recording for two articles (Haddington 2002a and the original article IV) and include actual examples in them. Game Night is a tape of approximately 20 minutes that has friends gathering at one of the interlocutors’ home to play Pictionary. The interlocutors talk about various everyday issues, such as family and art. The other tape is called Never in Canada (Haddington & Kärkkäinen 2003). It is a 60-minute tape which was recorded in Oulu in February 2003 as a data collection and transcription exercise in an undergraduate class on conversation analysis. The recording has three native speakers of English, who were visiting the University of Oulu as exchange students, talking about various everyday issues, such as doing laundry, attending classes and so on. Never in Canada was recorded with two cameras, which proved to be an excellent decision for the investigation of the co-participants’ embodied actions. In section 4.3.2 I discuss an example which comes from an audio recording called Burp. It was recorded by me in the United States in autumn 2001. It is four hours long, of which approximately 20 minutes have been transcribed. In the recording a family has gathered for dinner and they discuss various everyday matters, such as work, school, friends and so on.

All the other original articles (I-III and V) investigate stance taking in news interview data. The data for these articles were recorded in the United States and in Britain between October 1999 and March 2004. The corpus contains 38 separate programs and approximately 50 interviews. The duration of the individual interviews spans from a couple of minutes to a full hour, which amounts to a total of approximately 20 hours. In the United States most of the interviews were recorded from cable channels such as CNN (Larry King Live, Crossfire, Late Edition, etc.), FOX (O’Reilly Factor, Fox News) and

---

6 I want to thank Prof. Dale Kunkel from the Department of Communication at UCSB, my colleague Tiina Keisanen in Oulu, and my brother Gerry Haddington, for their help in gathering supplementary data for the corpus.
PBS ("Newshour with Jim Lehrer"), and nationwide channels such as ABC ("This Week, Nightline"), CBS ("Face the Nation") and NBC ("Meet the Press"). Some of this material contains live interviews on the day of and the days after the terrorist attacks in New York, Washington D.C. and Pennsylvania on September 11, 2001. The news interviews from 9-11 and the following days are not part of any particular program, but, due to the character of the events, just part of the general newsflow on different channels. The data from British television contain episodes of "HardTalk with Tim Sebastian" (BBC World), which is broadcast globally, and "Newsnight" (BBC 2). The interviews in the database were first recorded in VHS format, after which they were digitized, compressed and copied to several hard drives. After this the programs or selections of them have been transcribed.

The interviews in the database were selected with the following two criteria in mind. First, the interviewees in them are either political figures or other “public” figures who possess expert knowledge on particular issues (e.g. scientists). These people appear in the programs as advocates of particular lines of thought or as representatives of certain institutions with particular public viewpoints. The rationale behind this selection criterion is that political figures and other experts are invited to the programs specifically to answer questions regarding their viewpoints and opinions and to defend them. In other words, they are invited to take stances. Consequently, the following types of interviews are excluded from this work:

1) witness accounts and newsmaker interviews;
2) talk shows (e.g. CBS’s "Late Show with David Letterman" or The Oprah Winfrey Show) because these revolve around the guest’s personality, rather than a topic;
3) hybrid interviews—e.g. a combination of news interview and radio call-in genre in which viewers participate in the interview through phone-calls (cf. Clayman & Heritage 2002, 8);
4) sound bites and news items.8

Second, since news interviews are for the most part unedited and unscripted, the interactional situation is contingent on what the co-participants do and say. Even though there is a pre-conceived agenda at the outset of a program and the interviewer has prepared questions beforehand, the unfolding of interaction in itself is not predetermined, but emergent (Clayman & Heritage 2002, 6).

There is also a significant benefit of using news interview data which relates to the Observer’s Paradox (Clayman & Gill 2004; Duranti 1997; Labov 1972 xvii-xviii). The

---

7 See for example Linton (2001) for a study on how late night show hosts construct candidate answers for their guests in entertainment interviews.

8 Sound bites and news items provide an interesting research topic if one wants to consider how interviewee answers and comments are first decontextualized (pulled out of context) and then recontextualized (inserted in another context) (cf. Ekström 2001). Ekström (ibid.) examines these types of data and shows how journalists can change the context of interviewee answers by editing the answers and re-inserting them in particular contexts. As Ekström (ibid.) claims, this naturally has the potential to affect the public opinion. However, due to the focus here on stance taking as an interactive and sequential phenomenon between the participants in news interviews, these kinds of data have been excluded.
Observer’s Paradox refers to the problem that a field worker cannot be sure if the process of observing affected the observed situation, or if the interactional moment being recorded would have proceeded in the same way in case it had not been recorded. With news interview data the Observer’s Paradox dilemma is basically non-existent, because news interviews are already designed to be viewed by an outside audience. Whether or not someone views, records, and analyses an interview does not have any impact on how the actual interactional situation unfolds.

The use of the term *news interviews* here is slightly different from that in conversation analytic literature (cf. Clayman 1988; Clayman & Heritage 2002; Heritage 1985; Heritage & Greatham 1991). Conversation analytic studies have shown among other things that *news interviews* have a constrained turn-taking order in which interviewers are primarily responsible for asking questions and interviewees for answering these questions (Clayman & Heritage 2002). They have also shown that interviewers adhere to a principle of maintaining a neutralistic stance in the course of questioning (e.g. CNN’s *Larry King Live*) (Clayman & Heritage 2002). The conversation analytic studies claim that these characteristics, among other things, constitute the *institutionality of news interviews*. However, as is discussed by Hester and Francis (2001), this extrapolation is problematic. In some interview programs the “special” turn-taking organization and the principle of interviewer neutralism are breached; in some programs only occasionally (e.g. *HardTalk with Tim Sebastian* on BBC World or *Newsnight* on BBC 2), whereas in others the program format is based on an openly confrontational debating ambience (e.g. CNN’s *Crossfire*). *Crossfire* has received some attention in discourse analysis (Firth 2003) and is also mentioned by Clayman and Heritage (2002, 336, 340) as an example of a program that *resembles* news interviews. In spite of the fact that there are differences between the above interview programs, the unifying character of the corpus is that all the programs belong to the genre of televised interview programs and the findings are valid in this specific context. Furthermore, it seems that based on the increasing number of debate programs and news interviews and perhaps also the shift in the latter towards a more hostile interviewing style, one might ask whether the genre of televised interviews in general is undergoing a gradual change toward a more confrontational interviewing style. This is quite possible, as articles (I-III) suggest.

---

9 Both of these concepts are discussed in greater detail in the following chapters.
10 For a brief discussion on this, see Clayman and Heritage (2002, 30).
3 En route to dialogic syntax and stance taking: The benefits of using Discourse Transcription

Doing transcriptions and using them as helping tools in the analysis of naturally-occurring talk are both theory-laden practices. The data in this project have been transcribed by using the Discourse Transcription (henceforth DT) notation system, which has been developed by Du Bois et al. (Du Bois 1991; Du Bois et al. 1992; Du Bois et al. 1993). The rationale for using the DT transcription system (and not for example the system used in conversation analysis) is that together with dialogic syntax (see section 4.3.2) it helps foreground and make explicit the evidence for linguistic engagement, recycling and alignment between interlocutors in talk-in-interaction, and therefore it helps support the claim that stance taking is intersubjective.11

Choosing and preferring one way of “writing interaction” over another of course means that the analyst acknowledges certain theoretical principles (what is transcribed and how is it transcribed) (Bucholtz 2000; Coates & Thornborrow 1999; Du Bois 1991; Ochs 1979). However, this condition of theoreticity is not necessarily a disadvantage, but there are also benefits from it (cf. Duranti 1997, 114). Some aspects of the DT system are discussed below and the full explanation of the transcription conventions can be found in the Appendix.

The DT system owes a lot to the transcription system developed by Gail Jefferson, which is also sometimes called the conversation analytic notation system (Atkinson & Heritage 1984b; Sacks et al. 1974; Schenkein 1978). There are indeed many important similarities between the two in the theoretical, notational and practical methods of preparing transcriptions of talk. However, there are also some fundamental differences, which reflect the differences in the theoretical and analytical goals between conversation analysis and the work that uses the DT system. The main difference is that whereas the former’s primary unit is the turn, in DT it is the intonation unit (also sometimes referred to as tone group, tone unit, or intonational phrase, see Kärkkäinen 2003; Park 2002, 637).12 One line in DT transcriptions represents one intonation unit.13

---

11 This does not mean that the Jeffersonian transcription used in conversation analysis could not be used for investigating these phenomena, but that with the DT system it is easier to identify them.
12 See also Cruttenden (1986, 35-63) and Matsumoto (2003).
Although intonation units are prosodic units, they are advantageous as a practical and theoretical starting point for the functional analysis of spoken language in three “non-prosodic” ways as well. First, intonation unit has usually been characterized as a cognitive unit, which is important in terms of information flow. Second, intonation units are syntactically relevant, and third, they are interactionally relevant. As Kärkkäinen (2003, 29-30) claims, practices and actions in interaction are sensitive to intonation units and are thus related to the emergent, incremental and directional nature of turns (see also Park (2002)). Even though these advantages are presented individually below, they are by no means mutually exclusive in any way. As I show, they are elements of language use and interaction that are interconnected and arise from each other—and therefore motivate the use of the DT transcription system.

As was mentioned above, intonation unit is a prosodic segment of speech in that it bears a single intonation contour (e.g. Chafe 1987, 22; Du Bois et al. 1993, 47). Intonation units are also frequently marked by features which act as cues for identifying boundaries between them. Some of these cues are prosodic: changes in tempo (accelerated uttering of unstressed syllables at the beginning of an intonation unit, i.e. anacrusis; prosodic lengthening at the end of an intonation unit) and a rise in the baseline pitch level between two units, i.e. pitch reset. Some of these cues are not prosodic: pauses, different uses of the vocal tract (inbreaths, creaks), and particular interactional cues (turn beginnings, turn endings) (Chafe 1998; Du Bois et al. 1992; Du Bois et al. 1993). However, intonation units and the process of identifying and transcribing them is not devoid of problems. These problems are primarily related to the fact that even though prosodic features can to a certain extent be measured, they are still influenced by subjective perceptions. Consequently, identifying intonation units and intonation unit boundaries is difficult, and different analysts can have variable interpretations of them. In most cases this is most likely due to the lack of a sufficient number of available boundary cues.

13 There are several reasons why conversation analysis does not use intonation units as the primary unit in its transcriptions. In several lectures and discussions with conversation analysts I have come across the following ones: First, it is impossible to require knowledge of intonation units in transcribing. Second, marking pauses and silences in the DT system is problematic. This is because quite often DT system marks a pause on the same line just before the following talk. Marking pause in such a way can be interpreted to mean that the pause “belongs” to the speaker who takes the next turn although in principle anyone could have selected the turn. This problem can be avoided by marking pauses on separate lines. Third, transcribing talk into intonation units is already analysis of talk and some conversation analysts suggest that this must be avoided. Finally, some conversation analysts do not want take a stand on the theoretical presuppositions connected to intonation units (for example, the notion of information flow and its relation to cognition). However, there is an interesting correlation between the notions ‘intonation unit’ and ‘turn-constructional unit’ (TCU), which is one of the central terms in conversation analysis (see below). For example, Park (2002, 641) claims that the intonation unit is a crucial unit in interaction, upon which turn constructional units are based. This question is also discussed in Ford, Fox and Thompson (Ford et al. 1996), Ford and Thompson (Ford & Thompson 1996) and Selting (Selting 1996). Although intonation units and TCUs have different theoretical backgrounds, there are striking similarities between them.
Another pertinent problem is the marking of transitional continuity, i.e. whether an intonation unit ends with final or continuing intonation. Du Bois et al. (1992, 28) say that transitional continuity is a functional category, i.e.

when a speaker arrives at the end of an intonation unit, poised to continue on to the next—or not continue—the intonation contour usually gives a fairly clear indication of whether the discourse business at hand will be continued, or has finished. This is “transitional continuity”: the marking of the degree of continuity which occurs at the transition point between one intonation unit and the next.

This claim suggests that even though transitional continuity is a functional category, the prosody provides an indication of whether the intonation unit is final or continuing. This question whether an intonation unit is produced with continuing or final intonation is the point where intonation units interface with the interactional turn-taking organization (Ford et al. 1996; Ford & Thompson 1996; Park 2002; Selting 1996; Selting 2000). If a speaker designs an intonation unit as final, it is likely that a co-participant understands that the current turn is coming to an end, and that after it has come to a completion, it is possible to take the next turn. In other words, as Park (2002) claims, “turn completion is to a large extent projected by prosody.” Traditionally two specific pitch contours have been claimed to signal turn-finality in everyday conversation: the fall-to-low and the high rise. Many others have been considered to signal incompleteness and doing turn-holding (see for example Chafe 1987; Du Bois et al. 1992; Du Bois et al. 1993). However, Szczepan Reed (2004) suggests that there are several other pitch contours that can be seen to signal turn-finality and that occur at turn-endings, of which she identifies ‘the step-up’, ‘level pitch’, ‘rise-to-mid’, and ‘the musical interval’. She (2004, 113) states that

these findings reveal the notion of terminal pitch movement as indicating turn-finality or turn-holding to be questionable. If turn-final intonation occurs in such a variety of forms, is it really oriented to by participants as an indicator for turn-taking opportunities?

Although Szczepan Reed (2004) does not consider much of the interactional context (e.g. what action an immediately prior turn is doing and what it projects), her analysis is very interesting. It suggests several new questions regarding the role of prosody and intonation in interaction. It can also be seen to open up a rarely addressed question of how different combinations of (embodied, syntactic and prosodic) practices signal turn-finality and turn-continuity.14 This type of analysis can provide interesting new findings of the different “rules” that are at play in turn-taking.

---

14 This idea is also central in Ford, Fox and Thompson (1996). They claim that there are several cues (pragmatic, syntactic and prosodic) that simultaneously indicate and signal turn ending.
The above questions are pertinent with respect to talk in the media. In media talk\textsuperscript{15} it seems that quite often the function (whether the speaker could stop at the end of the intonation unit or not) is not supported by the prosody (finality or continuity) and vice versa. To begin with, an interviewer turn often ends with low rising intonation—not enough to suggest appeal intonation—or even level pitch, thus suggesting (in the traditional view) continuity. However, in many cases these intonation contours occur in turn-final position, and thus together with syntactic and turn-construction-related cues (i.e. the interviewer has asked a question) indicate that the interviewer is finishing her turn. It is then quite possible that it is these other cues and the activity the participants are engaged in, and not only the transitional continuity or the pitch movement, that help co-participants to plan their subsequent actions on-line (cf. Szcepek Reed 2004, 114). Similarly, interviewees seem to be aware that a fall-to-low pitch at the end of an interviewer’s intonation unit which is not doing questioning, does not indicate a transition place. These differences in transitional continuity between everyday talk and news interviews are likely to be due to the different turn-taking “rules” and turn allocation in news interviews (see section 5.3). Therefore, it is quite possible that in news interviews there is a different or alternative set of pitch movements (compared to everyday conversation) which together with syntactic cues signal turn finality.\textsuperscript{16}

The above remarks naturally affect and complicate the transcriber’s decision between marking final or continuous transitional continuity. This only proves the well-known fact that analyzing the way people actually talk deflects peremptory truths, and when it comes to language and language use, we sometimes need to acquiesce to this fact. In this study, in the very robust cases of transitional continuity in which the traditional view of the function of certain prosodic contours and the overall function of the unit agree, transitional continuity is marked accordingly. The ambiguous cases are judged, as far as it is possible, on the basis of the pitch movement at the end of the intonation unit.

At any rate, the advantages of using intonation units and the DT transcription system surpass the problems in identifying those units. In fact, many of the findings made in this work would not have been possible with any other notational system. Consider intonation units, for example. Even though the name—intonation unit—already suggests a close connection between the units of speech and intonational/prosodic/phonetic features, as was already suggested above, there are also other “non-prosodic” advantages of using them.

‘Information flow’ refers to the connection between the ways in which speakers talk and what their mental and cognitive states are (Chafe 1987; Chafe 1993; Chafe 1998; Cumming & Ono 1997). One important aspect of this connection is “the movement of ideas into and out of consciousness of speakers and hearers” (Chafe 1993, 33), i.e. how speakers communicate thoughts to other people. Chafe (1987; 1998) has shown how speakers segment spoken language into intonation units. According to him (ibid.), one intonation unit expresses a single focus of the speaker’s consciousness. These foci can of

\textsuperscript{15} Cf. with Szcepek Reed’s (2004, 111-113) analysis of ‘the musical interval’ in the openings and closings of a phone-in program.

\textsuperscript{16} The above questions are of course extremely interesting, but unfortunately cannot be discussed here further.
course be of various kinds, such as objects, events, and properties, which Chafe uniformly calls “concepts” (Chafe 1979). What is important about these concepts is that they can have a different activation status at a particular time in the speaker’s consciousness, namely active, semi-active, or inactive (Chafe 1987, 48-49). Roughly speaking, active refers to a concept that is “given,” it is in the speaker’s mind at a particular moment; semi-active refers to a concept that is somehow schematically connected to the active concept and can quite easily be recalled to the active status; and inactive is a “new” concept, not in the speaker’s mind at the moment of talking. There are certain constraints on how these concepts can be invoked and used by speakers. This is where Chafe (1998, 100) makes the important connection between the phonetic features of intonation units and their semantic content. In brief, one intonation unit includes one new idea that has a new semantic meaning (Chafe 1987; Chafe 1993).17

The above ideas regarding information flow—that speakers communicate thoughts to their interactants and to a significant degree do this through intonation units—are also closely related to the ‘syntactic relevance’ of intonation units. In fact, by using the intonation unit as the basic unit for transcribing talk, it is possible to perceive in a detailed manner the sophistication of the ways in which interlocutors recycle each others’ morpho-syntactic (lexical, phrasal, clausal) and prosodic structures (see section 4.3.2 for a further discussion) and how these structures engage with each other. It seems therefore that the ways in which speakers communicate thoughts to their co-interactants, and the ways in which they recycle each others’ talk, are closely related to each other. Although repetition or recycling of different linguistic structures has been noted for example by Linell and Marková (1993), Marková (1997), Johnstone (2002), and Anward (2004) systematic discussion of the relationship between ‘information flow’ and these phenomena is still pending. It thus remains a task for future research to unravel.

One of the most recent treatments of syntax, engagement and dialogicality is Du Bois’s (2001c) research on “dialogic syntax” (see also Anward (2004)).18 Du Bois (2001c) describes and formulates dialogic syntax by connecting it, on the one hand, to formal questions of syntax and, on the other, to issues of semantics and culture. As we will see in section 4.2.2, the DT transcription system and the fact that it uses intonation units as its basic unit have been central to the development of dialogic syntax. Furthermore, dialogic syntax provides both a theoretical framework and a concrete method of representation for the analysis of the intersubjectivity of stance taking in talk-in-interaction. In fact, it was only after Du Bois (p.c.) noticed how speakers use syntax dialogically that he discovered that interlocutors simultaneously in fact engage in

17 There are of course very different kinds of intonation units, some very short, comprising only for example hesitation markers, such as “uhm.” The principle of one new idea in one intonation unit applies to so-called substantive intonation units, which usually comprise around four or five words (Chafe 1993, 39; Chafe 1994, 69; Park 2002, 639).
18 Du Bois, in addition to being one of the developers of the DT transcription system, has been involved in a wide spectrum of work ranging from anthropology (Sacapultec Maya) to syntactic issues (preferred argument structure) (Du Bois 1987; Du Bois 2003a; Du Bois et al. 2003b), collecting corpora (Du Bois et al. 2000; Du Bois et al. 2003a), and the relationship between discourse and grammar (Du Bois 1985; Du Bois 2001b; Du Bois 2003c).
continuous and fine-tuned stance taking (see section 4.2.1). Moreover, interlocutors do not take stances only by using the most obvious linguistic forms, but rather engage with what other speakers have said in the immediately prior talk (see also Anward (2004)).

Finally, the DT system has been designed so that it maximizes the readability and usability of the transcriptions, both during the transcription process and the actual analysis. Consequently, the majority of the symbols used in the transcriptions are easy to recognize. In many cases they are also self-explanatory due to their familiarity from written conventions; for example the marking of transitional continuity (notation within brackets): (--) for a truncated intonation unit (cf. a truncated sentence in novels), (.) for final intonation (cf. end of a sentence in writing), (,) for continuing intonation (cf. sentence continues in writing), and (?) for appeal intonation (cf. marking of questions in writing). Also the notation (...) for pause is widely used for example in novels for marking some kind of break in talk. Many of the symbols can be used iconically or are themselves iconic: square brackets [ ] are used to mark the exact positions where overlap occurs, (@), which resembles a ‘smiley face’, is used for transcribing laughter, and so on. I have consistently used the Courier New font in my transcriptions. Although it is not an aesthetically beautiful typeface, it is the most practical one in one major respect: since it is a monospace font, the process of aligning overlapping talk in the transcriptions is much easier than with proportional fonts such as Arial and Times.

In sum, the DT transcription system is highly beneficial for studying stance taking as an intersubjective activity. As is shown in the original articles and the following chapters, the theoretical background that it provides helps the analyst to foreground especially those linguistic phenomena which are endemic of the intersubjective and dialogic relationship between co-participants in interaction. Thus the DT transcription is a useful system for describing the intersubjectivity of stance taking in talk-in-interaction.

As was mentioned above, ‘stance’ and stance-related phenomena have received a fair amount of attention in linguistics and sociology. In the next chapter I describe these approaches and discuss their relationship to the present work.
4 Towards a functional-interactional approach to intersubjective stance taking

As has been suggested above, it is fruitful to consider ‘stance’ in relation to the notions of ‘subjectivity’ and ‘intersubjectivity’ and to tease out the difference between a subjective stance and intersubjective stance taking. Single pieces of work on stance-related phenomena in discourse can be placed on a ‘subjective-intersubjective’ continuum, in which at one end a stance is considered an expression of a speaker’s subjective and inner attitudes, whereas at the other end, stances are collaboratively constructed, and managed in interaction on the basis of other speakers’ actions and stances.

Although it is impossible to assign discrete categories for the stance-related research on this continuum, some generalizations can be made on the basis of their different starting points. At the ‘subjective’ end of the continuum, the starting point tends to be a linguistic or syntactic unit, which is seen to be expressive of a speaker stance. The analysis of these single-speaker contributions to stance, which can also be considered “purely” linguistic and linguistic-pragmatic approaches, is however insufficient for investigating the intersubjectivity of stance. Consequently, in this chapter I provide some theoretical and methodological preliminaries for analyzing intersubjective stance taking, which have been used in the original papers and in this summary. I suggest that the combination of conversation analytic work, the functional linguistic ‘theory of stance’ and interactional linguistics, which I here define as a functional-interactional approach, is fruitful for examining stance taking. This is because for conversation analysis, for the ‘theory of stance,’ and for interactional linguistics, the intersubjectivity of stance taking does not come as a surprise. The research undertaken in these academic fields looks at a vast amount of interactional practices (including linguistic items, prosody, and embodiment) and how these contribute to the accomplishment of particular actions and activities (cf. Schegloff 1997). I start by considering conversation analysis in section 4.1. It has contributed greatly to the understanding of social and sequential organization of talk-in-interaction. The reason why it is used in this work is that it provides helpful tools for considering how stances are expressed through particular interactional practices, actions and activities, and also how stance can be a matter of turn construction or turn design. I begin section 4.1 by giving a brief overview of the historical development of the ethnomethodological work that subsequently led to the rise of the conversation analytic
enterprise. After this I provide a more detailed account of the findings made in conversation analysis. In the same section I also briefly discuss the work undertaken in discursive psychology, whose contribution to the present work is sorely missing in the original articles (I-V). It is suggested that although discursive psychology is not used in this work, it can be used to further our understanding of stance taking in future research. Discursive psychology is discussed in the context of conversation analysis, because it uses the methods and tools of this approach. In section 4.2 I discuss three approaches that are central for the present work. First (in section 4.2.1), I discuss the functional linguistic ‘theory of stance’ (Du Bois 2002a; Du Bois 2002b; Du Bois 2003d; Du Bois 2004), which provides some of the central theoretical and methodological tools for the present work. One of the most important tools is dialogic syntax, which is discussed in section 4.2.2. Dialogic syntax is a theory and a research practice by which linguistic engagement, alignment and recycling can be examined in detail. Then in section 4.2.3 I discuss interactional linguistics, which has effectively combined the aims of conversation analysis and functional linguistics, and has looked at the use of particular linguistic and syntactic practices in their sequential contexts in interaction. This work takes into consideration the contingent and emergent nature of interaction and is most clearly used in the ‘denial’ paper (III). In section 4.3 I consider previous linguistic work that has concentrated on the single-speaker contributions to stance and problematize its approach in light of the theoretical and methodological background provided in sections 4.1-4.2 and with the help of some examples. Finally, in section 4.4 I give a brief outline of the functional-interactional approach used in this work and give some reasons why I have coined a new term for this approach rather than using the old ones.

Even though the above two approaches—single-speaker contributions to stance and the functional-interactional approach to intersubjective stance taking—are strongly juxtaposed in this chapter, they should not be considered mutually exclusive but supportive of each other in important ways. Moreover, there is no clear boundary between these categories but rather the research done within them overlaps in many respects. Nevertheless, this division of research on stance into the two above sub-categories has certain advantages to it. For one, it provides a platform for those interested in stance taking in discourse, but who come from very different backgrounds, to have a glance at what others with an interest in similar phenomena have done. Second, it indeed mirrors the (often implicit) assumptions of the origins of the stances speakers take. The linguistic work (see section 4.3) seems to presuppose that stances originate from the speaker, i.e. that the stances that the speakers take are inherent and subjective expressions of how the speaker thinks and feels. However, the more interactionally oriented stance research that is outlined here pays attention to the practices of stance taking in their sequential context, as part of the actions and activities interlocutors engage in, and importantly engage in with each other. In other words, the interactionally oriented body of work is interested in stance taking as an intersubjective activity.
4.1 Conversation analysis and intersubjective stance taking

As has been mentioned above, the present work applies conversation analytic methods for investigating stance taking. Conversation analysis is central for this work, because it focuses on the practices, actions and activities that interlocutors produce as part of the sequentially unfolding interaction. Moreover, it focuses on how interlocutors orient to their co-participants’ behavior and modify their own actions on the basis of that. By concentrating on how interlocutors orient to their co-participants’ behavior when they produce actions, conversation analysis also assumes that interaction is inherently intersubjective.

The rise and development of conversation analysis was a response to the sociological research of the 1960s (Sacks et al. 1974), and on the other hand, to language philosophical work on speech acts (Searle 1971b; Searle 1977). The founders of conversation analysis felt that most of sociology assumed a viewpoint that “members of society do not ‘really’ know what’s going on around them” (Hutchby & Wooffitt 1998, 33) and that its research relied too heavily on intuition at the expense of a more observational view of what really happens in the world. Conversation analysis has been influenced by and draws on both Erving Goffman’s research on the interaction order and Harold Garfinkel’s ethnomethodological program (Heritage 2003b: 2; Hutchby & Wooffitt 1998: 4). Goffman’s interests remained in the ritual properties of interaction order, i.e. issues of face, role and identity (Heritage 2003b, 3; Hutchby & Wooffitt 1998, 28), and therefore differs from conversation analysis in two important aspects: it does not start from the sequential order of talk-in-interaction and it does not systematically rely on collections of real recorded data (Hutchby & Wooffitt 1998, 29). Garfinkel’s ethnomethodological approach, on the other hand, was related to Goffman’s work, but was still distinctly different from it. He studied the common-sense knowledge or the “rule-governed” behavior that members of society rely on in order to make sense of and act in their everyday lives (Garfinkel 1972, 309; Heritage 1984b, 4). Similarly to conversation analysis, he admitted that language plays a crucial part in this tacit sense-making process. The central claims that Garfinkel made and which are also centrally important for understanding stance taking are that individuals orient to an intersubjectively constructed experience and that members in society share and jointly construct the cultural and social world. The procedures that members apply in particular contexts function as a basis for understanding actions, but also as resources for the production of actions (Heritage 1995, 393). However, in spite of the intersubjective experience and the fact that members of society assume identical perspectives and experiences of the world, they can never actually share them completely (cf. Heritage 1984b, 55; Scannell 1991, 5).

In addition to the above claims, the ethnomethodological work contributes to the present work on stance taking by first, acknowledging that language, which people use to maintain and negotiate the practices they engage in, is inherently indexical. As Garfinkel (see Hilbert 2004) has noted, linguistic indexicality plays an important role in interaction. This is because the background knowledge and the indexical linguistic items that people rely on in order to refer to or use that knowledge, are central elements of their routinized
practices. Moreover, the indexical elements are context-dependent and become understandable only in situ. This becomes very clear in the ‘identity’ paper (V), in which linguistic engagement and resonance between linguistic indexicals foregrounds an ethnic category and renders it understandable in a dubious sense. Second, the ethnomethodological work shows that the various practices that interactants engage in when they communicate and interact with each other are actively oriented to and organized in situ, in those moments in which they occur.

This second aspect of interaction has been further discussed in conversation analysis, which emerged in the late 1960s as a sociological approach to the study of talk and its role for accomplishing social order in interaction. Conversation analysis has indeed been strongly influenced by and draws on both Erving Goffman’s research on interaction order and Harold Garfinkel’s ethnomethodological program (Heritage 2003b, 2; Hutchby & Wooffitt 1998, 4). The beginning and foundation of conversation analysis is usually connected with the work of Harvey Sacks, in collaboration with Emanuel A. Schegloff and Gail Jefferson (Sacks 1992a; Sacks 1992b; Sacks et al. 1974; Schegloff 1968). Harvey Sacks was both a student of Goffman and had worked with Garfinkel, so the connection between conversation analysis and Goffman’s and Garfinkel’s work is not a surprise (see Schegloff 1988).

Conversation analysis provides rigorous methods and important findings which can be used to describe how the various interactional organizations of talk impinge on stance taking. It is a qualitative research method, although its quantitative side becomes evident in the way conversation analysts compile collections of interactional phenomena. Whereas traditional linguistic research on stance (as we will see in section 4.3) has generally speaking overlooked the issue of intersubjectivity, in conversation analysis social actions and activities are considered to be the basic building blocks of intersubjectivity. As Heritage (1984b, 259) claims

\[...\] conversational interaction is structured by an organization of action which is implemented on a turn-by-turn basis. By means of this organization a context of publicly displayed and continuously up-dated intersubjective understandings is systematically sustained. It is through this ‘turn-by-turn’ character of talk that the participants display their understandings of ‘the state of the talk’ for one another.

In other words, conversation analysis maintains that no instances of talk should be looked at in isolation (cf. the discussion of intersubjectivity in section 1.1). Rather, the shared understandings between interactants can be traced by interactants and analysts alike by considering the sequential context and what a speaker does with an utterance (cf. Schegloff 1996b). Moreover, conversation analysis also acknowledges that intersubjectivity is basically “two-directional.” In other words, current talk is seen to be shaped by the immediately preceding turn and to display some (intended or unintended) understanding of an action in the previous turn. Hutchy and Wooffit (1998, 15) call this
the next-turn proof procedure.\footnote{19} Participants indeed normally address themselves to preceding talk and, most commonly, to the immediately preceding talk (Sacks, 1987, 1992; Schegloff, 1984; Schegloff and Sacks, 1973). So in this simple and direct sense, interlocutors’ talk can also be perceived to be context-shaped (Heritage 1997: 162). However, in conversation analytic work, also the intersubjective relationship between interlocutors from “this point onward”—the context-shaping quality—has been acknowledged. By doing some current action, participants normally project (empirically) and require (normatively) that some ‘next action’ (or one of a range of possible ‘next actions’) should be done by a subsequent participant (Schegloff, 1972).\footnote{20} Talk can thus be seen to also create, maintain, or renew a context for the next interlocutor (e.g. Heritage 1997, 162; Schegloff 1972) or make relevant (Schegloff 1968) an appropriate next action—whether the action is there or “noticeably missing” (ibid.).

The conversation analytic understanding of talk as context-shaped and context-renewing resonate with the notions ‘backward-type intersubjectivity’ and ‘forward-type intersubjectivity’ respectively. However, it is important to bear in mind that the conversation analytic notions refer to how actions are related to each other in the sequentially unfolding interaction, whereas the two notions used in the present work are specifically intended to refer to the activity of stance taking, i.e. what stances are indexed, expressed and negotiated through the use of different practices. In addition to this, in conversation analysis one can concentrate on any action or turn, examine its relation to the interactional context, and not necessarily consider how they are related to stance taking.

Interlocutors can also design a turn specifically for a particular recipient. In conversation analysis this is called recipient design and it refers to the multitude of respects in which the talk by a party in a conversation is constructed and designed in ways which display an orientation and sensitivity to the particular other(s) who are the co-participants (Sacks et al. 1974, 727).

‘Recipient design’ is also related to forward-type intersubjectivity. However, whereas ‘recipient design’ indeed refers to the construction of talk that takes into account the recipient, the forward-type intersubjectivity refers to the idea how a stance which is encoded, evoked or presupposed in a turn takes into account the intended recipient.

\footnote{19} The next-turn proof procedure is not only a tool for co-participants in interaction to understand each others’ actions. It is also a tool for analysts to unravel the understandings that the co-participants display by their actions.

\footnote{20} Schegloff and Sacks (1973) also use the term sequential implicativeness to refer to context-shaping quality of action. “By ‘sequential implicativeness’ is meant that an utterance projects for the sequentially following turn(s) the relevance of a determinate range of occurrences (be they utterance types, activities, speaker selections, etc.). It thus has sequentially organized implications” (Schegloff & Sacks 1973, 296, fn. 6).
Conversation analysis is distinct from Goffman’s and Garfinkel’s work in two major respects. First of all, conversation analysis has taken its own specific direction in sociology by examining the interactional organization of human conduct in talk-in-interaction. Second, in contrast to Goffman’s ethnographic methods and Garfinkel’s experiments, conversation analysis relies on actual recorded instances of talk-in-interaction together with detailed transcriptions made of the recordings (including words, non-verbal sounds, pauses, overlapping talk and sometimes also non-verbal embodied practices). The transcriptions do not only recover the minuteness of the interactional organization for a disciplined analysis, but they also provide others a possibility to check the validity of the findings (Sacks 1984). Conversation analysis maintains that data should not be used to seek support for theories, but rather that they should be studied with no preoccupations in mind (Sacks 1984, 27). The use of real spoken data is also a central part of the present work. Stance and especially stance taking have rarely been studied with the help of real spoken data.

The central aims of conversation analysis are, first, to provide a coherent analytic framework for examining and describing the interactional and sequential organization of human conduct in talk-in-interaction. Second, conversation analysis also examines the ways in which talk-in-interaction is systematically and sequentially organized. This includes looking at how the interactants’ practices, in their situated context, become understandable for the interactants themselves, and how they co-ordinate their actions in relation to each other. Indeed, conversation analytic studies have shown that interactants share various interactional practices of turn-taking, sequencing and repair that organize and structure the course of interaction. Interestingly, one can find affinities between conversation analysis and one aspect of how Bakhtin approaches the study of language. Bakhtin (Vološinov 1973, 95) suggests that the study of language should start from “the forms and types of verbal interaction in connection with their concrete conditions.”


The term ‘talk-in-interaction’ comes from Schegloff (1999a) who uses it to differentiate between ‘conversation’ and ‘talk-in-interaction.’ In short, ‘talk-in-interaction’ refers to all kinds of oral discourse (including news interviews and other forms of institutional talk), whereas ‘conversation’ refers to a particular kind of talk-in-interaction, that of everyday or ordinary conversation.

This is an interesting question, since during the last thirty years a considerable body of conversation analytic work has already accumulated and added to our knowledge of talk-in-interaction. This work is therefore bound to affect how scholars approach interactional data. In addition to this, for example the above list of previous work on stance-related phenomena in itself partly represents work that not only functions as a relevant background, but also affects, both theoretically and methodologically, the way in which the current topic is approached in this work. Therefore, Sacks's original idea although worth aspiring to, is very difficult to carry out in practice.

However, see Du Bois (2002a; 2002b; 2003d; 2004). Du Bois’s work is described in section 4.2.1.
Moreover, he (Vološinov 1973, 86-87) claims that an utterance is not only socially oriented but also determined by the participants of a specific situation in a speech event. Furthermore, it is the situation that “shapes the utterance, dictating that it sound one way and not another – like a demand or request. (Vološinov 1973, 86-87)” Even though it is difficult to determine what exactly Bakhtin meant by this, these claims resonate strongly with conversation analysis that concentrates on the import of speaker actions in their sequential context in interaction.

Conversation analysis has productively concentrated on several types of phenomena that are seen to significantly organize interaction. The most basic and important interactional organizations are 1) turn-taking organization, 2) repair organization, 3) sequence organization, 4) preference organization, and 5) turn design.

Turn-taking organization (Sacks et al. 1974) refers to the basic rules of how interactants take and allocate turns in interaction (see also for example Drew 2005; Heritage 1984b; Heritage 1995; Hutchby & Wooffitt 1998; Lerner 2003). Speakers’ turns are composed of units called turn constructional units (TCUs). TCU is the basic unit out of which speakers set out to construct talk. One TCU can constitute a recognizably complete turn and can characteristically be lexical, phrasal, sentential or clausal. These linguistic characteristics of turn construction provide for the “projectability of a turn.” This refers to the possibility of the co-participants to realize and understand what is under way and to project the possible completion point of the TCU (cf. Sacks et al. 1974). In a similar way, Ford (2004, 28) suggests that projectability refers to “the ability for a current moment of unfolding talk to provide for anticipation of possible and even likely aspects of talk yet to come.” In everyday conversation one speaker is allocated one TCU at a time with a series of options, after which the right for who will take the next run has to be “negotiated” (Sacks et al. 1974, 706). In everyday conversation, i.e. the types of conversations in which all of us engage with our family members, friends and so on, this negotiation is usually accomplished through such procedures as “current speaker selects the next speaker” or “next speaker self-selects the next turn”. In other words, everyday conversation proceeds on a turn-by-turn basis and the different features of turn-taking and turn organization are not fixed in advance but accomplished by the interlocutors as they talk. However, different rules apply in institutional talk. For example, in news interviews questions and answers tend to be longer than one TCU. This is because the interviewers often provide relevant background for the question they are going to ask and the interviewees are expected to support their answers with longer rather than short turns. In fact, minimal answers (e.g. one-TCU answers) are generally treated as inadequate in news interviews and tend to be followed by requests for more information.

The difference between the turn-taking organizations in everyday talk and institutional talk provides interlocutors with different resources for taking stances and engaging with stances taken in prior talk. As Sacks et al. (1974, 730) claim, the speech-exchange type

---

25 ‘Possible completion point’ is indeed a possible completion point. In other words, it is not equal to an actual completion of a turn but rather talk can go past a possible completion. Thus, ‘possible completion point’ refers to the possible places in a speaker’s turn in which the possible next speaker can self-select the next turn. The possible next speaker thus needs to analyze and monitor the on-going talk for the possible places where she can assume the next turn. (Schegloff 2001)
and the turn-allocation system constituting it affects the turn-size of individual speakers: the more restricted the turn-allocation system, the longer the turns tend to be. This greatly affects the activity of stance taking in different speech-exchange systems. For example, the simple fact that in institutional talk (cf. news interviews) interlocutors are entitled and even expected to produce multi-TCU turns enables them to construct complex stances that engage with a prior stance in different ways. For example, the particular turn allocation system and the possibility to construct multi-TCU turns in news interviews provide the interviewees for the possibility to produce the linguistic NEG + POS pattern by which the denial + account and the claim for insufficient knowledge + account action combinations are produced (original article III). This linguistic pattern in these action combinations is basically absent in everyday talk. The special features of institutional talk are further discussed in section 5.1.

Repair organization refers to the ways in which interactants deal with some type of interactional trouble or a need to produce a different action from what the talk prior to the repair was on its way to produce (Hutchby & Wooffitt 1998, 59; Jefferson 1974; Jefferson 1987; Schegloff 1979; Schegloff 1992c; Schegloff 1997; Schegloff 2000; Schegloff et al. 1977). In other words, repairs “replace potentially inappropriate talk with material that is more appropriate or suitable to the immediate social situation” (Clayman & Heritage 2002). Repairs can characteristically be either self-repairs or repairs initiated by others.

Sequence organization refers to the ways in which interactants “line up” actions as meaningful sequences of actions or other type of patterned conduct. One such sequence of actions is the question-answer adjacency pair, in which the question action strongly projects an answer, and indeed the second speaker is expected to give an answer to the question (Sacks 1992b, 521-569). In other words, since the first item (e.g. question) has been produced it makes the second (e.g. the answer) conditionally relevant (i.e. expectable) (Schegloff 1968, 1083). However, this does not mean that the second must occur (because interactants are still free to do what they want), but that the absence of a second can be observed, i.e. it is “officially absent” (Schegloff 1968, 1083). In the present work, the news interview data is fundamentally organized into sequences of questions and answers (see original papers I, II, III and V). The interviewers and interviewees orient to this sequence organization and rarely depart from it. Therefore, this particular sequence organization also organizes the host’s and the guest’s stance taking in news interviews.

Another good example of a strong sequence organization is the organization of assessments into first and second assessments (Goodwin & Goodwin 1987; Goodwin & Goodwin 1992; Pomerantz 1984). Pomerantz (1984) shows that the production of assessments in talk is a highly constrained and organized sequential pattern. She claims that the action of producing an assessment is a routinized way of partaking in a social activity, in which a speaker first invokes a referent of which she has direct knowledge and then produces an assessment of it. Assessments tend to be sequentially organized into first assessments and second assessments. A first assessment prefaces and engenders a next action by a co-participant, namely a second assessment, in which she reports either an agreeing or a disagreeing stance about the same referent. The way in which the first assessment is designed invites either a preferred or a dispreferred second assessment (Pomerantz 1984, 63). Heritage and Raymond (2005) claim that in the production of
assessments, interlocutors also display their epistemic rights to make those assessments by designing them in different ways. In sum, co-interactants use different practices to produce assessment actions, which are then frequently patterned as sequences of assessments (first assessment, second assessment). Assessment sequences are highly organized and display the interlocutors’ rights to assess the objects they are talking about. The conversation analytic work on assessments is very important for the present work, because it shows that certain practices and actions themselves involve taking a position or taking a stance.

The ‘gaze’ paper (original article IV) supplements the previous work on assessments by investigating some recurrent gaze patterns and how they are intertwined with and organized relative to the different kinds of assessment sequences.

Preference organization, which is an element of sequence organization, refers to ways in which certain actions are designed to “prefer” particular kinds of responses or next actions (Lerner 1996a; Sacks 1987 [1973]). For example, invitations prefer acceptances, and indeed declinations to invitations tend to be designed as dispreferred next actions (with hesitation markers, delays, etc.). As Sacks (1987 [1973]) suggests, people tend to design, for example, their questions so that they exhibit a preference for a particular answer and render a dispreferred answer difficult to produce. Preference organization is an important feature of stance taking primarily because of its closeness to the idea of “forward-type intersubjectivity” (see section 1.1). However, similarly to the “projectability” discussed above, ‘preference’ as understood in conversation analysis is not related to a speaker-specific ‘stance’, but rather to the structural features of interaction. So in contrast to the notion “structural preference,” I have decided to use the term “forward-type intersubjectivity” to refer to the ways in which a speaker stance (e.g. an interviewer stance in news interviews) is designed and how it invites a particular stance in response. ‘Preference’ and ‘forward-type intersubjectivity’ are discussed in greater detail in the ‘theory’ paper (I) and the ‘positioning’ paper (II). In those papers I discuss how ‘preference organization’, as it is understood in conversation analysis, contributes to the stance encoded or evoked in the interviewer’s question, and therefore contributes to the ‘positioning’ activity that it does.

Finally, turn design refers to two things. First, it refers to speaker’s selection of the action that the turn is designed to be doing. Second, it refers to the practices (phonetic, lexical, grammatical, and embodied) of which the turn has been composed (Drew 2005, 82-86). These two things about the design of the turn together contribute to the turn’s role and its meaning-making potential (e.g. a stance) in a particular position in the sequentially unfolding interaction. In other words, by designing a turn to be doing a particular action, the speaker orients to the immediately previous interaction and displays her understanding of it. However, by designing the turn in a particular way, the speaker also constructs a context for the upcoming talk. Turn design is an important element of the findings reported in the ‘positioning’ paper (II) and the ‘denial’ paper (III). In the ‘positioning’ paper (II) it is claimed that interviewer’s turn as a whole sets up a position for the interviewee to take a stance, and also the interviewee’s turn as a whole that aligns with the position set up in the question. In other words, it is the way in which the turn as a whole is designed that contributes to these stance-taking activities.
In order to understand the interlocutors’ stances in the interactional context we need to understand how they relate to the action that is being produced. ‘Action’ is perhaps the most central and fundamental notion in conversation analysis and it is closely connected with turn design. ‘Action’ has been understood and described in numerous different ways, the most famous in linguistics being perhaps Searle’s speech act theory (Searle 1971a; Searle 1971b; Searle 1977; Searle et al. 1980). However, as Schegloff (1996a, 167) notes, these accounts overlooked the speakers’ own understanding of what they or their co-participants are actually doing in interaction. In contrast to this, Schegloff (1996a, 172) claims,

 [...] there must be a grounding of this formulation [of what action or actions are being accomplished] in the “reality” of the participants. Here the investigator undertakes to establish that the formulation is not an academically analytic imposition on conduct that may have been quite differently understood and experienced by the participants. This requires some demonstration that the interlocutors in the data being examined have understood the utterances (or other conduct) in question to be possibly doing the proposed action(s) or that they are oriented to that possibility—a demonstration ordinarily grounded in the interlocutors' subsequent talk or conduct [...]. This immediately subsequent talk, being appropriate to—or even responsive to—what preceded it, ordinarily displays an understanding of what that preceding talk was “doing.”

Therefore, in conversation analysis, an action is not a hypothetical, theoretical or analyst-oriented notion (Schegloff 1996a, 172). Rather, it is understood in terms of what the interactants themselves understand their own and their co-participants’ actions to be performing—inviting, questioning, answering, blaming, agreeing, disagreeing, and so on. As Drew (2005, 74) suggests, “these and other such activities are the primary forms of social action, as real, concrete, consequential, and as fundamental as any other form of conduct.”

As was mentioned above, actions are implemented by “particular orderly, routinized and patterned organizations of practices” (Schegloff 1997, 499-500, italics added), which in turn can be linguistic or embodied, or for example the repetition of a linguistic form. The different practices that interlocutors use are the devices by which they construct talk. Finally, the relationship between the notions ‘action’ and ‘activity’ is here understood in a similar way as Heritage and Sorjonen (Heritage & Sorjonen 1994), Kärkkäinen (Kärkkäinen 2003) and Drew (2005, fn 6) outline it: an action is performed at utterance or turn level, whereas an activity is “relatively sustained topically coherent and/or goal-coherent course of action” (Heritage & Sorjonen 1994, 4), and which consists of a series of connected actions produced across a sequence of turns. This definition is here augmented with an understanding of activity as a series of connected TCUs in a multi-unit turn (for example in news interview talk). This understanding of ‘activity’ is important for understanding stance taking, because especially with news interview talk, it is often difficult to describe the individual TCUs in terms of the action they are performing. Rather, what often seems to be more relevant is a larger sequence of talk and what it does, for example in terms of stance taking.
Conversation analysis has been criticized for setting the boundaries for what is relevant analysis and what is not, too much within the boundaries of the conversation and utterance, and thus for not taking into account larger macro-social or contextual factors (such as gender and ethnicity) that may be consequential for interaction (Wetherell 1998). These charges have probably risen from the fact that the notion of ‘context’ embodies a vast array of meanings in different disciplines. Many disciplines such as sociolinguistics indeed consider broad macro-contextual or societal issues such as gender and race in their analysis,26 and it is true that in this sense conversation analysis assumes a context-free method of analysis (Hutchby & Wooffitt 1998, 35). This question of where to set the boundaries of context or prior text in interactional analysis is also pivotal for understanding stance taking. It was already addressed in section 1.1 together with Bakhtin and his view of the two-dimensionality of dialogicality. Indeed, it was noted that for Bakhtin a discoursal moment can have an endless number of prior texts. However, this is problematic for conversation analysis. One reason why traditional conversation analysis does not start from a macro-social perspective is that particular interactional practices are seen to recur freely across different kinds of interactional and social situations no matter what the larger circumstances may be. In other words, the same form can be take different meanings in different contexts (cf. Drew & Heritage 1992a).

In addition to this, sociality is seen to be dynamically constructed in interaction in and through the practices in which humans engage with each other when they talk to each other. Therefore, in order to be able to disentangle these practices, the analysis must start from a micro-level, and to follow an inductive research method. The analyst should not invoke just any context that they feel is appropriate or relevant in a speech event, because it might not be of concern to the interactants. Rather, the analyst needs to show that participants indeed organize their actions around the contextual variables described by the analysts (Schegloff 1992a).

In conversation analysis “context-sensitivity” is thus indeed narrowed down to concern 1) the analysis of the immediately surrounding interaction; i.e. the surrounding turns, actions and sequences, and the practices that the participants themselves can be shown to orient to,27 but also sometimes to 2) such topical and sociocultural aspects that the interactants can be shown to orient to. The latter point is obvious in some conversation analytic work on news interviews (Clayman & Heritage 2002), in which for example the understanding of how interviewers incorporate presuppositions in hostile and adversarial questions needs to consider the topical aspects of the question.

Both of the above points regarding the question of how much context the analysis can and must take into account are also central for the analysis of intersubjective stance taking. In fact, the ‘positioning’ paper (original article II) connects hostile and adversarial questioning in news interviews closely with a forward-type stance taking activity called ‘positioning’. Although “context” in the ‘positioning’ paper is understood narrowly, the descriptions of the phenomena in it benefit from the consideration of the knowledge that

26 This issue is also discussed briefly in section 7.1 in relation to the use of category terms and the construction of identities in talk-in-interaction.
27 This view is discussed and debated in detail in the dialogue between Michael Billig (1999a; 1999b) and Emanuel A. Schegloff (1999b; 1999c).
the participants possess about the relevant topical and sociocultural issues.” In the ‘identity’ paper (original article IV), on the other hand, the assumed “context” is broader (consideration of an ethnic identity), but still the analysis focuses on how the ethnic category is first topicalized and rendered noticeable for the participants, and then followed by intersubjective stance taking in which the (dubious) ethnic identity is constructed as part of the interlocutors’ practices and actions. Consequently, the ‘identity’ paper is able to address a broader vantage point for the workings of ideology, and how these workings emerge in situ in discourse, but importantly, it does so by relying on close empirical analysis of interaction.28

By limiting ‘context’ in the above manner I do not intend to undermine work that adopts broader views of “prior texts.” Stance and stance taking can naturally be examined in a beneficial way by considering broader social issues. One approach that does so, and which uses the conversation analytic methodology is discursive psychology. In the same way as conversation analysis was a response to the mainstream sociological research, and functional linguistics was a response to formalist views of language as an internal and abstract phenomenon (see section 1.1), discursive psychology is a response to the widely held view in the mainstream experimental social cognition tradition that evaluative expressions are related to people’s cognition and in that sense expressive of their “underlying” attitudes (Potter 1998).29 The aim of discursive psychology is “to avoid psychological theorizing in favour of analysis based in the pragmatics of social actions” in discourse and interaction (Edwards 2005, 260). Similarly to conversation analysis, discursive psychology avoids prior hypotheses or formulations of particular questions and is an empirical and qualitative enterprise which focuses on real and natural data, for example naturally occurring interactional talk (Potter 1998).30 Thus discursive psychology adopts a

discursive approach where versions of events, things, people and so on, are studied and theorized primarily in terms of how those versions are constructed in an occasioned manner to accomplish social actions (Edwards & Potter 1992, 8).

In short, discursive psychology can be defined “as the application of principles and methods from discourse and conversation analysis, and increasingly CA, to psychological themes” (Edwards 2005, 258). That is, it looks at these psychological themes from a discursive viewpoint, by concentrating on the interactional activity itself, and the interlocutors’ own understandings of the practices within the sequentially unfolding interactional activity (Potter 1998). It investigates such psychological phenomena as ‘memory’ (i.e. ‘what happened’), ‘attribution’ (i.e. ‘why it happened’), attitudes, evaluation, and emotion, but respecifies them, and investigates them from an ‘action-oriented’ and an ‘anti-cognitivist’ point of view. In other words, rather than talking for example about the internal aspects of ‘remembering’ or ‘evaluation’, discursive social

28 This is also discussed in the ‘theory’ paper (original article I, 113).
29 See also Wiggins (2003), Edwards and Potter (1992), and Edwards (2005).
30 Note that discursive psychology does also investigate written data, such as newspaper reports and articles (Potter 1998).
psychology focuses on explaining how people talk about the event they are remembering or what practices they use when they evaluate things.

Discursive psychology approaches data from two perspectives: the first one considers how people construct their worlds through their accounts and descriptions as they talk, and the second approach (which indeed differentiates it most clearly from conversation analysis) examines how these accounts and descriptions are related to the broader ideological and social contexts (Edwards & Potter 1992, 28; Potter 1998). One way to approach data in discursive psychology is to start from looking at a word, a term, or a construction, which stands for a particular psychological concept (such as ‘angry’, ‘happy’, ‘love’, and ‘like’), and then to proceed by considering their use in conversation, and how the use of these terms reflects the emotional and cognitive states of the speakers.31 This approach resembles research in interactional linguistics, which similarly takes up a particular linguistic item and investigates its use in real interactional situations. However, in discursive psychology the focus is more on how the use of these terms relates to the cognitive or psychological states of the speakers that are expressed through their use in interaction than on the interactional use of the linguistic items per se (te Molter & Potter 2005).

One possible direction that the functional-interactional perspective to stance taking could take in the future is to consider the ‘psychological’ and social aspects of interaction, similarly to discursive psychology (cf. Edwards & Potter 1992, 28). In fact, the ‘identity’ paper (V) touches on a discursive psychological theme, by looking at the discursive construction of attitude and stance in a news interview program, and how this activity further constructs an ethnic identity. Indeed, the aims of discursive social psychology can be closely linked with the examination of stance taking in talk-in-interaction. Not only can ‘stance’ be considered in a sense a psychological concept, but, similarly to the approach undertaken here, discursive social psychology also considers ‘attitudes’ from a discursive point of view; i.e. how interlocutors engage in practices of evaluation in various interactional settings (cf. Potter 1998). Therefore, in the future, a closer amalgamation of the type of functional-interactional linguistic research undertaken here with discursive psychology may yield interesting insights into how stances, attitudes and evaluations are intersubjectively constructed in talk-in-interaction.

In sum, in this section I have described how conversation analysis provides the tools for analyzing what interlocutors are doing in interaction; i.e. what type of practices they rely on in order to produce particular actions that are meaningful for themselves and bear on their understanding of each other’s actions. As we have seen above, many of the above-mentioned conversation analytic findings and notions have an important role in the examination of stance taking in talk-in-interaction. Indeed, in my work, this has been of prime importance for two reasons. First, the practices and actions that interlocutors produce or organize their stance taking. Second, some practices and actions that interlocutors produce are themselves expressions of speaker’s stances and examples of stance taking (cf. assessments and assessment sequences in Goodwin & Goodwin 1987; Goodwin & Goodwin 1992; Pomerantz 1984). I also briefly discussed the principles of

31 See for example Potter and Hepburn (2003), who investigate the use of the construction “concerned about x” in the early parts of the turns in calls to a child protection helpline in Britain.
discursive psychology, and suggested that it addresses issues that are relevant for the examination of stance taking and provides a rich source of research for pursuing stance research towards broader questions related to the psychological themes of emotion and attitudes. In the following sections I consider how functional linguistics and interactional linguistics contribute to the research approach adopted here by providing tools for analyzing the role of morphosyntax, prosody and lexis in intersubjective stance taking.

4.2 Linguistic-interactional approaches to stance taking

4.2.1 The functional linguistic ‘theory of stance’

This section concentrates on a functional linguistic approach to stance that has been outlined and developed by John Du Bois (2002a; 2002b; 2003d; 2004). I have called this framework ‘theory of stance’ (see original article I).32 Together with conversation analysis it forms the backbone of this work. Whereas conversation analysis provides the tools for analyzing the sequential organization of talk into actions from a dialogical and intersubjective perspective, Du Bois’s work provides a rich linguistic framework that can be used to supplement conversation analytic findings and methods.

The aim of functional linguistics in general is to discover motivations for why certain linguistic patterned forms and structures emerge and are used in natural discourse settings.33 Functional linguistics can be seen as a response to the formalist and sentence-centered views of language and grammar as innate and abstract, and as decontextualized and separate from issues of culture (cf. Cumming & Ono 1997; Du Bois 2001b, 87; Du Bois 2003c; Hopper 1987; Hopper 1998). Functional linguistics rejects the view that grammar exists prior discourse or that it is “an object apart from the speaker and separated from the uses which the speaker may make of it” (Hopper 1987, 141). In contrast to this, Hopper (1998) proposes that grammar is emergent, which means that “structure, or regularity, comes out of discourse and is shaped by discourse as much as it shapes discourse in an on-going process.” In view of this, functional linguists aim to find functional explanations for particular uses of syntax and, importantly, do this by studying real discourse (Hopper 1987).

In functional linguistics ‘context’ is thus understood more broadly than it is understood in conversation analysis (cf. the discussion of context in section 5.1). For example Du Bois (2003c, 54) argues that “by context we understand no less than the sum of the salient situation, prior discourse, interlocutors’ actions, models of mutual knowledge, cultural frames, and even the publicly visible processes of marshaling cognitive resources in acts of communication.”

---

32 Note that even though I here use the notion ‘theory of stance’, Du Bois’s findings are based on observation of everyday talk, not on theoretical presuppositions.
Although functional linguistic work is primarily linguistic, it draws from and is influenced by research in quite different research fields. One of these is the linguistic anthropological work in the United States (Boas 1995; cf. Sapir 1994; Sapir 1996; Whorf 1995), which has relied on extensive field work and emphasized the connections between linguistic and grammatical structures and culture (Duranti 1997; Duranti 2001; Duranti 2002; Foley 1997).

Du Bois (2002a; 2002b; 2003d; 2004) has also examined how interlocutors construct stances in everyday talk. He claims that “stances come in pairs and it takes two to take a stance” (Du Bois 2004). Stance taking should therefore be perceived as an intersubjective and emergent activity in which two subjectivities co-construct and negotiate their stances (feelings, emotions, attitudes, etc.) in relation to what a co-participant has said in the immediately preceding discourse. Therefore, Du Bois’s work can be seen as a critical response to the widely held view in linguistic research that particular linguistic (syntactic or lexical) forms encode or act as expressions of speakers’ internal and subjective attitudes (see section 4.3 below for an overview).

For Du Bois, stance taking is very much a linguistic process in which interlocutors use the reservoir of the linguistic (morphosyntactic, prosodic, lexical) tools that are available for them. This framework draws primarily from previous functional linguistic research, conversation analysis (e.g. Goodwin & Goodwin 1987; Pomerantz 1984) and linguistic anthropology34 (e.g. Besnier 1993; Haviland 1989; Ochs 1992), but it also owes theoretically to Benveniste’s (1971), Bakhtin’s (Holquist 1990; Vološinov 1973), and Tomasello’s (1999a; 1999b) work (see section 1.1).

Du Bois (2002a; 2002b; 2003d; 2004) understands stance taking as a tri-act activity. The stance-taking activity (see Figure 1 below) consists of three “elements”35: Subject1, Subject2, and Stance Object. The first two stand for two interlocutors, whereas the Stance Object refers to the topic or focus of talk (a person, an event, an issue, or the like) that the interlocutors are talking about. It is important to notice that in a stance-taking activity the Stance Object can remain the same or it can change without disrupting the coherence of the activity (cf. the example below).

---

34 For overviews, see Duranti (1997; 2001; 2002) and Foley (1997) inter alia.
35 This activity has been described in detail in the ‘theory’ paper (I), but since it is a central part of the present work it is reproduced here.
In the following example, which comes from a tape called ‘Burp’ (see chapter 2), Allison’s mother Mindy and older sister Rebecca have teased Allison about her crush on a boy called Mike. After the teasing Allison displays her dislike to Mike, to which Mindy responds by producing two evaluative stances about Mike and his mother.

(2) ‘Burp’ – 5-minute segment (2:26): He wants to be friends

38 → MINDY: ...(1.2) <VOX> He 'wants to be ^fri=ends.
39 → His ['Mom] is ^super 'swee=t </VOX>.

In other words, Mindy (Subject1) says something about two Stance Objects (Mike and his mother) and “puts a stance on the table” (Du Bois 2004). Whilst Mindy (Subject1) takes the stance, she pursues two things. First, she establishes a relationship between herself and the Stance Objects, i.e. positions herself in relation to them. The fact that the arrow in Figure 1 is pointing back to Subject1 indicates that this positioning is a result of the evaluative act. In other words, “I take this stance by evaluating the Stance Object in this particular way.” Second, the interlocutor creates a discourse/interactional context for Subject2. In the above example, this context is created by the evaluative action and topical content. Although the next speaker can do anything, the evaluative action makes relevant a subsequent action that somehow builds on it. In news interviews (see original article II), interviewers often create a context—by encoding, evoking or presupposing a stance—for the interviewee so that it constrains the way in which the interviewee should design her responsive stance (cf. the idea of forward-type intersubjectivity).
Next, a short moment later, Gerry (Mindy’s husband) responds to and simultaneously builds upon Mindy’s evaluation and says:

(3) ‘Burp’ – 5-minute segment (2:26): He wants to be friends

44 GERRY: [3His ‘Mom is ‘nice,
45 but ‘he is ‘no=t3].

In other words, Gerry (Subject2) responds to Mindy (Subject1) action by evaluating the same Stance Objects (Mike and her mother). By doing so, he positions himself relative to the same Stance Objects and simultaneously aligns with Mindy (Subject1). In other words, Gerry agrees with Mindy’s evaluation about Mike’s mother, but disagrees with the implication that Mike’s mother’s niceness would mean that Mike is nice. Consequently, Gerry’s stance becomes clear through his engagement with Mindy and the engagement between the linguistic forms they have used to express their stance. Finally, by displaying his stance, Gerry attests the intersubjective connection between himself and Mindy, but this time by “looking backward,” and by reflecting Subject1’s utterance and the stance therein.

However, it is noteworthy that the Stance Object need not remain the same, but it can change slightly. Consider the following segment from the above example, in which Allison (Subject2) in lines 42-43 responds to her mother’s stance.

(4) ‘Burp’ – 5-minute segment (2:26): He wants to be friends

38 MINDY: ...(1.2) <VOX> He ‘wants to be ‘fri=ends.
39 His [‘Mom] is ‘super ‘swee=t </VOX>.
40 REBECCA: [He --]
41 MINDY: I think ‘you [2’should2] --
42 → ALLISON: [2His2] ‘Mo=m is ‘nice,
43 → [3so ‘let’@s be ‘friends with ‘him3].

In Allison’s response His ‘Mo=m is ‘nice, so ‘let’@s be ‘friends with ‘him. in lines 42-43 it is Mindy’s evaluation per se that becomes the Stance Object. Nevertheless, not only does Allison engage with Mindy, but also the words and the syntactic forms in their turns engage with each other. From this engagement a particular resonating relationship rises, which then creates a new meaning. In other words, by building on the two TCU’s in her mother’s turn, Allison produces a cheeky and humorous stance. Allison’s stance would not receive the same meaning if it were produced in isolation. It is only after it is juxtaposed with Mindy’s previous stance that the participant’s (and we as analysts) understand its meaning.

Although in this work stance taking is considered a sequential activity, Du Bois (2003d) claims that there are also other resources from which stances can arise. For example, as we can see in the following examples, a participant can rely on an intertextual reference and use it as a resource for taking a stance. In September 18, 2001, one week after the terrorist attacks in New York and Washington D.C., the President of
the United States of America, George W. Bush, made the following famous statement about Osama bin Laden, the alleged leader of Al-Qaida.

(5) White House press conference, Sep 18, 2001: Dead or alive

IR: White House correspondent, IE: President George W. Bush

1 IR: Do you want bin Laden dead.
2 IE: (H) I want him hel- --
3 I want,
4 I want justice.
5 ... And uh,
6 .. Uh,
7 There's an old poster out west,
8 as I recall,
9 → that said wanted,
10 → ... dead or alive,

Three months later, in CNN’s Crossfire, the participants are discussing whether the United States should capture and bring bin Laden and another Al-Qaida leader Mullah Omar to justice or alternatively, try to kill them where they are. At one point in the interview, one of the participants says:

(6) CNN, Crossfire, Dec 27, 2001: Dead or dead

IR: Tucker Carlson, IE: Charlie Wilson (003 / 2 / 1:09)

1 → IE: ...(H) I would like to see them both dead.
2 → ... Dead dead.
3 ?: ... [(Hx) @]
4 → IE: [I don't s]ay dead or alive,
5 → I say dead or dead.
6 ...(H) Uhm,
7 .. but I <MRC>do=n't really want</MRC> them as
8 prisoners.
9 ...(H) And I had just as soon,
10 ... that ... it not be an American that pulls the
11 trigger.

It is quite possible that in news interviews intertextual references, for example third-party statements (Clayman & Heritage 2002), are used as resources for taking stances. In terms of sequential analysis, the grounding of such findings of course remains a problem, unless one is able to find the example with which a participant engages (cf. example (5) above). Due to this problem, this work concentrates on stance taking as a sequential activity.

In sum, the triangle is a useful way to depict that 1) interlocutors actively engage with each other when they take stances, 2) stance taking is not internal, and 3) interlocutors calibrate over stances in relation to a relatively constant object (Du Bois 2003e).

The notion of ‘alignment’ (see Figure 1 above) requires some further explanation. First of all, it does not mean the same thing as it does in conversation analysis, in which an “aligning” action is an appropriate or preferred next action, which fulfils the expectations raised by the previous action. Thus, for example, an acceptance aligns with a request, whereas a refusal disaligns with it (Heritage 1984b, 269). Second, ‘alignment’ here is not synonymous to ‘agreement’. Rather, it explains the range of possible types of convergent and divergent positions that interactants can take in relation to each other. These positions are negotiated with the help of language. Therefore, as Du Bois (2004) claims, alignment is a process of calibration between stances and it is essentially a dialogic linguistic activity in which interactants use morphosyntax, lexis and prosody to construct their stances in relation to a prior stance. However, in addition to the fact that alignment is understood as a dialogic linguistic activity, in the present work alignment is also displayed by particular ways of doing actions and of turn design.

The central and most important aspects of the above stance-taking framework and the ‘stance triangle’ for the present work are that it, first of all, emphasizes and focuses on how interactants design their stances dialogically and intersubjectively with the help of various syntactic, lexical and prosodic practices. Second, it focuses on language by considering how the interlocutors’ language engages across turns (and by implication how the interactants and their stances engage with each other) and how this affects stance taking. And it is the level of detail that is accorded to the subtleties of lexis and morphosyntax which epitomizes the important difference between Du Bois’s ‘theory of stance’ and conversation analysis. Even though conversation analysts carefully consider the details of language use and by no means exclude any linguistic element from analysis, their focus is mostly on how language contributes to the structure and social order of interaction, and thus they often give little linguistic attention to the lexical, morphosyntactic and prosodic detail, and how they impinge on intersubjective stance taking.37 Finally, the theory of stance also acknowledges the importance of the surrounding world, the larger sociocultural issues and stances that have been uttered and taken in other interactional contexts and which may thus affect the interlocutors’ present stance taking. However, these are given fairly little attention in the present work, since the aim has been to concentrate on how stance taking is managed locally in interaction. Nevertheless, Du Bois’s framework plays a central role for the present work because it can be used fruitfully together with the conversation analytic work for combining the benefits of both linguistic and interactional aspects of stance taking.

One central element of linguistic alignment, as Du Bois (2001c; 2003b) further notes, is that interlocutors surprisingly often use and recycle linguistic elements from their co-participants’ utterances (see section 3.1). Du Bois (2001c) calls this ‘dialogic syntax’, and it is discussed in greater detail in the following section.

37 However, for exceptions of conversation analytic work that focuses on syntax, see Lerner (1991; 1996b) and Lerner and Takagi (1999).
4.2.2 Dialogic syntax: Recycling linguistic elements in interaction

The discussion of dialogic syntax in this chapter continues two themes developed in the previous sections. First, dialogic syntax (Anward 2004; Du Bois 2001a; 2001c; 2003b) is both a theory and a research practice that examines and describes how various types of morphosyntactic, prosodic and lexical units and structures are recycled, reused and borrowed across turns and utterances in spoken interaction. Dialogic Syntax is thus a way of understanding and representing speakers’ dialogical use of language (Du Bois 2001c). By using Dialogic Syntax it is possible to gain insights about how the linguistic units and structures by one interactant engage with the units structures of another. Such linguistic engagement displays the participants’ syntactic knowledge and their understanding of each other. By such engagement, they can modify their stances and create new meanings. Thus, the type of linguistic recycling that is described by dialogic syntax is pivotal for understanding stance taking and the fine-tuned intersubjective engagement between interlocutors and their stances in talk-in-interaction. It is worth noting that Anward (2004) describes the very similar linguistic phenomenon of recycling linguistic elements. Second, this section develops the theme in chapter 3, i.e. the reasons for using the ‘Discourse Transcription’ system in this work.

As Du Bois The framework of dialogic syntax has been developed on the basis of findings made from actual data, namely the Santa Barbara Corpus of Spoken American English (SBCSAE) (Du Bois et al. 2000; Du Bois et al. 2003a; Du Bois & Englebretson 2004). In fact, Du Bois (2001c) claims that the type of ‘recycling’ described here is frequent in spoken discourse. However, it is important to note that ‘recycling’ is not considered to be equivalent to ‘repetition’. Although the co-participants recycle linguistic elements, they modify and change the prior talk for their own purposes (e.g. for designing a particular stance). Dialogic syntax has many interfaces with Bakhtin’s ideas of how texts and prior texts are connected (Holquist 1990; Vološinov 1973), and the same point about ‘recycling’ was also noted by Bakhtin when he claimed that in conversation “both speakers are different from each other and the utterance each makes is always different from the other’s (even when one appears to repeat the “same” word as the other); and yet all these differences – and many more – are held together in the relation of language” (Holquist 1990, 40).

Bakhtin’s notion of dialogism indeed offers a helpful theoretical frame and context for dialogic syntax. However, whereas Bakhtin understands ‘dialogism’ quite theoretically, one of the most important aims of dialogic syntax is to provide means to empirically show that interactants indeed do use syntax dialogically and to provide tools for unraveling empirically the functions of linguistic dialogicality.

In the following extract (same that was discussed in section 4.2.1.) I show an example of linguistic and dialogic recycling. Rebecca, Allison and Joyce are sisters, Mindy is their mother and Gerry is their stepfather. Julie is a friend of the family. In lines 1-26 Rebecca
and Mindy tease Allison for her previous crush on Mike, after which the participants begin to evaluate Mike and his mother.38

(7) ‘Burp’ – 5-minute segment (2:26): I don’t like Mike anymore

1 MINDY: What2] did 'we 'do=,
2 ~^Rebecca.
3 We 'did ^something to her.
4 GERRY: [You went and ^knocked on] her 'door.
5 MINDY: [We did 'like a ^test].
6 ALLISON: .. [2And you're ^all2],
7 [2<CLAP + + + CLAP>2]
8 ... <VOX> 'Give me the ^phone </VOX>,
9 or `somet[thing],
10 GERRY: [Yeah].
11 MINDY: [Oh],
12 [2^yeah@ 0@02].
13 ALLISON: [2And ^I was 'like2],
14 .. <VOX> <F>No=(0.6)</F></VOX>.
15 [0][0] ((SNIFF))
16 MINDY: [0]
17 I'm 'all,
18 <VOX> 'Give me the ^phone=, 
19 I 'need it right ['no-w]. </VOX>
20 REBECCA: [That] was 'rea-lly ^funny.
21 MINDY: And 'then,
22 REBECCA: <VOX> <F> No=(1.1) </F>.
23 I'm 'ta=king] to ~^Mi=ke. </VOX>
24 MINDY: [@@@]
25 GERRY: (H) @[2@2]
26 MINDY: [2002]
27 ALLISON: [2' Shut2] ^u=p,
28 God I ^don't 'like ^-Mike anony=me.
29 REBECCA: ... [But 'you ^di=d].
30 JOYCE: ['-Mike's ^stupi=d].
31 ALLISON: ...(0.8) ^He 'bugged me,
32 #Always ^'tried to ^ta=lk to me,
33 and 'I'm like --
34 ... (SMACKS LIPS)
35 GERRY: .. Oh,
36 he ^tries to ^talk to ^you?
37 MINDY: ...(1.2) <VOX> He 'wants to be ^fri=ends.
38 His [^Mom] is ^super 'swee=t </VOX>.
39 REBECCA: [He --]
40 MINDY: I think ^you [2'should2] --
41 ALLISON: [2His2] ^Mom is ^nice,
42 [3so ^let'@s be ^friends with ^him3].
43 GERRY: [3His ^Mom is ^nice,

38 The names of the participants in all the examples from everyday talk here and in the original article (IV) have been changed.
but he is not.

JULIE: @

MINDY: ... He's nice.

JOYCE: [He is too nice.]

ALLISON: [2 He was mean to me=2],

JULIE: [2 He's nice=2],

ALLISON: [3 What he did last year=3],

JULIE: [3 And that's=3] [4 good=4].

GERRY: [3 You think so=3]?

ALLISON: [4 though=4].

GERRY: ... 'I don't know.

Some of the things she said about him,

ALLISON: [5 I don't think he deserves my friendship=5].

GERRY: ... <CRK> made me wonder a little bit</CRK>.

At first sight, the above transcription does not readily reveal how the interlocutors recycle each other’s linguistic units and structures as resources in their own talk. However, the activity of recycling becomes clearer when we consider the recycled elements in a *diagraph*. A diagraph is an abstracted representation of talk and can be used to depict and illustrate structural relationships and parallelisms between linguistic elements across (and within) speakers’ turns of talk (Du Bois 2001c). One line in a diagraph represents one intonation unit.39 In a diagraph parallel utterances are aligned vertically, which clearly demonstrates and foregrounds the relationships between parallel linguistic elements. With the help of the following diagraphs I show first how the interlocutors can engage with each other’s language, and second, how such engagement can cause new meanings to emerge.

As has been outlined above, it is necessary to examine the sequentially unfolding interaction and the actions the interlocutors produce when they take stances. Therefore, I connect the description of the recycled language with a general account of what the participants are doing with the utterances they produce when they participate in a stance-taking activity (cf. Anward 2004). Rebecca’s and Mindy’s teasing of Allison builds on Allison’s crush on Mike. This can be seen in the way in which Allison responds to Rebecca’s action in lines 22-23 in which Rebecca uses direct reported speech to repeat what Allison has said: <VOX> <F> N[o=(1.1)] </F>, I'm 'talking] to Mike. </VOX>. Rebecca’s teasing action simultaneously evokes a stance that Allison can be seen to have taken (that she likes Mike) when she produced the talk that Rebecca now reports. Although this stance does not become explicit in Rebecca’s turn, Allison makes it explicit in her next turn in line 29. In this turn Allison denies that she likes Mike anymore by saying *God I don’t like Mike anymore*. By producing the denial Allison produces an interpretation about Rebecca’s previous turn (that Mindy was teasing her because she

---

39 At this point it is important to recall the discussion in section 3 about the importance of intonation units in this work. The most useful and advantageous unit for making diagraphs is the intonation unit. This of course suggests that intonation units are meaningful units for interlocutors, because the recycling of linguistic elements overwhelmingly happens across intonation units and not, for example, across turns.
likes Mike) and thereby tries to vitiate Rebecca’s and Mindy’s teasing (cf. Ford 2002). As we can see in the following diagraph, the linguistic design of Allison’s denial and the fact that it engages with Mindy’s utterance, and aligns with it, act as further evidence of its connection with Mindy’s tease.

Diagraph (1)

23 REBECCA: I’m talking to Mike.
29 ALLISON: God I don’t like Mike anymore.

Du Bois (2001c) claims that when interlocutors recycle a linguistic unit or a form used by a co-participant in the immediate prior discourse, these parallel units or forms are brought into a paradigmatic relation to one another. In such a relation, they engage with each other and begin to resonate. The term ‘resonance’ here refers to the emergent ‘activation’ of relationships between recycled structures and the (semantic, social, or cultural) meanings they incorporate, when they are being juxtaposed with equivalent structures with their own meanings in prior discourse (Du Bois 2001c). This type of resonance between linguistic structures (and their referential and indexical features) may bring with it and foreground a new meaning which is invisible to the structures in isolation (ibid). Even partial parallelism of linguistic structures can invoke resonance which in turn can generate new locally constructed meanings. The best examples, as Du Bois (2001c) suggests, occur when one speaker constructs an utterance based on the structure or linguistic units used in the immediately prior utterance of a dialogic partner.

In diagraph (1) above, we can see how Allison’s utterance engages with Rebecca’s previous utterance. First, although the first person pronouns are uttered by different speakers, they are co-indexical. The fact that Rebecca’s utterance directly reports something that Allison has said earlier enables her to use the first person pronoun that still indexes Allison. Second, the proper name Mike naturally indexes the same referent. In addition to this, the diagraph shows how a cultural and social meaning is invoked by the resonance between Rebecca’s and Allison’s turns and is used for stance taking. Allison takes a stance in line 29 by building upon and making an interpretation of Rebecca’s utterance. This is explicit in the way in which Allison’s verb like engages with Rebecca’s verb talking to. In other words, the linguistic engagement of like with talking to displays Allison’s interpretation of the meaning of talking to (a boy on the phone) to index “liking.” Allison’s divergent alignment with the stance taken in Rebecca’s turn can be seen in the negative n’t (in the auxiliary don’t) and the particle anymore.

The above is a rare example of how the NEG + POS pattern is used in the denial + account action combination in everyday talk (Ford 2002). Allison constructs her denial by using an epistemic stance marker I don’t like together with a linguistic element that is partly recycled from the previous turn. The denial projects and makes relevant an explanation or an account for the denial (Ford 2002), which Allison gives in lines 32-33 when she says ’He ‘bugged me, #Always% ’tried to ^talk to me,. It seems that Allison

---

40 It is naturally also possible that this type of engagement occurs within a turn or within one speaker’s speech.
uses the NEG + POS pattern for producing the denial, because of the problematic position she has been put in by the teasing action. The only other example of this pattern in everyday talk that I have encountered occurs after a question that is also problematic for the recipient. It seems therefore that this linguistic pattern is used for denying stances that are imposed on the individual who produces the denial, and which are personally problematic for her. This is course a frequent problem for interviewees in news interviews, who continuously face problematic questions and stances. Therefore, the NEG + POS pattern is also much more frequent in news interviews than it is in everyday talk.

Allison’s account is composed of a personal negatively affective stance “He ‘bugged me,” which is followed by an extreme case formulation (Pomerantz 1986) of a factual event #Always% ‘tried to ^ta=lk to me.. Gerry responds to the latter part of the account in line 36-37 with an oh-prefaced utterance “Oh, he ‘tries to ^talk to ‘you?, in which the particle Oh marks the content of Allison’s explanation (that Mike still tries to talk to her) as news for him (Heritage 1984a). Now consider the way in which Gerry accomplishes this action in Diagraph (2) below.

Diagraph (2)

33  ALLISON:      Always  tried to   talk to      me  
37  GERRY: he           tries to   talk to       you

As we can see in the diagraph, Gerry recycles the structure of the utterance almost verbatim except for the object me which is changed to you, but which still maintains the same indexical referent, namely Allison. The very fact that Gerry uses similar verb phrases (try to do, talk to x) displays his engagement with Allison’s account. Simultaneously, Gerry uses syntax dialogically for identifying the part that he is commenting on and engaging with.

In addition to the fact that interlocutors can use syntax dialogically for identifying a part with which they engage and for displaying an understanding of a prior linguistic element, recycled syntax can also construct new meanings. After Gerry’s turn Mindy says “He ‘wants to be ^ fri=ends. in line 38. Mindy’s turn is sequentially relevant, because it formulates a possible reason why Mike still tries to talk to Allison, namely, he wants to be friends. After this Mindy shifts the discourse topic slightly and produces an assessment about Mike’s mother: “His ‘Mom] is ^super ‘swee=t. As we can see in Diagraph (3) below, although there is basically low resonance between Mindy’s assessment and the prior utterances, the subject argument in the assessment, “His ‘Mom” is co-referential and engages with the pronoun he in previous talk.

Diagraph (3)

33  ALLISON:      Always  tried to   talk to      me  
37  GERRY: he           tries to   talk to       you

As Du Bois (2001c) suggests, two utterances do not always engage and resonate with each other. Consequently, resonance is contingent, just like any other syntactic phenomena in interaction. It must be achieved, and when it occurs, it is used for a reason.
Nevertheless, Allison in her next turn, partly humorously and partly cheekily, puts two and two together and says [2His2] ^Mo=m is ^nice, [3so ^let'@s be ^friends with ^him3].

There are three things that simultaneously make Allison’s humorous remark and the stance therein possible. The first one is that Mindy’s utterances in lines 38-39 are produced in sequentially consecutive intonation units. Although the two TCUs in Mindy’s turn are doing separate actions (line 38 provides an explanation and line 39 is an assessment) it seems that the vicinity of these two TCUs contributes to the understanding that Allison communicates and the stance she takes in lines 42-43: *His ^Mom is ^nice, so let’s be friends with ‘him.*, in which the particle *so* indeed marks that she connects the two apparently separate actions in her mother’s previous turn. This aspect of how the design of a previous turn affects the stance in the following turn is closely related to the idea of ‘positioning’ (original papers I and II). That is, the closeness of two apparently separate actions within a turn can contribute to the way in which that turn’s stance as a whole can be understood. It is important to note that this is not an example of an ‘action combination’ (cf. the original article III and Ford 2002). In action combinations, as we saw with Allison’s denial and account in example (7) above, the first action strongly projects the second one (within one turn), whereas in the above example and in the examples of positioning (original paper II) the separate TCUs can be understood to be connected in terms of how they communicate the speaker’s stance. As we can see in example (7) above, Allison indeed provides an understanding of what she thinks her mother’s stance is, although her mother did not specifically take the stance Allison reports. And moreover, this understanding is communicated with recycled language.

The second is that the two reference forms *He* and *His Mom* can be seen to be categories of the same collection (or categorization device) (Sacks 1972b; Sacks 1979), namely that *He* is the category of ‘son’, *His Mom* is the category ‘mother’, and they belong to the collection or device ‘family’. In other words, even though the two TCUs in Mindy’s turn in lines 38-39 are doing separate actions, the fact that the two turns contain categories from the same collection enable Allison to hear that the two TCUs are connected, a possibility that she indeed voices in her subsequent turn.

---

42 Curly brackets in the diagraphs indicate that some unit has been moved or removed for representational purposes.

43 Cf. Sacks (Sacks 1972b) and his analysis of “The baby cried. The mommy picked it up.” See also the discussion of membership categories and collections in chapter 7.
However, there is also a third element, the parallel use of some of the linguistic units and phrases, that Allison draws on (the two engaging linguistic units He in line 38 and His Mom in line 39 which occur in the same linguistic position in the intonation units) and uses in her humorous remark. As was mentioned above, Du Bois (2001c) and Anward (2004) have noticed how speakers construct utterances on the basis of linguistic elements used in prior utterances. And indeed, Allison uses the possibilities given to her by the prior utterances. She draws a parallel between be friends (in line 38) and super sweet (in line 39) in her mother’s turn (see Diagraph 3 above) and explicitly voices the unintended causal relationship she has heard between those units by repeating the relevant elements and linking them with the adverbial so in line 43. Several elements in Mindy’s and Allison’s turns engage with each other (‘His mom’, ‘supersweet and nice’, and ‘be friends’), but they are not used for expressing a similar stance. On the contrary, in this context the engagement displays the divergent stances between Mindy’s and Allison’s turns.

After this and partly in overlap, Gerry participates in the stance taking by saying [3His Mum is ‘nice, but ‘he is ‘not] in lines 44-45. As we can see in diagraph (5) below, also Gerry engages with the language and the stances in the previous turns, aligns with them in order to produce his own stance about the issue.

Diagraph (5)

| 39 MINDY:         | His Mom    is          super sweet |
| 42 ALLISON:       | His Mom    is          nice        |
| 44 GERRY:         | His Mom    is          nice        |
| 45              | but he         is    not           |

The above analysis of example (7) shows and further supports Du Bois’s (2001c) and Anward’s (2004) findings regarding the importance of examining how interlocutors recycle language from their co-participants’ talk. Furthermore, it shows how such dialogic and intersubjective orientation to each other’s talk can contribute to meaning-making and stance taking in situated discourse. The above type of parallelism between utterances can occur at lexical, phrasal, morphosyntactic and prosodic levels (Du Bois 2001c). It is worth noting, as the above example shows, that recycling linguistic structures does not necessarily coincide with agreements. Rather, parallel structures in talk-in-interaction can be used for producing very different actions and activities, which can range from convergent alignment between interlocutors to total divergent alignment. Thus, participants can use recycled syntax for expressing very subtle differences in their stances. Du Bois (2001c) calls these subtle differences stance differentials and Anward (2004) recycling with difference. Moreover, as Anward (2004) (42) claims, for “recycling with difference to work, participants must […] remember fully detailed experienced exemplar turns and use them as models for new exemplar turns.” Consequently, dialogic syntax stands as a good example of the emergence and intersubjectivity of language in interaction.

44 Stance differential is also closely connected to the notion of ‘alignment’ described in section 4.3.1.
In sum, dialogic syntax provides a new and interesting vantage point for investigating language as it is used in spoken interaction. Dialogic syntax brings with it a fresh way to approach questions that have concerned linguists interested in, for example, the cognitive linguistics and language learning. However, the most interesting aspect of dialogic syntax for the present purposes, as we saw above, is that it offers tools for analyzing stance taking, and how meanings are intersubjectively constructed in different types of stance-taking activities. Traditionally syntactic analysis has concentrated on the linear and hierarchical construction of meaning within sentences. However, as was already mentioned above, the meaning of an utterance is partly dependent on the way in which the linguistic structures in it are paralleled with and mapped onto the structures in the immediate interactional context. Therefore, Dialogic Syntax forces us to view the traditional sentence-centered linguistic and syntactic analysis in a different light.

In the original papers (I-V) dialogic syntax and diagraphs are used to exemplify and clarify several phenomena, such as the construction of two action combinations in news interviews (III), the relationship between embodied practices and recycled language (IV) and the negotiation of an ethnic identity (V).

4.2.3 Interactional linguistics

The functionalist view in linguistics interestingly meets with conversation analysis (and also sometimes with linguistic anthropology) in the interdisciplinary and empirical body of research that calls itself interactional linguistics. Interactional linguistics concentrates on describing the relationship between linguistic structures (i.e. morphological, phonetic, phonological, semantic, syntactic) and naturally-occurring spoken language in both everyday and institutional settings (Couper-Kuhlen & Selting 2001a, 1-3; Ford et al. 2003; Ochs et al. 1996). In general it aims to bring together the strengths of the above disciplines in order to study naturally-occurring spoken language in its actual interactional settings (Keevallik 2003, 22). Scholars in interactional linguistics consider ‘language as a system’ and ‘everyday language use’ as inextricably intertwined (Thompson 2001, vii). Interactional linguistics aims to relate linguistic phenomena to findings in conversation analysis and vice versa (Thompson 2001, vii) and focuses on “how certain syntactic and other structures can be attributed to, and motivated by, the accomplishment of interactional tasks in situated use of language” (Keevallik 2003, 23). Interactional linguistics uses both qualitative and quantitative research methods and investigates the relationship between linguistic detail and interaction from two starting points:

(i) what linguistic resources are used to articulate particular conversational structures and fulfill interactional functions and (ii) what interactional function or conversational

---

45 De Saussure’s distinction between langue and parole might act as a helpful analogy in understanding the difference between language and everyday language use.
structure is furthered by particular linguistic forms and ways of using them? (Couper-Kuhlen & Selting 2001a, 3)

One of its most interesting and important findings has been that linguistic structures are emergent and not only closely related to the context in which they occur, but also dependent on and shaped by the actions speakers do with their talk (Thompson 2001: vii). It has also been noted that many linguistic elements are frequently and repeatedly used by interactants to solve various interactional problems when they talk with each other (Thompson 2001: viii). This general finding is crucial in understanding how and for what purposes speakers use language in their everyday interactional situations. It also provides insights into how interlocutors use their knowledge of grammar in interaction.46

The interactional linguistic research tradition has been used to examine phenomena in different languages, for example English, Finnish, German, Dutch, Japanese (Couper-Kuhlen & Selting 2001b; Ochs et al. 1996) as well as Estonian (Keevallik 2003). This body of work has looked at for example the use of negative markers in TCUs and how they are regularly followed by a unit that elaborates the TCU that contained the negation (Ford 2001);47 how for example the use of the verb remember in real interactional situations differs from the ways it is typically considered to be used (Tao 2001). In other words, rather than being used as a complement-taking predicate Tao found that “(a) remember rarely takes a complement clause and (b) the environments in which remember occurs are unlike those in which other members of the illusory class of cognitive verbs occur” (Ford et al. 2003, 121); how particular finite verb forms in Estonian are used for example for producing particular actions (Keevallik 2003); how the use of a full NP to refer for example to a person, after that person has already been talked about and referred to with pronouns, occurs in contexts in which the speaker wishes to display that previous sequence has ended or is now continuing (Fox 1987); how the adverb actually has different roles depending on its placement in a turn in informings, self-repair and topic movement (Clift 2001). She shows that if actually is produced TCU-initially it tends to reserve the right for the speaker to adopt the subsequent turn and is thus invulnerable to interruptions and overlap. However, if it is produced TCU-finally it is more vulnerable to overlapping talk by other participants. Clift (2001) further claims that the variable uses and functions of actually are made possible by the flexibility of grammar. Nevertheless, as Clift (2001, 252) further notes it cannot be placed anywhere. Rather, its use is closely related to the activities produced in the sequential context in which it occurs; and so on.

The interactional linguistic vantage point can further our knowledge about stance taking by offering tools for understanding how different linguistic patterns in some sequential contexts may function as expressions of stance. As Wu (2004, 17) claims

the significance of sequential positioning lies in the fact that it can play a constitutive and deterministic role in giving meaning to certain linguistic practices and/or to the stances they embody.

46 Cf. the outline of dialogic syntax in the previous section.
47 This is closely connected to the production of denials and accounts in everyday talk (Ford 2002) and in institutional talk (original article III).
Wu (2004, 19) further continues:

like all other human actions, the displaying of stance commonly requires the simultaneous deployment of a multiplicity of linguistic (and non-linguistic) resources, such as the lexical, syntactic, prosodic and sequential aspects of a turn design.

Keisanen (under review) takes an interactional linguistic view of language to describe the activity specific features of stance taking as they are manifested in the doing of requesting confirmation and challenging with a tag question or a negative yes-no interrogative. Along similar lines, Wu (2004) investigates two final particles in Mandarin spoken in Taiwan (a and ou) by looking at how speakers display their stances by using these particles. Also Kärkkäinen (2003, 271) looks at the stance marker I think and claims that stance markers do not simply index utterances to speakers, but are engendered by the ongoing speaker-recipient interaction. Kärkkäinen (under review) further argues that stance taking requires participants’ joint engagement. Furthermore, Rauniomaa (under review) investigates the use of two related Finnish stance markers, minun mielestä (‘in my opinion’) and minusta (‘I think’), and how they are used in the context of assessments. She demonstrates that these stance markers do not reveal the speaker’s stance but suggest ways in which the recipients are to understand the upcoming assessment in that particular sequential environment. She identifies three functions: projecting disagreement in a second assessment, marking transition to a first assessment across turns and marking transition to a first assessment within an extended turn.

In the present work, the interactional linguistic viewpoint is most explicitly used in the ‘denial’ paper (III), which considers the use of the linguistic pattern NEG + POS and how it is recurrently used in news interviews by interviewees for constructing two action combinations: the denial + account and the claim for insufficient knowledge + elaboration.

4.3 From single-speaker linguistic contributions to stance to intersubjective stance taking

As I have suggested above, the investigations of stance within linguistic research usually start from a particular linguistic form, for example a syntactic unit or a lexical item. Furthermore, this research describes the ways in which a linguistic form, from a pragmatic or semantic vantage point, expresses the speaker/writer’s inner attitudes and feelings, i.e. her subjective stance. In these studies the investigated linguistic elements tend to be decontextualized from their interactional environments. This linguistic research also uses various alternative terms for referring to ‘stance’, such as ‘appraisal’ or ‘evaluation’. This section briefly outlines and discusses some of this work relative to the functional-linguistic approach outlined above. With the help of examples from my news interview corpus I show how the mere investigation of single-speaker contributions to
stance tells us little about the intersubjectivity of stance taking, and how interlocutors engage and align with each other when they take stances.

The appraisal framework within systemic functional linguistics is one of the most extensive approaches to stance-related phenomena in discourse. Its main aim is to account for attitudinal meaning in spoken and written discourse by looking at evaluative lexis (see Martin 2000; 2003). It also aims to develop a comprehensive framework of the semantic resources used for negotiating emotions, judgments, evaluations or appraisal in discourse (Martin 2000, 145). The appraisal framework identifies different systems of choices that are used to express appraisal. These are ‘affect’, which is a resource “deployed for construing emotional responses” (Martin 2000, 145), ‘judgment’, which is “deployed for construing moral evaluations of behaviour” (Martin 2000, 145) and ‘appreciation’, which “construes the ‘aesthetic’ quality of semiotic text/processes and natural phenomena” (Martin 2000, 145-146). Extract (8) below gives an example of affective evaluation in light of the appraisal framework.

(8) CNN, Larry King Live, Sep 13, 2001: Angry?
    IR: Larry King, IE: Bradley Burlingame (009/7/6:08)

1 IR: Angry Bradley?
2 IE: (TSK)(H) My= .. emotions are across the board,
3       .. I'm sad,
4       .. I'm in disbelief,
5       (H) ... Yes I'm angry.

In example (8) the interlocutor’s affective evaluation is instantiated three times in the utterances (in lines 3-5) .. I’m sad, .. I’m in disbelief, and I’m angry. The adjectives in these utterances display the speaker’s displeasure toward the fact (i.e. the appraised) that his brother was a pilot who died on the flight that crashed into the Pentagon in the terrorist attacks on September 11, 2001. Within the appraisal framework the above affectual adjectives, which have negative cultural connotations, would perhaps be said to be used as a resource by the speaker for construing or evoking an emotion that displays the speaker’s current mental state and negative disposition towards the recent unforeseen events. In sum, this type of analysis concentrates, first, on the semantic resources the speaker relies on, and second, on the subjective state of the speaker. However, it rarely considers appraisals within their interactional contexts or as contingent responses that result from the intersubjective relationship between the co-participants. In the above example, the interlocutor’s affective stance is indeed occasioned by the interviewer’s question and recycles a word from it (angry). It can be argued that the interviewer agrees with the interviewer’s yes/no question, since he in line 5 explicitly displays his agreement by saying Yes. However, the construction of his affective stance also contains linguistic elements that are used for negotiating his stance relative to the question. In other words, he is not just angry, but that his anger is mixed with emotions across the board, i.e. sadness and disbelief. In addition to this, as we can see in diagraph (6) below, many of the linguistic elements in the two turns engage with each other. The different linguistic elements are either co-indexical (proper name Bradley and the first person pronouns) or
semantically parallel and resonate with each other (the predicate adjectives angry, sad, and disbelief).

Diagraph (6)

1 IR: Bradley {Angry}?
3 IE: I 'm sad,
4 I 'm in disbelief,
5 Yes I 'm angry

In light of the above, therefore, it cannot be said that the interviewee just agrees with the question. Rather, he negotiates his stance by aligning with the content and structure of the question.

Another example of the examination of single-speaker contributions (such as lexical and grammatical resources) for expressing stance is the work by Biber et al. (1999), Biber et al. (2003) and Conrad and Biber (2000). They look at and identify numerous stance expressions by using the Longman Corpus of Spoken and Written English. This corpus contains discourse from four different registers: conversation, fiction, newspapers and academic prose. The authors (Biber et al. 2003, 460) define stance as "overt expressions of personal attitudes or feelings towards the content of a clause." Similarly, Biber et al. (1999, 966) claim that both speakers and writers commonly express stance meanings overtly, using either grammatical or lexical means. With grammatical marking, some grammatical device is used to express a stance relative to another proposition.

This research suggests that stances are expressed by using for example evaluative lexical items, such as adverbs (unfortunately in line 3 in example (9) below), different types of combinations such as premodifying adjective + noun (terrible disaster in line 12), verbs or modal markers (has to in line 19) that indicate affect, certainty, and doubt (Biber & Finegan 1989; Biber et al. 1999, 968), or complement clauses that are controlled by for example verbs, adjectives or nouns (Biber et al. 1999, 969).

(9) CNN, Sep 11, 2001: He has to be it
IR: Aaron Brown, IE: Peter Bergen (002 / 1 / 1:26)
1 IE: (H) A=nd,
2 u=h,
3 unfortunately,
4 %uh,
5 he must <A>top of the list of</A>,
6 %u=h,
7 the %persons,
8 sop'histicated enough,
9 in terms of 'operations,
10 (H) %U=h,
11 to 'bring off these kinds of %uh,
Furthermore, Biber and Finegan (1988) and Conrad and Biber (2000) have concentrated on investigating stance adverbials and their occurrences in large corpora. Conrad and Biber (2000, 56-8) claim that stance adverbials are grammatical devices that are used to frame the status of information presented in the main clause (proposition). For example, the attitudinal adverbial\textsuperscript{48} \textit{unfortunately} in line 3 in example (9) conveys the interviewee’s value judgment about what he says in lines 4-12.

Biber and Finegan (1988), on the other hand, base their categorization of stance adverbials on the semantic distinctions in Quirk, Greenbaum, Leech and Svartvik (1985, 612ff), although they make the important point that the functions of these categories often differ from their literal meanings. Similarly to Conrad and Biber (2000) above, Biber and Finegan (1988, 34) claim that adverbials express attitudes towards the content of the sentence. For them, \textit{unfortunately} belongs to a group of “amazingly” adverbials, which the authors claim are infrequent in their data (Biber & Finegan 1988, 13) and occur mostly in personal memoirs and autobiographies.

Another similar example, but this time of the use of the adverb ‘actually’, is given in example (10) below.

\begin{example}
IR: Tim Sebastian, IE: Dan Plesch (031 / 3 / 11:54)
\begin{verbatim}
1 IR: Dan Plesch,
2 in London (GLOTTAL),
3 The fact is that, ...
4 ... treaties don’t work,
5 The world has never had more treaties,
6 than we have at the moment,
7 (H) Never been more unstable,
8 has it.[t].
9 IE: \[We\]'ll,
10 \[H] actually,
11 I think the world is a lot more stable,
12 than it uh was,,
13 during the cold war,
\end{verbatim}
\end{example}

\textsuperscript{48} ‘Attitudinal adverbials’ are one of the three semantic subcategories of adverbial classes. The others are ‘epistemic adverbials’ (e.g. ‘probably’, ‘maybe’, and ‘certainly’) and ‘style stance adverbials’ (e.g. ‘honestly’ and ‘more simply put’) (Conrad & Biber 2000, 58).
According to sentence or utterance level linguistic analyses of the adverb ‘actually’ in line 10 it can be claimed that it expresses the interviewee’s stance towards the subsequent proposition the world is a lot more stable, than it uh wa=s, during the cold wa=r, in lines 11-13. Thus, the adverb is here an epistemic stance marker, which comments specifically on the actuality of the subsequent proposition. It is also grammatically realized as a single adverb and placed in turn-initial position. These characteristics of ‘actually’ are very common in spoken discourse (cf. Conrad & Biber 2000, 64). However, these findings only explain the adverbial’s position or function within a sentence/Clause/Utterance. In addition to this, one could see the adverb ‘actually’ to be expressing a conflict between the speaker and a stance proposed in the prior turn. This possibility is also noted by Chafe (1986) and Biber and Finegan (1988). In other words, they acknowledge that ‘actually’ can also be used intersubjectively. As we saw in section 4.2.3., Clift (2001) has studied the interactional and intersubjective use of ‘actually.’ The above example resembles the use of actually in the following way. First, it occurs after a question that projects a yes-answer and together with the disagreement marker well (cf. Clift (2001, 270). Second, the TCU in which it occurs resembles a counterinforming action that corrects an informational error (‘world has never been more unstable’) in the previous turn. However, what is interesting is that whereas in the above example the ‘actually’ occurs turn and TCU-initially after a question that projects an agreeing no-answer, Clift’s data contained no such uses of ‘actually.’ Rather, in Clift’s data the turn and TCU-initial uses of ‘actually’ marked contrast with the information in the speaker’s own prior turn. The reason for the “different” use of ‘actually’ in example (10) is that it occurs in news interviews and is sequentially placed after a question that sets up a difficult position for the interviewee (cf. original articles I and II).49 The interviewee uses ‘actually’ for aligning with the difficult question. In addition to this, here the TCU that contains the adverb ‘actually’ is performing a similar type of action that the ‘denial’ and the ‘claim for insufficient knowledge’ (see the ‘denial’ paper (III)) are doing; i.e. it refers to a stance built in the interviewer’s question and the inadequacy of that stance, and then provides an elaboration in form an account. In addition to this, the interviewee’s stance linguistically builds upon the design of the question and modifies it in order to display a divergent stance toward the stance in the question. Consider diagraph (7) below.

Diagraph (7)

5 IR: The world has never had more treaties,
7 Never been more unstable,
9 IE: Well,
10 actually,
11 I think the world is a lot more stable,

49 In fact, Clift (2001, 270) shows that the adverb ‘actually’ is used turn and TCU-initially after claims. As she (Clift 2001, 272) further argues, this use of ‘actually’ indicates a highly confrontational stance with regard to the prior turn and further shows that the interviewee in example (9) interprets the question to contain a strong stance.
Diagraph (7) shows that the interviewee produces a stance that diverges from the
stance in the question but importantly, does so by linguistically engaging with linguistic
elements in the question (the NP the world and the comparative more). Moreover, the
slight differences in the answer are used for displaying the interviewee’s stance
differential (the change from the past tense to present tense and from unstable to stable).
Finally, the particle well, the adverbial actually and the stance marker I think display the
fact that the answer does not provide a preferred answer (an agreeing ‘no’ answer). In
sum, in order to describe stance taking one needs to describe the actions that interlocutors
do, how the linguistic elements function in the production of these actions, how the two
turns engage with each other and finally how the stances align with each other.

Subjective stances can also be expressed with varying degrees of evidentiality. Chafe
(1986, 262) defines evidentiality broadly as “a range of epistemological considerations
that are linguistically coded in spoken and written English,” but which also go beyond
grammatical marking and include ways by which speakers encode their attitudes toward
the reliability and source of knowledge. Chafe (1986) further claims that evidentials
function in different social situated contexts in three ways: they evaluate the degree of
reliability of knowledge, specify the mode of knowledge, and mark the contrast between
knowledge and expectation. Chafe’s work tackles the same linguistic forms as for
example Biber, Finegan and Conrad do. In other words, speakers and writers express
different degrees of evidentiality (reliability, belief, induction, sensory evidence, hearsay
evidence, deduction and expectation) with adverbs, modals and verbs.

Barton (1993) and Downing (2001) use Chafe’s definitions of evidentiality as a
starting point in their work. Barton (1993) considers evidentiality and stance in written
argumentation between academic writers and student writers, and shows differences in
the epistemological stance (i.e. the degree of knowledge) that these groups have or are
ready to express. She lists various linguistic items that can function as evidentials, such as
modals, sentential adverbs, conjunctions, and prepositional phrases. Downing (2001)
investigates the use of the adverb surely and how it co-occurs with subject personal
pronouns in the constitution of attitudes in British English. She (2001, 277) claims that
surely displays the speaker’s confidence of the fact that the content of her claim prefaced
by surely is true, reasonable and reliable (see lines 1 and 3 in example (11) below). In
other words, as Downing (2001, 254) suggests, surely smuggles in opinion and stance
through grammar and lexis.

(11) BBC2, Newsnight, Dec 17, 2003: Surely
IR: Jeremy Paxman, IE: Alistair Darling (033 / 1 / 1:53)

1 → IE: [{(H) Surely] y --
2 IR: [So the s-] --
3 → IE: Surely the longer,
4 th- the ^bigger question %i=s.
5 ^firstly to make sure,
6 (H) that in the ^longer term,
7 av[iat]ion ^does meet the costs of the environmental
8 damage.
9 IR: [#]
Downing (2001) further argues that the use of surely is triggered by a psychological event of “coming into awareness,” thus displaying the “surprisal” quality and difference in the status of knowledge between the speaker and the hearer. Downing thus hints at the intersubjective aspect of surely in interaction. Nevertheless, she (2001, 254) seems to back down from this argument and concludes that “surely in its main present use is subjective in meaning, at least partly grammaticalised and is the result of historical development through frequent use.”

Summarizing thus far, the above research mostly concentrates on the semantic categorization of stance and only occasionally refers to contextual factors. This can be problematic, because, as one of the primary claims in this study goes, stances are not just expressed by words and phrases with particular inherent evaluative meanings. As has been suggested above and will be shown in the following chapters, stance and stance taking is also an intersubjective activity, and thus managed also in other ways than the use of particular lexical or phrasal structures.

Channel (2000) takes a step away from pure semantic categorizations of explicitly stanced meanings. In her corpus linguistic approach she embarks upon examining how words and phrases receive positive or negative evaluative meanings as a result of their frequent use in particular pragmatic contexts. Her main agenda is to emphasize the importance of collocations in relation to evaluative lexis, especially as these collocations are used—often unconsciously—by speakers to express evaluative meanings. Example (12) supports Channel’s (2000) claim about the use of the noun regime in line 12.

(12) CNN, Sep 11, 2001: Taliban regime
IR: Wolf Blitzer, IE: Porter Goss (013 / 2 / 1:26)

1 IR: Chairman Go=s,  
2   let me,  
3   u=h,  
4   wrap up by asking you=,  
5   u=h,  
6   briefly a question,  
7   that Greta,  
8   ... u=h,  
9   asked Senator Biden,  
10  If in fact uh,  
11  Osama bin Laden was responsible,  
12  (H) And if the Taliban regime in Afghanista=n,  
13  .. is protecting him,  
14  .. harboring his organization,  
15  .. Should the US,  
16  go to war,  
17  not only against Osama bin Laden,  
18  ... but against Afghanista=n itself,
Channel (2000) suggests that in addition to the dictionary meaning of the noun *regime*, it displays the speaker’s disapproval of the government or system of authority so designated (Channel 2000, 46). This claim is proved by example (12) above, in which *regime* collocates with such proper nouns as *Osama bin Laden* and *Afghanistan*, and with the phrase *go to war*, which all invoke, at least in the Western world in that particular historical moment it was uttered (September 11, 2001), negative rather than positive connotations.

Fox and Clifford (1991) and Fox (Fox 2001) take a critical stance toward Chafe’s work and other related work on evidentiality. They claim that these studies look at how language encodes the ways in which “the speaker has come to know the proposition expressed by an utterance” (1991, 1), and by doing so, they focus on “the internal cognitive workings of speaker in his/her relationship to the factual world” (1991, 8). Fox and Clifford (1991), on the other hand, aim to give evidence of the ways in which evidential marking in English responds to the speaker’s social *authority* to make an assertion (ibid.). In the examples in their paper, they show that a speaker’s relationship to knowledge does not always correspond with the evidential marking that she uses. They also show how a single speaker produces one proposition with different evaluative markings, on different occasions and to different recipients. In sum, the authors do take into account the contingencies of the social context and how it affects evidentiality (cf. section 4.2.

The notion of ‘evaluation’ is frequently used as a near-synonym to stance. Research reported in Thompson and Hunston (2000) contains investigations of how ‘evaluations’ are expressed in discourse. Thompson and Hunston (2000, 2) argue that the expression of (the writer’s or speaker’s) opinion is an important feature of language. They use the term ‘evaluation’ to cover the expressions of attitudes and stances toward entities or propositions that the speaker/writer is talking about. The authors (2000, 14-22) acknowledge that some lexical and grammatical items are very clearly evaluative, but they also talk for a broader understanding of evaluation, claiming that evaluation is important also in three other major respects. First, evaluative acts express opinions which also simultaneously reflect and build up the socio-cultural value-system (Thompson & Hunston 2000, 6). Second, evaluative acts construct and maintain relationships between the speakers/writers and the hearers/readers (2000, 8). Finally, evaluative acts organize discourse on a metadiscursive level (2000, 10). The second claim is especially important for how stance taking is approached in the present study. This is because it acknowledges that ‘evaluation’ is something that happens between interlocutors.

Kärkkäinen (2003, 19) claims that some grammatical units that relate to the speaker’s commitment to a proposition have been grammaticized in discourse (see also Stubbs 1996). Such expressions of linguistic subjectivity include linguistic markers that index the speaker (e.g. the subjective first person pronoun *I* in line 29 in example (13) below) together with a particular cognitive verb (*think* in line 29); such markers frame the point of view or attitude that the speaker encodes in an utterance (lines 29-33). (cf. Finegan 1995; Stubbs 1986; as cited in Traugott & Dasher 2002, 20)

---

50 “A method or system of government” or a more derogatory use to refer to “a particular government” (Concise Oxford Dictionary, 1990).
As Kärkkäinen (2003) notes, many recent studies that examine spoken language suggest that grammar is extensively used to express subjective views and to display subjective identities, feelings and attitudes (e.g. Scheibman 2001, 61-62; Thompson & Hopper 2001, 53). But much of this work has concentrated on stance (or related phenomena) as a speaker’s attitude toward a proposition in her own utterance, clause or sentence, as is also argued by Kärkkäinen, Sorjonen and Helasvu (to appear) and by Kärkkäinen (under review). However, as is shown in the ‘positioning’ paper (II), in which example (13) is also analyzed, the epistemic stance marker is also a linguistic element that is routinely used for producing the denial + account action combination. In other words, it is not used just to express the interviewee’s subjective stance but rather it is a linguistic element that is used for responding to a position set up in the question and aligning with that position.

In sum, the linguistic research has primarily concentrated on investigating the semantic categories and individual linguistic forms and structures that express the speaker’s stances. Such an approach thus presupposes that taking a stance is a subjective act (cf. Thompson & Hunston 2000, 21). What I intend to make even clearer in the following chapters is that stance taking is also an intersubjective activity, which the above research does not consider. However, my intention is not to undermine the above work or suggest that it should be done in a different manner. On the contrary, the above work provides valuable insights into the lexical and phrasal practices of expressing a stance. These linguistic practices can be used as resources for locating sequences and activities of interactional and intersubjective stance taking that take place around these linguistic items.

### 4.4 Summary

As has hopefully become clear in the above sections, stance has been approached and investigated in very different ways in linguistics and sociology. The aim of the present work is to adopt a functional-interactional viewpoint that combines the benefits of conversation analysis, functional linguistics (especially the functional linguistic ‘theory of stance’) and interactional linguistics, and with the help of this framework to investigate how interlocutors take stances in talk-in-interaction by relying on various linguistic, interactional and embodied practices in interaction align and engage with stances taken in prior discourse. The fundamental aim is to search for patterns and tendencies of stance taking in which the actions and the language are closely connected. Alignment is defined
as a process of calibration between stances. It is both a dialogic linguistic activity (cf. Du Bois p.c.) and displayed by particular ways of doing actions and designing turns. Engagement refers to lexical and syntactic engagement between linguistic units. Interlocutors engage with each other for various reasons, but the two major reasons probably are, first to identify a place in prior talk with which a speaker wants to engage, and second to engage with the stance in prior talk and to express one’s stance in relation to that.

In section 4.3 above it has become clear that traditional linguistics work primarily concentrates on single words or phrases and how they represent the speaker’s stance toward something. Traditional linguistic work does not investigate stance taking, i.e. how interlocutors negotiate and co-construct stances in interaction. Consequently, I have outlined a functional-interactional approach which aims to address questions that are important for understanding stance taking. First, what is the relationship between the actions that the interlocutors produce and the stances that they take? Second, are there any linguistic, interactional or embodied patterns (or combinations of them) that interlocutors use for taking stances? Third, how might the (sociocultural) context be used as a resource for taking stances and oriented to in situ?

It is worth noting that the term ‘functional-interactional approach’ has been coined very late in the present project. It is not for example used in the ‘theory’ paper (I). The main reason for this is that the findings of stance taking in this work are based on the empirical analysis of data. Only after the analysis has it been found out that there are many practices that organize stance taking, practices which interlocutors orient to when they take and understand co-participants’ stances. In other words, in contrast to conversation analysis, functional linguistics and interactional linguistics, the focus is neither on the description of an action or an activity nor on a linguistic phenomenon alone. Rather, the different approaches together are used to investigate different stance-taking activities in a way that supplements their individual aims.

In the following chapter I discuss the previous linguistic and interactional work which has focused on news interview data and discuss how the functional-interactional approach can be used to examine stance taking in particular.
5 Intersubjective stance taking in news interviews

There is a large amount of research on the televised media and news interviews. In this chapter I give an overview of the linguistic, social psychological and conversation analytic research that is relevant for building an understanding of how stances are constructed intersubjectively in news interviews. I suggest that this body of work can be divided into two categories, in which the linguistic and social psychological work considers stance as a subjective phenomenon (primarily related to interviewee evasion as a subjective action), whereas the conversation analytic research examines the co-participants’ practices, actions and activities with a view to how they are managed intersubjectively in the course of the interview. I suggest that the linguistic and social psychological research on news interviews has overlooked the interactional contingencies which contribute to the co-participants’ construction of actions and stance taking. Conversation analytic research on news interviews, on the other hand, has given scant attention to the morphosyntactic and prosodic aspects of the co-participants’ construction of stances, or even to the intersubjective engagement between the interviewers’ questions.

51 A critical approach is adopted by for example van Dijk (1985) and Fairclough (1995). They focus on unraveling the connection between language use and the unequal power relations/ideological assumptions in media discourse, and examine how interviewing practices are shaped by and shape wider social and cultural shifts (Fairclough 1995, 22-23, 54). Simon-Vandenbergen (1996) in her systemic-functional study investigates the communicative strategies and the ways in which politicians in news interviews use language for gaining or retaining social power and for creating an image of an intellectually powerful and reliable person. Thornborrow (2002) combines the issue of ‘power’ with an interactional analysis of news interviews, and examines ‘power’ as a “locally produced and analysable phenomenon in the social interaction between participants in talk in institutional settings” (Thornborrow 2002, 138). In addition to this, other interesting and related work has been done in pragmatics (Jucker 1986) and in text linguistics (Felderer 1986), on rhetoric in the media (Atkinson 1984), radio talk (Hutchby 1991; Hutchby 1996; Hutchby 1999), television debates (Leudar & Nekvapil 2000), and on news interviews in France (Léon 2004) and in Finland (Berg 2001a; Berg 2001b; Berg 2003a; Berg 2003b; Kajanne 1999; Kajanne 2001a; Kajanne 2001b; Nuolijärvi 1994; Nuolijärvi 1999; Nuolijärvi & Tiittula 2000).
and the interviewees’ responses. Thus, I claim that the functional-interactional approach to stance taking outlined above is able to provide new insights into how interviewers and interviewees intersubjectively take stances in news interviews. This theme is developed by a more detailed analysis of examples of stance taking in news interviews in the ‘theory’ paper (I), the ‘positioning’ paper (II) and the ‘denial’ paper (III).

In news interviews, the primary task of the interviewers is to elicit relevant answers from the interviewees to the questions they ask (Clayman & Heritage 2002, 7). A great deal of the prior research on news interviews starts from the question-answer activity and aims to unravel, for example, how questions and answers are designed syntactically, how questioning and answering is managed sequentially, and whether the interviewees’ responses provide direct and relevant answers to the interviewers’ questions (Clayman & Heritage 2002). Two different types of definitions of what counts as a question and what counts as an answer can be found: one that focuses on the role of syntax in questions and answers, and one that focuses on the contingencies of interaction.52

The first one, which Clayman (2001) calls an analyst-oriented approach, and which is more closely discussed in section 5.3, tends to rely on the analyst’s perspective in order to code and classify, among other things, different instances of how interviewees (especially politicians) refrain from giving direct answers to the interviewers’ questions (see for example Bull 1994; Bull & Mayer 1993; Harris 1991). This work, however, gives little attention to the social organization and sequentiality of interaction, and how these contribute to interviewee equivocation. Thus, similarly to the research described in section 4.3, this work presupposes that stance (e.g. being evasive) can be attributed to the participants’ subjective aims, which they are claimed to express by the use (or lack of use) of particular linguistic forms. The second approach, on the other hand, which Clayman (2001) calls participant-oriented approach, describes the interactional practices which interviewers and interviewees rely on in order to manage for example hostile and adversarial questioning or in order to produce evasive answers. This approach, which relies on the conversation analytic methods (see section 4.1 above), has addressed various interactional phenomena (such as neutralism, adversarialness, hostility, evasion) that provide important insights into how stance taking, as it is manifested in news interviews, can be understood.

News interview data provide fruitful and interesting objects for the examination of stance taking, because they are bound to be full of various stance expressions and recurrent practices, actions and activities of stance taking. In the following sections I first (section 5.1) give an outline of news interviews as an institutional form of interaction, whose turn-taking organization and turn-allocation practices differ in many respects from mundane conversation. In section 5.2 I consider the central findings made in conversation analysis regarding practices, actions and activities in news interviews. This section provides important background information for the original articles I, II, III, and V. Finally, in section 5.3 I concentrate on the phenomena of interviewee equivocation and evasion. I suggest that previous research on these phenomena has given fairly little attention to the recurrent and patterned intersubjective linguistic practices of stance taking and the ways in which the interviewers’ and interviewees’ turns engage with each

52 For a paper that examines questioning from both sides, see Heritage and Roth (1995).
other, even when the interviewee seems to be evading the question. These questions are dealt with in greater detail in the original papers I, II, and III.

5.1 News interviews as institutional interaction

The difference between news interview interaction (and other task-related or institutional forms of interaction) and everyday talk has been discussed and addressed to a great extent in the conversation analytic literature. There are two major reasons why it is important to address this question here. First, the participants orient to and construct the institutionality of news interview interaction intersubjectively. Second, institutionality constrains the ways in which interviewers and interviewees can express stances, even to the degree that practices are used in news interviews which are rare in mundane talk (see the discussion of the NEG + POS practice in articles II and III).

The interactional phenomena described above in section 4.1 (e.g. the turn-taking organization and assessments) have mostly been investigated relative to everyday conversation. This is due to the fact that many conversation analysts see everyday conversation as the primordial form of talk-in-interaction, in which talk is “not subject to functionally specific or context-specific restrictions or specialized practices or conventionalized arrangements” [original italics] (Schegloff 1999a, 407). Very early on conversation analysts noticed that different speech-exchange systems are organized in different ways, i.e. the turn-taking or the turn-allocation techniques in ceremonies, debates, trials, doctor-patient consultations, trial examinations and news interviews differ from everyday conversation (Sacks et al. 1974, 729). Subsequently, the term institutional interaction has been widely used for referring to the use of talk in a variety of professional and institutional settings. The aim of this work has been to show how interactions are organized as institutional interactions (ten Have 1999, 162). As Heritage (1995) says, this can be methodologically challenging, because in addition to the analysis of the interlocutors’ conduct and the organization of their activities, one must show how the participants orient to and display their understanding of the institutional task at hand through the practices and actions they produce (Heritage 1995, 407).

Consequently, conversation analysts have considered institutional forms of talk as “versions” of everyday talk. This view has been questioned and criticized by some scholars (Billig 1999b; Hester & Francis 2001; McElhinny 1997; Thornborrow 2002, 135). Some claim that this dichotomy furthers the idea of everyday interaction being “normal” and institutional interaction somehow “deviant”. This criticism has also pointed out that it is very hard to draw a line between “ordinary” talk and “institutional” talk, and that many institutional forms of talk are in fact very conversational in nature.

53 See also Schegloff (1992c, 1296).
The issue of “normalcy” and “deviance” aside, Sacks, Schegloff and Jefferson (1974, 730) already noted that these two types of interaction can be considered as two polar extremes on a linear array and that different interactional types can be located at different points on such a continuum. In other words, Sacks et al. (1974) acknowledged that the borderline between everyday and institutional talk can be hazy. Consequently, for example, family meal-times can be considered as a type of institutional interaction, because as Thornborrow (2002, 4) suggests—based on Ochs and Taylor (1992) and Blum-Kulka (1997)—family members can engage in different roles and identities during meal-time conversation, which affect the organization of the interaction.

Iedema (2003, 73-80) criticizes conversation analysis for predetermining the different types of interaction, when it at the same time claims to avoid prior theorizing. Iedema (ibid.) further claims that Schegloff presupposes “an individual agent” in ordinary conversation to be in the position to affect the overall organization of interaction, turn-taking and action production, whereas in institutional interaction, it is a “structural agent” that “forces” the individuals to act in certain ways. Iedema’s criticism seems to be somewhat of a misinterpretation, because in contrast to his claims, conversation analytic work in fact acknowledges that individuals in the different types of institutional forms of interaction accomplish their roles in those interactions (cf. Drew & Sorjonen 1997, 97) and contribute to the course of interaction on a turn-by-turn basis (Clayman & Heritage 2002, 13). These previous studies and findings on institutional interaction show this as well, and it is hardly the case that Schegloff’s view differs from this.

Even though conversation analysis refers to institutional interaction as a version of everyday interaction, this should not simply be understood as a request to compare these two types of interaction. Rather, by invoking this dichotomy, the aim has been to set up a foundation for studying the practices and actions and the social organization of institutional interaction in its own right (cf. ten Have 1999, 163; ten Have 2001, 3-4), and to explore how interlocutors use particular practices for doing specific tasks or maintaining particular roles and identities in institutional interaction (cf. Heritage 1995). Heritage (1997, 164) suggests some possible places where ‘institutionality’ can be investigated. These are turn-taking organization, overall organization of interaction, sequence organization, turn design, lexical choice, and epistemological and other forms of asymmetry, i.e. the very same issues that were mentioned in section 4.2.2 to be important for the analysis of everyday conversation.

A great deal of the institutional forms of interaction are based on sequences or activities in which one participant is mainly responsible for questioning, and the other participant for answering or responding to the questions. This suggests that the set of actions that interlocutors produce in institutional talk are limited. This type of turn-type pre-allocation is asymmetrical and reflects the roles that the participants have and the identities they locally construct in the interactional situation (e.g. doctor/patient, journalist/guest, and employer/applicant). The turn-type pre-allocation also constrains them in terms of what actions they are expected or allowed to produce (see also chapter 7). Participants in different institutional interactions can of course anything they want, but in such a case the interactional situation would turn into something else. For example, Schegloff (1992b) studies a famous example of when the turn-taking rules and turn-type
pre-allocation were breached in a news interview in which Dan Rather interviewed George Bush, the former President of the United States of America in January 1988. In that paper Schegloff (1992b) shows that when in that particular interview the “rules” regarding the news interview speech-exchange system and the turn-type pre-allocation are breached, the interview becomes disorganized and it becomes impossible to bring it to a successful conclusion.55

As was mentioned above, prior work has shown how the participants’ behavior is shaped by the constraints in institutional talk and that it is organized in a notably different way than everyday talk. The most evident differences between everyday and institutional conversation are the following:

1. In contrast to everyday talk, the participants in institutional talk have pregiven, but asymmetrical rights in two different ways: First, the types of turns and actions they are “entitled” to produce are different. Second, the distribution and allocation of the turns between the participants is different.
2. In contrast to everyday talk, the participants in institutional talk are allocated more than one turn-constructional unit.
3. In contrast to everyday talk, participants in institutional talk have different identity or participant roles (e.g. interviewer / interviewee), which they are seen to adopt, construct, and “talk into being” (Clayman & Heritage 2002, 21) in that interactional situation.

Language itself and the linguistic resources that the interlocutors rely on have important roles in how participants manage their institutional roles. As Drew and Sorjonen (1997, 92) claim,

the study of ‘institutional dialogue’ is [...] the study of how people use language to manage those practical tasks, and to perform the particular activities associated with their participation in institutional contexts [...]. Thus when investigating institutional dialogue, we are focusing on linguistic resources at various levels – lexical, syntactic, prosodic, sequential, etc. – which are mobilized for accomplishing the interaction work of institutions.

Drew and Sorjonen (1997, 99) further point out that for example the particular uses of pronouns as indexical person references, and lexical choices more generally, display the participants’ understanding of the situation they are in. In addition to this, as we will see in the following section, various grammatical forms are also often linked to the participants’ roles in institutional talk. For example, different types of interrogative syntax are available for participants in those types of interactions in which one of the interlocutors is responsible for questioning (e.g. interviewer, doctor, teacher). Consequently, as Drew and Sorjonen (1997, 102) suggest, the study of talk in institutional settings naturally benefits from a vantage point in which the linguistic details are examined relative to and together with the sequential context in which their occur. In the

55 See also Clayman and Whalen (1988/89).
present work this becomes evident especially in the ‘denial’ paper (III) in which it is argued that the linguistic NEG + POS pattern is frequent because of the institutionality of news interview interaction.

As regards news interview interaction in particular, conversation analysis has approached it “as a form of spoken interaction and examines the recurrent communicative practices that constitute it” (Clayman 1988, 474). As Heritage (1985, 95) points out, such an approach is also of central importance in the investigation of how opinions (cf. stance and stance taking) are interactionally generated in news interviews. News interview interaction has been shown to have its own distinct organization. This cannot only be perceived in the differences between the ways in which news interviews and everyday conversation are socially organized, but, as was suggested above, also when that organization is breached.

The reason why it is important to here acknowledge the difference between everyday conversation and institutional talk is that this difference influences, structures and constrains the way in which interlocutors in news interviews can and do take stances. Although stance taking as it has been outlined in chapter 4 above has been studied to some extent in everyday conversational data (Du Bois to appear; Kärkkäinen under review), stance taking has not been studied at the other end of the continuum of interaction types. For such a study it is important to have a closer look at the conventions and practices of conduct in news interviews, which will be turned to in the next section.

5.2 News interview interaction

In the following I give an overview of the conversation analytic research on news interview interaction. The section considers three central aspects of that work and how they relate to stance taking: the overall organization of news interviews (section 5.2.1), turn design features of questions (section 5.2.2) and the features of answering, especially interviewee evasion and equivocation (section 5.2.3). Finally, I suggest that the conversation analytic work can be supplemented, as is done in the original papers (I, II, III, and V), with a more detailed linguistic and grammatical analysis of the participants’ conduct, and moreover, with a special inspection of how the participants linguistically engage with each other when they construct their stances. This functional-linguistic view is used as the starting point in the original articles I, II, and III, which investigate the activities of positioning and alignment, and how these activities are accomplished in news interview interaction. It is also adopted in the ‘identity’ paper (V), which shows how an ethnic identity is locally constructed in a televised interview.

5.2.1 The sequential organization of news interview interaction

News interviews are publicly broadcast interactions, in which the journalist’s primary task is to ask questions from the guest, who is usually a politician or other type of public figure (e.g. Scannell 1991, 4). Through this questioning/answering activity, journalists
and their guests publicly and jointly discuss, negotiate and debate topical and often controversial topics. Conversation analysis considers news interview interaction emergent, even though the questions and the topical agendas have usually been decided beforehand. The rationale behind this is that there are tacit social conventions in news interviews, which construct news interview interaction as an organized social institution (Clayman & Heritage 2002, 6). It is through these conventions that the participants construct news interview interaction as an intersubjective accomplishment. In addition to this, as Heritage (1985) argues, many of the recurrent practices in news interviews display an orientation towards the overhearing audience.

One of these conventions relates to the turn-taking organization of news interviews (Clayman 1988; Clayman 1991; Greatbatch 1985; Greatbatch 1988; Greatbatch 1992; Heritage 1985; Heritage 2003a). As conversation analytic research on everyday conversation shows, one speaker is basically allocated one turn-constructional unit at a time (see section 4.1). However, in news interviews, in which the turn-taking system and allocation of turns differs distinctly from everyday conversation (e.g. Heritage & Greatbatch 1991), a different set of “rules” apply. First of all, the interviewers are responsible for opening and closing the interviews, introducing and organizing topics, and allocating the turns in multi-party interviews. Second, news interviews have an asymmetrical turn-taking order, in which particular turn-types have been pre-allocated to particular participants: interviewers confine themselves to producing actions that are at least minimally recognizable as questions, and interviewees are expected to give relevant answers to the interviewers’ questions (Clayman 1988; Clayman & Heritage 2002; Greatbatch 1985; Greatbatch 1986; Greatbatch 1988; Heritage 1985; Heritage & Greatbatch 1991; Heritage & Roth 1995; Schegloff 1988/89). It is noteworthy that also

---

56 The issue of pre-planned questions and topical agendas, and their role in news interview interaction has not been considered in conversation analysis to a great deal. However, Tiittula (1999, 124) has been able to follow on location the planning and setting up of a live panel discussion, and she notes that the discussion in the program faithfully followed a pre-written script planned by the hosts. She also gives an interesting example of what happens when the planned script does not hold and the guests continue the debate that the host has tried to interrupt and stop.

57 Thornborrow (2002: fn 4, 106) argues that the term “overhearing audience is something of a misnomer.” She further suggests that “ratified but indirect recipients of the talk is perhaps a more accurate description” of viewers, who indeed watch the news interviews at home and not at the studios (ibid.).

58 Some interview programs tend to breach the “rules” of turn-taking and pre-allocation. A good example of this is CNN’s Crossfire. As Firth (2003, 133) suggests, Crossfire represents a program in which “all parties are cognizant of the modus operandi of the programme, a mode that calls for the production and performance of hostile questioning, highly interruptive turn-taking, and a robust defence of one’s political viewpoint. Voices are raised, fingers are pointed, ‘questions’ are mocked, and ‘answers’ are openly castigated.” In other words, Crossfire is an example of an argumentative interview format, in which also interviewers can explicitly express viewpoints and stances that are based on political viewpoints. Clayman and Heritage (2002) claim that such programs do not belong to the ‘news interview’ genre. Crossfire is nevertheless included in my corpus and also examined in detail in the original articles.
interviewees orient to this pre-allocation of turn types. They tend not to start their answering turns until the interviewer has completed her turn with a clearly identifiable question. The fact that the interviewees withhold the production of the answer until a question has recognizably been produced gives the interviewers the right and the possibility to build almost any kinds of arguments and statements into the question (within the limits of neutralism, see below in section 5.2.2.).

The sequential organization of news interviews not only constrains the participants in terms of their actions, but also structures their conduct (Heritage & Greatbatch 1991, 95) and provides them with very different resources for dealing with each other (Clayman & Heritage 2002, 96). For example, interviewers have the right to interrupt the interviewee if a “minimally adequate ‘answering’ component has been uttered” (ten Have 1999, 166) or if the interviewee is not giving an appropriate answer. As Piirainen-Marsh (2003) shows, some interviewer interventions tend to occur at or near possible completion points. The function of the interviewer interventions is to foreground some aspect of the interviewee’s utterance, provide some interpretation of it, or to oppose it. Interviewers can also intervene in the interviewee’s turn by producing an anticipatory or a collaborative completion that connects lexically or syntactically to the interviewee’s turn (ibid.). This type of action can both try to bring the interviewee’s utterance into a closure and frame the next question as symbiotic and responsive to an argument that the interviewee has just made in the preceding utterance (ibid.).

Although the present work does not specifically address questions related to the sequential organization of news interviews, it is important because it helps understand that stance taking and the ways in which interviewers and interviewees construct stances in news interviews are markedly different from everyday talk. One such difference is the possibility for the participants to construct long and complex turns. This is considered in greater detail in the following section.

### 5.2.2 Turn design features of questions

Turn design, or the organization of multi-unit turns out of several turn-constructional units, which are sequentially relevant to each other, is one of the central concerns of conversation analytic research (see section 4.1). The question of turn design is also an

---

59 This is related to the lesser interactional importance of TRPs (transitional relevance places) in news interviews compared to everyday conversation, as a place where possible turn transition can occur. In other words, since the participants tacitly share the understanding that the interviewees do not begin to answer the question before a clearly identifiable question has been produced, the interviewers have the right to produce several TCUs (with neutralized transition relevance places) without the possibility that the interviewee attempts to take the turn. It is not the possible completion point that interviewees “analyze,” but other turn design features that act as cues for a closure of the interviewer’s question.

60 These could also be called ‘second position questions’ in that they can be designed to clarify the interviewee’s prior utterance.
important part of the analysis of stance taking in news interviews, both regarding the practices by which interviewers design difficult questions for the interviewees (see the ‘positioning’ paper II) and in terms of how the interviewees align with the interviewers’ questions (cf. the discussion of action combinations in the ‘denial’ paper (III)). The question of turn design and its relation to intersubjective stance taking is dealt with in greater detail in the ‘positioning’ paper (II) and the ‘denial’ paper (III), but some of the findings are also discussed below.

Many studies have examined news interview questions by starting from particular linguistic units or structures (e.g. interrogative syntax) and the role they play in the production of questions. Heritage and Roth (1995) provide a detailed analysis of how questioning is accomplished in news interviews. They start by considering the syntactic form of questioning, but importantly, also consider other ways of how interviewers’ turns can be seen to accomplish questioning. They (1995, 9-10) mention that interviewers can use for example imperative (directive) utterances (although very rarely) and B-event statements,61 which tend to be produced with final rising intonation, for accomplishing questioning. Heritage and Roth (1995, 13) also claim that questioning is related to turn organization and show how questioning can be accomplished across multiple TCUs, in what they call “question delivery structures.” They (1995, 39-41) further claim that the differences between various “question delivery structures” can be seen in the types of activities that are accomplished in question prefaces, i.e. in those segments of interviewer’s turn that precede the actual questioning TCU. All in all, these non-interrogative practices of questioning prove that questioning is not just a syntactic or prosodic phenomenon, but largely dependent on the ways in which interlocutors design their actions and activities.

Interviewers tend to be careful in how they construct their questions. This is also reflected in the following comment by Sam Donaldson, the former White House correspondent for ABC.

The wording of questions is very important. If you say to a president, “Would you care to comment on X,” he can always answer, “No, thanks,” or he can say anything he wants to and call it a comment. (Donaldson 1987, 7)

One of the most challenging tasks that interviewers face in interviews, and which is also directly relevant for stance taking, is how to force the guests to give direct answers, but at the same time to remain impartial and neutralistic toward them. Indeed, one important and central feature of news interviews is the expectation that interviewers

---

61 “B-event statements are declarative utterances in which the speaker formulates some matter as one to which the recipient has primary access. At the minimum, such an utterance makes a recipient’s confirmation or denial relevant in the next turn.” (Heritage & Roth 1995, 10) For B-events and other types of events related to who has primary access to information in therapeutic discourse, see Labov and Fanshel (1977).
maintain an impartial or neutralistic stance in the course of interviewing. Interviewers should refrain from making antagonistic or supportive statements relative to the interviewees and their arguments. Nor should the interviewers in any way assess or affiliate with the topic (or Stance Object) of the interview (Clayman 1988; Clayman 1992; Clayman & Heritage 2002, 126-131; Heritage & Greatbatch 1991). As Clayman and Heritage (2002, 188) claim, the interviewer’s questioning action should not assume a substantive position as to whether it agrees or disagrees with the interviewee. This is indeed the reason why interviewers work hard to package their actions as “questions,” and may invoke this packaging to defeat interviewee claims that they are pursuing some kind of agenda of their own (Clayman & Heritage 2002, 188).

The term ‘neutralism’ does not refer to the information that the interviewers’ utterances contain, nor does it mean that the interviewers are absolutely neutral towards the interviewees and their statements (Clayman & Heritage 2002, 151). On the contrary, interviewer questions are fraught with encoded beliefs, standpoints, challenges, accusations, judgments, opinions, presuppositions and so on. The term ‘neutralism’ refers to the various practices interviewers rely on in order to display a “neutralistic stance” (Heritage & Greatbatch 1991, 107) or construct themselves as without attitude (Potter 1998). The interviewers “write out” their own subjectivity and display a neutralistic stance towards the topical content and their guests by indeed formulating their turns almost exclusively as questions, but also by attributing statements to third parties (Clayman 1988, 482; Heritage & Greatbatch 1991). To a large extent, interviewers also refrain from using any affiliative actions or linguistic forms (e.g. newsmarks, assessments, news receipts, continuers) which would display (dis)affiliation with their guests (Heritage & Greatbatch 1991, 117-118).

Interviewer neutralism is, however, not only implemented through the interviewer’s actions, but is also strengthened and validated by the interviewee’s compliance with the turn-taking system and the turn-type pre-allocation. For example, the fact that the interviewees do not take the floor before the interviewer has recognizably completed the questioning turn shows that interviewees treat the interviewer’s question as to be coming from a formally neutralistic position (cf. Clayman 1988). Thus interviewer neutralism is very much a collaborative and interactional achievement, and it bears an important role in the intersubjective construction of stances in news interviews.

However, in spite of the apparently neutralistic posture, interviewers can exert pressure on their guests in various ways. One discursive strategy for doing that is called a ‘formulation’ (Berg 2003a; Heritage 1985; Thornborrow 2002). Interviewers use

---

62 Interviewers are also—at least in Britain—legally coerced to remain objective and impartial. However, as we will see, neutralism is also “enacted” and oriented to by the co-participants in the interview situation.

63 Conversation analytic work has used this term ‘third-party attributed statements’ I have used the term ‘third-party stance’ in the original paper and the present summary.

64 See also Hutchy and Wooffit (1998, 152-153).
‘formulations’ as a resource for controlling and directing the topical agenda in the news interview. A ‘formulation’ is the third part of a sequence (i.e. question + answer + formulation). It sequentially follows and builds upon an interviewee answer by producing a gloss or a summary of the central claim in the answer (Heritage 1985, 106). Usually a formulation displays an interpretation or an understanding of the interviewee’s (possibly vague) answer and thus projects either a confirmation or a disconfirmation from her.

Formulations can also be a resource for the interviewer to redirect or refocus the topical agenda. In Finnish news interviews, formulations tend to be declarative (84% of the examples) (Berg 2003a). Thornborrow (2002, 97) shows that formulations are often produced with the help of a prefacing frame that contains the following recurrent syntactic pattern: [(so) + you + verbal/mental process token], e.g. ‘so you’re saying that’ or ‘so you’re suggesting that’. According to Thornborrow (2002, 97), the purpose of these syntactic forms is to “tie the upcoming formulation to what the interviewee has just said in the prior turn by explicitly attributing the content of the interviewer turn to the previous speaker.” Formulations are also directed for the overhearing audience, because they foreground and emphasize some aspects of the interviewee’s answer and thus help the audience to interpret, infer or construct an understanding of the interviewee’s answer (Heritage 1985).

Interviewers can also put pressure on their guests more forcefully by adopting an adversarial stance (Clayman & Heritage 2002: 188). Indeed, they can incorporate stances in their questioning turns by adding particular agendas and topical domains which the interviewee is expected to address and comply with in the answer (Berg 2003a; Clayman & Heritage 2002; Heritage 2003a; Heritage & Roth 1995; Nuolijärvi & Tiittula 2000). These “statement turn components” (Greatbatch 1988, 407) are normally given in the question preface, i.e. prior to the actual question in the interviewer’s turn. They give the interviewers “room to maneuver” and allow them to “construct a context of their own choosing for the question they are about to put in play” (Heritage 2003a, 65). These statement turn components have a very special function for stance taking. As is shown in the ‘positioning’ paper (II) and the ‘denial’ paper (III) they can be used by interviewees as a reservoir of linguistic resources for constructing a responsive turn.

Question turns can also embody presuppositions and/or assert propositions which pose interviewees in self-contradictory terms, or in disagreement with themselves or other important and significant people (Heritage 2003a). Interviewers can do this by raising a particular contextual background65 for the question in a question preface. Interviewers can also, although rarely, confront the interviewees by accusing them of being directly accountable of something (cf. ‘Why did you X’ or ‘How could you Y’ type questions), or by producing a negative evaluation of an interviewee deed or an event which she has been involved in (Berg 2001b; Berg 2003a; Clayman & Heritage 2002, 221).

Questions can also be differently designed in terms of how broadly or narrowly they constrain the type of the following answer (Heritage 2003a, 69). Wh-questions in general establish a fairly loose agenda. However, negative interrogatives (e.g. “Isn’t it...,” “Don’t you...,” and “Won’t you...”) strongly prefer an agreeing yes-answer, tag questions prefer an answer that confirms the claim in the statement part of the question (Heritage 2002;

65 Cf. the discussion of ‘context’ in section 4.1.
Heritage 2003a), and questions that contain negative polarity terms also tend to prefer particular answers (Clayman & Heritage 2002, 212). Overall, interviewers use these features—in spite of the principle of ‘neutralism’—for embedding hostile, harmful and adversarial stances in the questions.

It is important to point out that even though conversation analytic research has investigated phenomena such as ‘neutralism’, ‘adversarial or hostile questioning’ and ‘evasion’, which that are closely related to ‘stance,’ the focus of these studies has mostly been on the turn design features of these phenomena and not so much on how they relate to the immediate interactional context and the co-participants actions and language use in the previous or immediately subsequent turns. Nevertheless, the above conversation analytic research on the features of question design that put pressure on the interviewees is very closely related to the notion of ‘positioning’ discussed in the ‘theory’ paper (I) and the ‘positioning’ paper (II). However, the term ‘positioning’ supplements the above work in the following ways. First, it refers to the way in which the interviewer builds up a stance within the question that the interviewee needs to take into account in her answer. Since various linguistic elements play an important role in stance taking, it is important to analyze how they figure in the ‘positioning’ activity. Second, the focus is not on individual turn statement components (cf. the conversation analytic view above) but on how these components together are used to build up a stance in the question. Third, ‘positioning’ is closely related to the idea of forward-type intersubjectivity. In other words, the position that is set up for the interviewee is specifically designed for that interviewee. Fourth, ‘positioning’ should always be considered together with how the interviewee engages and ‘aligns’ with the question and the position (or stance) therein. The relationship between the interviewer’s question and the interviewee’s answer is rarely addressed in conversation analytic research. The following section discusses previous linguistic, social psychological and conversation analytic work on interviewee answers. This body of work has primarily concentrated on interviewee equivocation. Therefore, I argue that a closer analysis of the linguistic and interactional relationship between the question and the answers is needed.

5.2.3 Turn design features of answers—an intersubjective view to interviewee equivocation

Of course, presidents don’t always agree with the way questions are framed [...] but when they don’t that too can be most revealing. (Donaldson 1987, 8)

The above observation by Sam Donaldson is most appropriate and telling. Conversation analytic work has provided descriptions of the ways in which interviewees answer in news interviews. Since questions project answers as relevant next actions, interviewees are bound by this interactional constraint to answer the question. Moreover, they are expected to give a direct and unequivocal answer. Failure to provide an answer is noticeable. It would also present the interviewee in a strange light and is usually met with the interviewer’s repetition of the question. There can be occasional departures from this
organization in the course of the interview, but these are usually met with overt sanctions by the interviewer, after which the participants quickly return to the ordinary chain of events.

Even though interviewees primarily give straightforward answers, they can also design their answers so that they simultaneously address the question as much as possible, but also avoid making claims or admitting to charges that would be harmful for them. In other words, interviewees design their answers with special prudence by trying to balance between the interviewer’s sometimes confrontational questions and the broader socio-cultural beliefs and audience expectations.

Interlocutors usually identify an answer on the basis of its sequential placement—it comes after a question (cf. Sacks 1992a, 49-56; Weber 1993, 14)—and based on the action that it does—it answers or responds to a question (Clayman 2001, 407). However, as is discussed in this section below defining an answer—especially answering in news interviews—is not so simple. In the following I concentrate on and give an outline of the linguistic, social psychological and conversation analytic work on interviewee evasiveness and suggest that the stance-taking framework and the view of stance taking as an intersubjective activity can supplement this prior work.

Within the social psychological paradigm (cf. the analyst-oriented approach mentioned above), Bull (1994) for example has examined interviewee equivocation by attempting to define and code replies and non-replies primarily on the basis of what the interrogative syntax of the question is (cf. Quirk et al. 1985) and whether the design of the answer corresponds with the way in which the question is syntactically formulated. Similarly, Bull and Mayer (1993) construct a typology of non-replies and use interrogative syntax as the factor that determines what kind of response is a reply or what is a non-reply. Furthermore, in a more linguistically oriented study, Harris (1991) similarly claims that “questions and answers can be analysed on the basis of syntactic and semantic features and coded into a framework which distinguishes between a direct answer and one which is not direct.” She further claims that the general belief regarding politicians’ evasiveness can be based on linguistic facts. On the basis of her data and coding criteria, indeed only 40% of the answers can be called ‘direct’.

Even though this is one way of identifying interviewees’ equivocal answers, it falls short of explaining the vast gray area in and the subtleties of the question-answer activity (Clayman 1993, 160; Clayman 2001, 440) at least in the following ways: First, by concentrating largely on interrogative syntax, it omits any considerations of turn design and what happens outside the immediate turn transfer area. Second, by ignoring the quality of the interviewers’ questions it assumes that they are by default neutral, appropriately designed and “right”, and thus unimportant and bearing no role in terms of how they set up a discourse context for the interviewee. By implication this means that interviewee equivocation is treated and understood to be a subjective action. The problem of the analyst-oriented approach is addressed by Bavelas et al. (1990), who suggest that one should not just assume that politicians knowingly evade questions. Rather, as the authors further suggest, one should understand that equivocation can be and often is a result of many pressures in the communicative situation. Third, the above work does not

---

66 See also Heritage (1995, 45) and Schegloff (1984b).
consider the question as an action, nor what the individual parts of the question (turn constructional units) are doing within the question and how they contribute to the understanding of the question. The problems of the view of evasion and equivocation that the analyst-oriented approach proposes are discussed in the 'theory' paper (I), the 'positioning' paper (II) and the 'denial' paper (III). The examples in the paper show that ‘evasion’ is not perhaps the best term to describe what interviewees do when they answer the interviewers’ questions. Rather it is shown that the design of the question and the stance that is encoded or evoked in it often sets up a difficult position for the interviewee. In spite of these difficult positions, interviewees still engage and align with the questions in different ways.

One reason for the oversight in the analyst-oriented approach is that it largely omits the contingencies of the sequentially unfolding interaction and how these affect the function and the meaning of the linguistic units. This of course relates closely to the analysis of stance taking in news interviews. In news interviews, as was already mentioned in the discussion of the turn-taking organization and turn allocation in section 5.2.1, the interviewees cannot just reply with single-word expressions, such as ‘yes’, ‘no’, ‘certainly’ or ‘of course’, which in the work described above would be considered appropriate answers to questions (cf. Bull 1994). Rather, in news interviews questions always project a broad and detailed answer. This is due to the fact that the interviewers’ and interviewees’ task is to negotiate, debate and discuss issues, and not just answer questions. News interviews are about finding a position where an interviewee stands in relation to particular topics and in relation to how the interviewer’s question has been formulated. The interviewees are in the limelight and, moreover, in a difficult spot, because the interviewers continuously exert pressure on them in order to dig out a stance or a position. Interviewers cannot do this by asking (syntactically) open questions, or easy and straightforward questions. This is why interviewers carefully design their questions so that they position (see the original articles I, II and III) and constrain the interviewees, and set bounds for them to design their answers.

In contrast to the work above, conversation analytic work adopts an interactional approach to answering. Conversation analysts maintain that defining an answer is not always straightforward; especially not in news interviews, where interviewees can formulate their answers in many different ways which accept, resist or reject preferences, presuppositions or topical agendas in the questions. Interviewees can deploy various resources for sidestepping (Clayman 2001), reformulating (Clayman 1993), and shifting for example the question’s topical agenda (Greatbatch 1986). These resources tend to be used especially by politicians or other public figures, who have their image, political agenda and reputation at stake (Clayman 1993; Clayman 2001; Greatbatch 1985). One practice that interviewees use for evading a question is the use of various kinds of repeats (e.g. word repeats) (Clayman 2001). However, recycling language is not just used for sidestepping the question’s agenda.67 As Heritage and Clayman (2002) have pointed out,

---

67 Cf. the aims of dialogic syntax (see section 4.3.2).
interviewees can also highlight an issue they want to address in the question by recycling the equivalent structure or phrase from the question.68

In spite of the close attention to the interactional context, conversation analytic work has examined interviewee evasiveness by concentrating primarily on how interviewers react to these evasive maneuvers in ‘third position’ (i.e. actions after answers) and thus try to keep interviewees ‘on topic’ (Clayman 1993; Clayman 2001; Clayman & Heritage 2002, 238-298; Greatbatch 1986). This provides convincing proof of equivocal behavior that is grounded on interactional evidence.

Indeed, it is a general belief that politicians often dodge and weave the questions they have been asked, and prefer to give equivocal and evasive answers. There is certainly no reason to doubt that this frequently happens. However, the question of what occasions interviewee equivocation has not yet been critically discussed.69 There are only passing references to those aspects of the interviewer’s question that bear on the evasive answer (e.g. above-mentioned hostility and adversarialness) and the impetus they give to and the role they have in interviewee evasiveness. For example Harris (1991, 78) at the outset of her article sets out to look at, among other things, “to what extent are politicians’ responses constrained by the syntactic and semantic properties of an interviewer’s question, and perhaps more importantly, to supply a particular answer,” but unfortunately does not address her own question in the remainder of the paper. Clayman and Heritage (2002, 188-237), on the other hand, refer to the impetus that the interviewer’s question has on an evasive answer only in passing.70

Consequently, much of the above, social psychological and conversation analytic work on equivocation and evasiveness attributes evasion to the interviewee, i.e. equivocation has been dealt with as the interviewee’s subjective action and stance, and as a phenomenon which is related to the lack of an appropriate answer to a particular interrogative form, or which is solely related to the design of the interviewee’s turn. In other words, since the individual participants are under a moral constraint to comply with norms, if the interviewee evades the interviewer’s question, it is the interviewee who is solely liable and can be held accountable for doing this action (cf. Clayman 1993, 160-161). Consequently, evasiveness is considered an individual and subjective action.

As I show in the ‘theory’ paper (I) and the ‘positioning’ paper (II), which discuss the stance-taking activities called positioning and alignment, interviewers apply both explicit and tacit strategies for setting up positions for the interviewees. By recognizing some of these strategies I have been able to discuss interactional reasons or the intersubjective “pressures” behind some instances of interviewee resistance and shed light on some

68 For everyday conversation see Schegloff (1996a) who notes that repeats often constitute an action of confirming an allusion.
69 It is important to note here, however, that although particular interviewer actions can be seen to occasion an interviewee’s evasive answer, no first action predetermines a subsequent action, because the second speaker can always choose what to do next. This point is also made in the ‘positioning’ paper (original article II).
70 However, see especially Heritage (2003a) and the analysis of examples 11 and 26, which are very close to the analysis of the ‘positioning’ activity in the ‘theory’ paper (I) and the ‘positioning’ paper (II).
practices by which ‘engagement in evasion’ is managed. The analyses of the two action combinations and the linguistic NEG + POS pattern in the ‘denial’ paper (III) also contribute to our knowledge of how interviewees, rather than bluntly evade the question, indeed engage and align with the stance that is encoded or presupposed in it.

5.3 Summary

To summarize chapter 5, previous research on news interviews has identified a number of practices by which interviewers exert pressure on their guests, and practices by which interviewees answer or fail to answer the question. Yet, it has given fairly little emphasis to

1) those cultural background assumptions which hostile presuppositions are built upon (cf. the question of context). As the ‘positioning’ paper (II) shows, interviewer hostility and adversarialness are not only achieved through particular uses of interrogative syntax. Rather, interviewers’ questions are fraught with different kinds of cues that help the interviewees (and the audience) to retrieve particular built-in presuppositions and assumptions in the questions. Interviewers use various grammatical, lexical and prosodic cues together with turn-organizational features, and against the contextual (socio-cultural) background, embed subtle stances in the questions;

2) the linguistic engagement and resonance between the questioning and the answering turns and what functions these have on the activities that the co-participants produce (original papers II and III, and especially the action combinations described in the ‘denial’ paper);

3) the action combinations and how these contribute to the alignment activities (original papers II and III).

The ‘theory’ paper (I), the ‘positioning’ paper (II), the ‘denial’ paper and also partly the ‘identity’ paper (V) all address the above points and thus contribute to and supplement the previous research on news interview interaction.
6 Embodiment and stance taking in interaction

This chapter gives an outline of research that has examined the roles and functions of gestures and other embodied practices in social interaction. The notions ‘embodiment’ and ‘embodied behavior’ are used to refer to non-verbal use of any body part (i.e. the use of gestures, gaze and body movements) as a resource for producing actions and activities in talk-in-interaction. Consequently, I make a conscious division between talk (verbal behavior) and embodiment (gesticulation, use of gaze, etc.). I am aware that many scholars consider this division unnecessary, because they consider both talk and gestures to together compose ‘embodied behavior.’ By doing this division I do not intend to separate ‘talk’ and ‘embodiment’, but indeed consider them together to compose communication and interaction. However, for a more linguistically audience the importance of gestures and gaze is perhaps not as clear as it is for the analysts of interaction. Therefore, by making an explicit distinction between ‘talk’ and ‘embodiment’ my aim is indeed to emphasize the importance of the relationship between language and embodied practices in talk-in-interaction.

In section 6.1 I give some background to the ‘gaze’ paper and briefly discuss how gaze relates to human ontogenesis and intersubjectivity. Then in section 6.2 I discuss the relationship between embodiment and talk. Section 6.3 considers briefly a pragmatic approach to gestures that investigates how gestures support or provide additional meaning to speech. Section 6.4 gives an overview of research that examines the relationship between embodied practices and turn-taking, turn design and action production. Finally, in section 6.5, I connect the previous interactional research on gesture and embodiment with the ‘gaze’ paper (original article IV), and, as is important for the present work, discuss whether interlocutors’ embodied practices can be seen to be related to the stances they take. It is worth mentioning here that the ‘gaze’ paper (original article IV) has only two examples of each interactional phenomenon it describes. Although having two examples of the same phenomenon is perhaps enough to indicate

---

71 Thus, I do not discuss gesture as an internal or psychological phenomenon. For a brief discussion of the psychological aspects of gesture, see Kendon (2001) and David McNeill’s work (e.g. McNeill 1992; McNeill 2000).
that what we have is a pattern, more work is required to get a thorough and detailed understanding of these phenomena.

In the following I do not claim that a particular gesture or any other embodied practice alone is expressive of a speaker stance. Rather, I suggest a way to examine how combinations of these practices contribute to the interlocutors’ intersubjective stance taking, i.e. I emphasize that in order to investigate the role that embodied practices have in stance taking, it is necessary to study how they are used together with the concurrent linguistic and interactional practices during a stance-taking activity. There are two ways to investigate this. First, one can look at whether a particular embodied practice (e.g. a gaze shift) recurs in a particular sequential position within the same action or activity across different examples. This type of embodied practice (notably the “cut-off” gaze in the ‘gaze’ paper) is closely linked to the utterance it precedes and can be seen to reflect the “meaning” in that utterance. The use of such a practice tends also to be occasioned by a prior action and to project a subsequent action. Second, one can look at whether an embodied practice recurs as a resource for the co-participants to arrive at joint understandings of an interactional activity. This type of practice (notably the ‘congruent gaze point’ and ‘mutual gaze’ in the ‘gaze paper) is not, in contrast to the first type above, closely linked to the meaning of the utterance, but participants can use it inseparably from talk and other modalities for regulating the ongoing activity and for accomplishing a shared understanding of the activity at hand (cf. Heath 1992, 121). One important element of regulating the ongoing activity is the interlocutors’ engagement in producing similar gestural movements, the way in which they bodily “get in sync”, as it were.

Consequently, both of the above types are important for looking at intersubjective stance taking. By using different embodied practices interlocutors can co-participate and tune in to each other’s actions and jointly come to an understanding of the “official focus” of the interactional activity. The examples in the ‘gaze’ paper (IV) supplement previous work on embodiment in interaction by showing how particular uses of gaze in interaction contribute to and facilitate the coherence of the interactional activity and the participants’ stance taking.72

Also Sandlund (2004) notices the paucity of interactional research on embodiment. She refers to research on how embodied actions correlate with particular emotional states, and claims on the basis of this earlier work (2004, 80-81) “that there is some agreement in terms of universal facial cues of happiness, sadness, fear, anxiety, surprise, and disgust/contempt.” However, she also notes that these studies are usually based on laboratory experiments in which certain emotions are induced artificially without a real interactional context. In addition to Sandlund, Biber et al. (1999, 967) note that “emotive and attitudinal stance meanings can be conveyed through a number of non-linguistic means (such as body posture, facial expressions, and gestures).” However, they do not describe in detail how such meanings are conveyed in interaction.

72 As Kendon (1985, 230) claims, this type of analysis is closely connected to Erving Goffman’s work on ‘frame analysis’ and ‘participation framework’, and the finding that participants tend to “co-operate together in order to establish and maintain an official or common focus of attention on the topic at hand.” I want to thank Adam Kendon for pointing this out to me.
As regards gaze in particular, although its role in interaction has been examined to some extent, its relationship with stance-related phenomena has received only little attention. For one, Argyle (1975, 232-237) suggests that “patterns of gaze play an important role in establishing relations between people,” and further that gaze can be perceived to signal liking, approval, dominance, status, friendliness, and threat. However, he does not look at the interactional contexts in which gaze is used. As to more intersubjective approaches to the role of head movement and gaze in stance taking, Streeck and Knapp (1992, 15) note that a head movement from a vertical to a horizontal motion tends to precede signs of disapproval of a co-participant’s action. In addition to this, Goodwin and Goodwin (1987; 1992) note that the use of a withdrawing gaze together with other paralinguistic features in the action of doing an assessment indicates that the speaker is moving toward a topic closure (Goodwin & Goodwin 1987, 38) or withdrawing from the assessment activity (Goodwin & Goodwin 1992, 169).

The ‘gaze’ paper shows that speakers use the concurrent organizations of assessments and particular gaze patterns as resources for stance taking in everyday conversation. It discusses stance taking in relation to three gaze patterns: congruent gaze points, which mark the Stance Object in the environment in an assessment sequence in which interlocutors communicate convergent stances, mutual gaze, which occurs at a particular sequential position when two interlocutors display agreement, and cut-off gaze, which is produced by an interlocutor as a reaction to a problematic stance by a co-participant, and which is followed by a divergent stance. The findings in the ‘gaze’ paper also provide supportive evidence for the following general findings made in the earlier research on embodied practices in interaction:

1) The gaze shifts discussed in the ‘gaze’ paper (IV) precede the action (in the ‘gaze’ paper the assessments) to which they are tied (cf. Streeck, 1992 and Schegloff 1984).
2) The gaze patterns in the ‘gaze’ paper (IV) are used for regulating co-participants’ attention and behavior and they provide information about the level and focus of involvement in the interaction (Kendon 1992).
3) In the ‘gaze’ paper (IV) the interactants share the gaze patterns (i.e. the gaze patterns are intersubjective), which coincides with the finding that people share movement forms, activities and gesture functions (cf. gaze following discussed by Tomasello 1999a).

In sum, it is a well-established claim that human beings do not just use language for interacting with each other in face-to-face situations but use both linguistic and gestural resources for communicating with each other (Streeck & Knapp 1992, 5). Consequently, researchers who are interested in interaction should also pay attention to the different modalities and how these modalities are used individually (what function a gesture serves in interaction, cf. Bavelas 1994; Kendon 2004) or in combination for achieving and furthering particular goals, and producing actions and activities in interaction (see for example Heath 1992, 102). This claim is discussed from different perspectives in the following sections. In the next section I give a brief overview of the previous research on the intersubjective use of gaze in interaction.
6.1 Gaze, facial expressions and intersubjectivity

People continuously use gaze in interaction for looking at each other, other objects or the general background (Argyle 1975: 232). Even though speakers are mostly unaware of how they use gaze, the different acts of looking can carry out very different tasks, ranging from the task of organizing interaction to the expression of individual emotions and attitudes. Thus, the facial area and the use of gaze are central elements of communication between human beings. As Goodwin (1986, 43) claims, the speaker’s face, which is the area that co-participants tend to look at, “might have special interactive significance.” This is already perceivable in the behavior of young children. Schaffer et al. (quoted in Schieffelin 1983) claim that at the age of two, children begin to use gaze in association with their vocal activity and that a child begins to look at the mother usually either during or immediately after her own utterances. These claims are supported by Tomasello (1999a), who claims that at the age of 11-14 months, which is a central or even revolutionary time period for human ontogenesis, children begin to recognize the world around them, to engage with their caregivers in different ways, and to understand that also other human beings (in addition to the child herself) have intentions and goals. In other words, during this period children begin to understand other humans as intentional and thinking subjects, i.e. develop an understanding of intersubjectivity. One indication of this change is the fact that infants begin to engage with their caregivers by imitating and mimicking them and by orienting to the same deictic reference points with them. One form of such engagement is ‘gaze following’ (Tomasello 1999a):

Most prototypically, it is at this age that infants for the first time begin to flexibly and reliably look where adults are looking (gaze following), to engage with them in relatively extended bouts of social interaction mediated by an object (joint engagement), to use adults as social reference points (social referencing), and to act on objects in the way adults are acting on them (imitative learning). In short, it is at this age that infants for the first time begin to “tune in” to the attention and behavior of adults toward outside entities. (Tomasello 1999a: 62)

After this period of human ontogenesis, gaze following remains a frequent, important and meaningful practice in human interaction. Interlocutors follow their co-participants’ gaze direction in order to trace and identify what their co-participants are looking at. Indeed, in everyday interaction one participant frequently follows her co-participant’s gaze in order to seek for and look at a reference (cf. congruent gaze points in original article IV). Gaze following is thus fundamentally an intersubjective action in that it is an expression of one participant tuning in with a co-participant’s action. It also displays an understanding of the co-participant’s gaze shift as an intentional act, which has a particular goal, and it is thus an important element of stance taking. Gaze following often has its own independent organization and is thus not connected to some other simultaneous organization. However, as is shown in the ‘gaze’ paper (original article IV) gaze following in the case of ‘congruent gaze points’ can also be an important element of recognizing and understanding the focus of talk (or Stance Object) in an assessment sequence, in spite of
the fact that it has little impact on the actual production of the assessments. In other words, as the examples in the ‘gaze’ paper indicate, ‘congruent gaze points’ are an important resource for constructing a joint focus and a shared participation framework in a stance-taking activity (cf. Goodwin 1996; Kendon 1990: 235).

### 6.2 Relationship between talk and embodiment

‘Gazing’ and ‘gesticulation’ are here considered to be elements of ‘embodied behavior’, which in turn is defined as the non-verbal use of any body part as a resource for producing actions and activities in talk-in-interaction. They share similar features in how they relate to talk in at least three ways. First, both can together with talk play a central role in the interlocutors’ action production. Second, both can be used for organizing the interlocutors’ turn-taking. Third, both can be intersubjectively oriented to, i.e. people frequently follow each other’s gaze shifts (gaze following) and similarly, people frequently share similar gesture movements. This last point is very important for the question of how interlocutors engage during a stance-taking activity and it is discussed in section 6.4.

However, there is at least one important difference between ‘gazing’ and ‘gesticulation’, which is also a pivotal question in the investigation of the role of gaze in stance taking. Previous research on embodiment in interaction has not described uses of gaze that would be related to the meaning of talk, and by implication to the stances that the interlocutors take, in the same way it has described how for example emblematic and iconic gestures are related to talk. However, the examples of the cut-off gaze in the ‘gaze’ paper seem to be rare examples of a gaze pattern that is “related to the meaning” of concurrent talk. The meaning of cut-off gaze stems from its sequential position in a stance-taking activity. In other words, an interlocutor’s gaze shift away from a co-participant, after she has produced a problematic action or stance, and the fact that the cut-off gaze precedes a divergent stance by the individual who produces the gaze pattern is also indicative of the close connection of the divergent stance and the cut-off gaze. The relationship between talk and ‘gazing’ and ‘gesticulation’ is discussed in greater detail below.

Although it is quite difficult to draw a definite line between what is a ‘gesture’ and what is not (cf. Kendon 1997, 109; Streeck & Hartge 1992), some extrapolations are possible. For one, McNeill (2000, 1-7) uses the term ‘gesture’ for describing a range of embodied behaviors that can be divided into four different groups: ‘gesticulation’, ‘pantomime’, ‘emblems’ (e.g. codified gestures such as the ‘OK’-sign\(^{73}\)), and ‘sign

\(^{73}\)‘Emblems’ are often conventionalized, but their meanings can vary from one culture to another. For example, the circular gesture which is produced with the thumb and the index finger stands for ‘OK’ in the Western world and thus has positive connotations. However, in some other parts of the world, for example South America and Russia, that same gesture is considered an insult. In a similar way, the hand gesture in which the back of the hand faces the recipient and the index finger and the middle finger together point upward in a v-shape while the other fingers are folded (cf. the
language’. These groups together form a continuum (a.k.a. Kendon’s continuum) regarding the ‘varying degree to which gestures can display features commonly regarded as ‘linguistic’ (Kendon 2001, fn. 1). That is, these groups differ from each other with respect to their relationship to speech (whether the gestures are produced co-presently with speech or not), linguistic properties (whether the gestures have linguistic properties or not), conventions (whether the gestures are conventionalized or not), and finally with respect to their semiotic differences between gesticulation and sign (McNeill 2000). The above definition, however, treats ‘gesture’ to some extent independent of other “less obviously bodied” practices, such as language and grammar (cf. Fox 1999). Consequently, it is not particularly useful for understanding the role of gestures in stance taking. Moreover, the above definition of gesture implicitly draws a line between gesture and other embodied behavior, for example the use of gaze.

Another, but not dissimilar definition of ‘gesture’ is provided by Kendon (2000, 49; 2004, 8). He suggests that ‘gesture’ is a body movement, and the actor who produces this body movement has at least some degree of voluntary control over the body part which produces the gesture. Kendon (2004, 13-15) also claims that movements which clearly contain features of manifest deliberate expressiveness tend to be most easily considered ‘gestures’. In other words, a ‘gesture’ should contain a certain degree of “intentionality” (i.e. the intention of having the gesture produced, not a specific “internal” or psychological intention) in order for it to be considered a ‘gesture’. Whether a ‘gesture’ is intentionally expressive depends on the interactional context and can only be judged as one by co-participants (Kendon 2004, 15-16). This definition provides for the possibility to understand a gazing practice as a “gesture” when a co-participant considers it to be relevant for the unfolding interaction.

Another crucial question that has puzzled researchers, and which is also a central question for understanding stance taking, regards the relationship between embodiment and talk. This relationship is complex and it can be very difficult to determine how and to what extent talk and embodiment are related to each other (Strueck & Hartge 1992). Although there is a shared understanding that the investigation of the co-production of speech and gesture is important, no consensus prevails regarding what the relationship actually is (McNeill 2000, 8-9) or whether such a relationship exists at all (cf. Kendon 1994a, 175). However, Kendon (2004, 15-16) claims that the relationship between ‘gesture’ and concurrent talk is an important question and that it is important to remember that gestural communication is not inferior to verbal communication. Gestures are not used only when “words fail” to offer an appropriate way of formulating an utterance verbally. Rather, gestures and talk tend to co-occur and support each other, and “contribute to the total effect of speaking” (Kendon 2001, 193). Therefore, I here assume, similarly to Kendon (1994a, 176), that

‘victory’ gesture) has been used (at least in Finland) in sports contexts to mark “two championships” or “two victories” in a row. However, in many other parts of the world this gesture is considered a rude and abusive gesture.

74 See also McNeill (1992, 75-104) and Kendon (1997, 110).
the gestures produced by speakers in association with their speech are produced as an
integral part of the same plan of action as the spoken utterance; and that they are but
another manifestation of the same underlying process or […] are produced by the
same “computational stage” as speech.\footnote{See also Kendon (1997, 111).}

However, it is also worth bearing in mind that, as Drew (2005, 78) points out, single
gestures or embodied practices are not likely to alter the action produced by talk.
Kendon’s and Drew’s points above together suggest that it is crucial to probe for the
relationship between non-verbal/embodied behavior and talk. Consequently, it is this
relationship that contributes to communication in interaction.

How does the relationship between talk and embodied practices contribute to the
understanding of how embodiment should be investigated in relation to stance taking?
The ‘gaze’ paper (IV) supports the finding and the claim that there is a close relationship
between talk and embodiment. This relationship is manifest in two ways. First,
interlocutors use particular gaze patterns as part of the stance-taking activity in order to
identify the Stance Object (congruent gaze point) and in order to display momentary
agreement (mutual gaze). Second, an interlocutor can produce cut-off gaze after a
problematic stance by her co-participant. In the examples in the ‘gaze’ paper (IV) the cut-
off gaze precedes a divergent stance by the same person who produces the cut-off gaze.
The cut-off gaze is also closely attached and contributes to the meaning of the following
divergent stance. Therefore, it differs from congruent gaze points and mutual gaze, which
are used for organizing the participants’ stance taking and for constructing a joint
participation framework.

These two different types of gaze patterns also reflect previous work on embodiment
in interaction. The first body of work (related to the cut-off gaze) assumes a traditional
pragmatic viewpoint to gestures and examines how a gesture in a particular moment in
interaction provides an additional meaning to the utterance conveyed through speech (cf.
Kendon 2004, 7-16). Such “pragmatic” gestures can be used to clarify and make the
verbal expression more precise or to provide additional substantive content to the verbal
utterance (Kendon 2000, 54). The pragmatic approach to gestures is discussed in section
6.3 below. I suggest, however, that the pragmatic viewpoint has given little attention to
the use of gaze in interaction and to how meanings become understandable in and
through interaction.

The second group of work considers how embodied practices “play a role in the
processes by which participants in interaction establish and maintain their orientations to
one another” (Kendon 2000, 54). This second group indeed looks at how gestures or
other embodied practices (e.g. gaze) are organized relative to the sequential organization
of interaction and the actions that interlocutors produce. It is thus closely related to the
investigation of congruent gaze points and mutual gaze in stance taking. Both the
pragmatic and the sequential viewpoint investigate some aspect of the “meaning” of the
embodied practices, but approach it from different angles.  

These approaches are briefly discussed in the next two sections.

6.3 Gestures as adding meaning to talk

Gestures have been shown to be often closely related to the meaning of concurrent speech. The linguistic forms that are connected to a particular gesture or embodied practice are called ‘lexical affiliates’ (Kendon 1997; Schegloff 1984a). As previous research has shown, the beginnings of gesture movements that are connected with speech tend to precede the words or the unit of talk they are tied to (Streeck & Hartge 1992). Moreover, gestures which have lexical affiliates in talk can be related to or prefigure the meanings, pragmatic functions, actions or topics that are conveyed in subsequent speech (Kendon 1994b, 171-172). Thus, as Kendon claims, gestures can be used for providing additional or supporting information about the content and the meaning of the utterance (e.g. Kendon 1995; Kendon 2001). This is of course very important regarding stance taking, because it opens up the possibility that gestures or other embodied practices are closely related to the interlocutors’ stances.

The gestures that are closely connected to talk and have lexical affiliates can further be divided into two types. The first are gestures that iconically represent the lexical affiliate (i.e. they resemble the thing they are intended to represent). As Streeck and Knapp (1992, 12) suggest, iconic gestures provide spatial representations of shapes, sizes, motions, etc., but these profiles are elaborated and become recognizable ‘as’ representations by virtue of the adjacency of other gestural units and by their coordination with speech.

Iconic gestures can also be used by interlocutors for clarifying a point made in an utterance. Such actions can be done through enactment, by using body parts for modeling shapes, by sketching shapes in the air, or by pointing towards a referent (Kendon 1997, 112).

Second, gestures can provide additional meaning to what is expressed through talk. One gesture of the latter type is described in Kendon (1997, 111). In this example a speaker, who is telling the story of Little Red Riding Hood, produces a swinging gesture in association of the verb “sliced” in the following sentence: “And he took his hatchet and with a mighty sweep sliced the wolf’s stomach open.” Kendon (1997, 111) suggests that “the precise timing of the arm-swing with the pronunciation of “sliced” could not have been achieved unless the speaker had begun to organize her gestural action in

---

76 See also Streeck and Knapp (1992, 5), who make a similar distinction between ‘symbolic communication’ and those processes which organize social co-presence.

77 See also Goodwin (1986, 37) and the ‘gaze’ paper (IV) in which the gaze patterns precede the actions to which they are connected.

78 This gesture is also iconic.
advance.” In addition to this, the swinging gesture is done horizontally, as if producing a longitudinal cut that opens the wolf’s stomach (Kendon 2000, 52). In other words, the gesture which co-occurs with the verb “serves to make the meaning of the verb [“slice”] more precise” (Kendon 1997, 111).

Kendon (1995; 2000) has also identified what he calls illocutionary marker gestures. He has given a profound account of two such gestures, namely the Mano a borsa (‘purse hand’) gesture and the Mani giunte gesture (‘praying hand’). The first one, Mano a borsa, is a widely-known and frequently used gesture in Italy (Kendon 2000, 56), and it tends to be used in connection with the action of asking a question. The Mani giunte gesture, on the other hand, is used for making appeals in different contexts (Kendon 1995, 259). Both of these gestures are used in particular contexts and therefore make the actions with which they are being produced visible and more explicit. These types of gestural forms have pragmatic functions and they

serve as markers of the speaker’s attitude toward what he is saying, or they express his expectations about how what he is saying is to be dealt with by his interlocutor, or they express the nature of the illocutionary intent of the utterance of which it is a part (Kendon 2000, 56)

Kendon (1995, 264) also discusses so called discourse structure markers. Such gestures are for example Finger Bunch, which is used for identifying the discourse topic in contrast to a comment about the topic, and Ring, which marks a discourse unit that is central for the argument being made.

As the above account suggests, Kendon is not a conversation analyst and does not consider the role that gestures have in their interactional and sequential contexts. Rather, his work quite explicitly (cf. the above quotation) views the gestures’ role for furthering the speaker’s subjective stance towards what she is saying. In this sense, even though the above work considers how gestures are used in interaction for furthering particular interactional goals, the view it adopts of gestures and embodiment is similar to the linguistic work on stance described in section 4.3 and the work on linguistic expressions of stance in news interviews described in sections 5 and 5.2.3, which approach their examples essentially as the producers’ subjective expressions.

What is more, the above work has not considered the roles that gaze in particular may have in adding meaning to talk. In fact, gaze and facial expressions have received little attention in this respect. Streeck and Knapp (1992) claim though that facial expressions are used to provide “metacommunicative commentary” about the unfolding talk. For example, the raising of eyebrows can mark an utterance as new information (Streeck & Knapp 1992). Thus facial expressions can be important for understanding and qualifying the meaning of the following utterances, or for commenting on current talk, and thus become “an important aspect in the fine-tuning of speakers’ and listeners’ mutual alignment” (Streeck & Knapp 1992, 17). The examples of the cut-off gaze in the ‘gaze’ paper (IV) provide a good example of such mutual alignment.

In spite of the limited attention that the research outlined in this section gives to the unfolding interaction, this work is useful in that it highlights and foregrounds the relationship between gestures and speech and shows that gestures can form indexical
connections with speech and provide information that adds a meaningful element to the spoken utterance. In the next section I consider the role that embodied practices can have as a resource for organizing interaction.

6.4 The role of embodied practices in turn-taking and action

In spite of the fact that since the 1970s embodied practices and their role in interaction has received more and more academic attention,79 fairly little attention has been given to embodied practices and their relationship to the language, sequential organization, turn-taking, turn design and action production in talk-in-interaction (cf. Fox 1999, 52).80 There are, however, some notable exceptions. First of all, Jürgen Streeck (1993) is one of those established researchers who emphasizes the importance of finding fundamental explanations for the use of different modalities in communication. He calls this approach ‘a naturalistic framework for the study of human communication and interaction’, and focuses on examining the coordination and collaboration of embodied practices with the other modalities (Streeck 1993, 275). Second, Charles Goodwin (e.g. 2000; 2003) has done pioneering work in showing that in addition to talk and embodiment, it is also important to consider other semiotic fields, such as the environment and the materials or artifacts that surround interactants, and look at how these are used as one important resource for producing actions and activities.

The foundational starting point for this type of analysis is of course that in face-to-face situations human beings are present and visible to each other as physical living bodies (Goodwin 1986; Streeck 1993). The work in the ‘naturalistic framework’ shows, first of all, that the different uses of gestures, body movements and gaze are intersubjective and therefore have implications for the interactional organization. Second, since it focuses on embodiment and its relationship with turn-taking and action production, it has been able to provide evidence for the fact that embodied practices actually can be used for managing and organizing the unfolding interaction.

As regards the turn-taking organization, certain patterned uses of gaze are tightly intertwined with it, and thereby have important synchronizing roles in interaction. Speakers can for example use gaze to indicate that the person they are looking at is the addressee of their utterances or the target of their actions (Goodwin 1981; Kendon 1992; Streeck & Knapp 1992, 332) or it can be used for handing over the floor to another speaker (Levinson 1983). Gaze can also play an important role in selecting next speakers. Lerner (2003) builds upon Sacks, Schegloff and Jefferson’s (1974, 717) remark that gaze direction can be used for selecting the next speaker in a first-pair part of a sequence-initiating action. Lerner (2003, 179-182) shows that gaze can indeed be used as an explicit form of addressing. He points out, however, that various simultaneous co-

---

79 A thorough account of the historical development of the study of ‘gesture’ and the relationship between ‘verbal and non-verbal communication’ can be found in Kendon (2004). See also Jones and LeBaron (2002), Kendon (1997; 2001), Streeck and Knapp (1992), and Streeck (1993).
80 The same point is also raised in a special issue of Journal of Communication (2002, 52(3)).
participant actions may impede the success of addressing a recipient, after which additional tasks may have to be accomplished in order to secure the addressing.

Furthermore, Kendon (1967) shows that in interaction interlocutors’ look at each other occasionally and that these occasionally converging gazes follow a cyclic pattern that is closely connected to turn taking. This pattern is asymmetrically distributed depending on whether an interlocutor has currently assumed the interactional role of a speaker or of a recipient. Kendon (1967) claims that a speaker tends to look away from the recipient at turn-beginnings. When the speaker talks the recipient tends to look at the speaker more than the speaker looks at the recipient, and the speaker’s glances at the recipient are shorter than *vice versa*. When the speaker begins to approach the end of her turn, she begins to look at her co-participant. However, at the end of the speaker’s turn, the recipient tends not to look at the speaker. Consequently, at turn transition, the next speaker is not looking at her co-participant and a new “gaze cycle” can begin. As the example of ‘cut-off gaze’ in the ‘gaze’ paper (original article IV) seem to suggest, ‘cut-off gaze’ can be considered a deviant case of this cyclic pattern.

Kendon (1967) has also shown that those moments in which the interactants’ glances meet in interaction tend to be related to the speaker’s aspiration to receive confirmation that the message is being understood. Bavelas et al. (2002, 576-577) have called this brief moment of mutual gaze a *gaze window*. They show that a listener tends to produce a backchannel or a nod when the speaker looks at her, i.e. when the gaze window is open. After the listener has produced the response, the speaker shifts her gaze and the gaze window closes. Thus, the interactants actively use the gaze window to regulate their actions. Finally, as regards hand gestures in particular, Schegloff claims both that a hand gesture by a non-speaker can indicate an intention to take the floor (Schegloff 1984a, 271) and that interrupted speakers can maintain a gesture in order to display that they consider their turn still in progress and intend to continue it after the interruptive stretch of talk has ended (1984a, 271-272).

Particular uses of embodied practices can also be important resources for producing actions and in that way can contribute to the unfolding activity. In relation to gaze, Bolden (2003) shows that it can be used as a resource together with other practices for collaboratively co-constructing single syntactic units of talk. She (2003, 205) claims that in her examples a recipient’s “production of a collaborative completion coincides with the full visual engagement of the two participants and seems to be responsive to the shift in the direction of the speaker’s gaze.” Furthermore, Goodwin (1980; 1981) describes some ways in which speakers use restarts to prompt the gaze of disattending listeners at turn-beginnings. He (1980, 284) further claims that a speaker can pause “to delay further production of his sentence until the gaze of a recipient is secured.” These findings are important because they show that restarts and pauses are not necessarily a problem of encoding, but emergent actions that aim to achieve a preferred state of affairs: that the intended recipient looks at the speaker when she starts a turn. Goodwin (1980, 293) also

---

81 Goodwin’s (1980; 1981) findings about the relationship between restarts and gaze and Kendon’s (1967) finding regarding the “gaze cycle” seem to be at least partly contradictory. Whereas Kendon (1967) claims that generally speakers at turn-beginning look away from the recipient, Goodwin (1980; 1981) argues that a mutual gaze seems to be a preferred state of affairs at turn-beginnings.
argues that within turns, if “the speaker gazes at the recipient before the recipient has begun to look at the speaker, a restart is produced”, whereas when “the speaker does not gaze at a nongazing recipient, the original sentence is continued after the phrasal break.” Goodwin (1986) also claims that gaze can be used as a resource for organizing mutual orientation, either to a referent in the distance or to a gesture made by a co-participant. The use of gaze in such contexts usually helps participants receive more information about the utterances they are related to. Along similar lines, Streeck (1993; 1994) shows that speakers frequently shift their gaze towards a gesture (for example a pointing gesture or an iconic gesture) that they have produced just before a keyword (or a lexical affiliate) in the utterance, and after the keyword has been uttered they shift their gaze back to the recipient. Streeck (1993, 294) suggests that by producing this gaze pattern, which seems to be universal, the speaker shows the recipient that the gesture is relevant in the here and now for understanding the content of the utterance that she is producing. Quite often, as Streeck (1993, 291) further claims, the recipients display their understanding of the gesture by moving their gaze toward the speaker’s gesture and by producing a verbal interpretation of that gesture. Sometimes these ‘interpretations’ occur in the context of word searches (Goodwin & Goodwin 1986). In word searches (Goodwin & Goodwin 1986), the speaker tends to gaze away from the recipient, whereas the recipient, in case she has not been looking at the speaker, tends to shift her gaze toward the speaker when the word search begins. By acting in such a way, the recipients visibly show that they are expecting an outcome of the word search. As Goodwin and Goodwin (1986, 54) suggest, this provides further evidence of the fact that co-participants in interaction are continuously oriented to each other’s actions and co-participating in the talk in progress. Interlocutors can also, as Goodwin (1986) shows, use particular types of gestures (e.g. a handclap), either alone or together with a linguistic unit that is indexically connected to the gesture (e.g. pronoun “this”), to request a gaze from a co-participant who is not looking at the speaker.

There is also some research about facial expressions in interaction and their relation to turn-taking and action production. Streeck and Knapp (1992) claim that facial expressions are used to provide “metacommunicative commentary” about the unfolding talk. For example, the raising of eyebrows can mark an utterance as new information (Streeck & Knapp 1992). Furthermore, Streeck (1992) and Streeck and Hartge (1992) provide evidence for the projecting and contextualizing features of a practice in Ilokanan conversation in which an [a]-face (a facial expression that can be associated with the production of the syllable ‘a’) can be used prior to a production of a news receipt token at turn transition, or as a recognizable turn-entry device. Thus, also facial expressions can be important resources for understanding and qualifying the meaning of the following utterances, or for commenting on current talk, and thus become “an important aspect in the fine-tuning of speakers’ and listeners’ mutual alignment” (Streeck & Knapp 1992, 17).

As regards gestures in particular, Heath (1992, 108-109) shows how a gesture produced by a speaker reattracts the recipient’s attention in case she has not been looking at the current speaker. Gestures can also play an important role in projecting and foreshadowing aspects of possible later productions (Schegloff 1984a, 267), for example for projecting turn completion (Fox 1999, 57).
As was mentioned in section 6.3 since a gesture tends to be produced before the lexical affiliate it has in talk, it can be argued that the lexical affiliate is already “in play,” before it has been uttered. Heath (1992, 107-108) shows that in a medical consultation a series of lateral head movements produced by a doctor simultaneously with a question, asking whether the recipient has any further issues to bring up, affects the meaning of the concurrent utterance, and encourages and prefers an answer from the recipient declining any further questions. Thus, it seems that a particular use of an embodied practice can also impact the understanding of the action that the simultaneous turn-at-talk is producing and therefore affect the preference organization of the activity. In another context, Heath (ibid.) shows that a lateral head nod of the speaker can also elicit a corresponding movement from the recipient before the verbal utterance has been finished. Thus, the corresponding movement displays the recipient’s alignment with and understanding of the speaker’s action and the meaning of an utterance without waiting for its linguistic completion (Heath 1992, 109).

The above research clearly foregrounds and provides empirical evidence of the fact that embodied practices in interaction are not idiosyncratic to one speaker, but rather that they are a central part of the regularized contingencies of interaction. Moreover, the research reported also indicates that embodied practices are used intersubjectively and for engaging with co-participants.

One central question that has preoccupied conversation analytic research on embodiment in particular is as follows: when is it possible to make the claim that a particular embodied practice is occasioned by a prior action, and further occasions and warrants upcoming actions? As was mentioned above, Drew (2005, 78) claims that in important respects nonverbal conduct is subordinate to the verbal conduct with which it is intermeshed; it’s probably true to say that none of the practices, or patterns identified in CA research are shaped or altered in any significant ways by accompanying nonverbal conduct.

Although for example Heath’s (1992) above example of the lateral head movements in medical consultation would seem to suggest the opposite, it is indeed a challenging task to try to find solid evidence that an embodied practice contributes to the unfolding interaction in a way that it is understood to be doing a particular action (cf. Goodwin 1986, 29; Goodwin 2000). There are two reasons for this. First, as Goodwin (2000) claims, there are several semiotic fields at play simultaneously and not all of them are necessarily relevant in the unfolding interaction. Second, interlocutors tend to respond to talk and not gestures (cf. Drew’s 2005 point above). However, the fact that interlocutors tend to respond to talk and not gestures does not mean that analysis should focus on embodied practices alone and disregard their connections and relationship with other modalities, such as talk, and the environment. On the contrary, as Goodwin’s (e.g. 2000) work shows, one should indeed consider how the different modalities and the environment are used together to organize interaction, but at the same time, a gesture’s relevance for interaction has to be shown and grounded in the participants’ actions. This has been done by, for example, Bolden (2003) (see above) and Goodwin (2000), who
convincingly show that an embodied practice produced by one interactant occasions an action by a co-participant and is also understood by her as doing that action.82

As we saw above, Bolden (2003) provides an analysis of how talk and embodied actions are used as a resource for collaboratively co-constructing single syntactic units of talk. She claims that action-sequential or grammatical resources may not always provide sufficient information of what action a speaker is producing, and recipients may therefore need other cues for understanding that action (cf. “cut-off” gaze in the ‘gaze’ paper). Bolden (2003) shows, for example, that both a gaze shift and a pointing gesture together with turn-constructional and sequential resources provide a recipient with information that helps her complete a co-participant’s turn-at-talk. Furthermore, Goodwin (2000) shows how during a game of hopscotch, one of the participants challenges the other for throwing a beanbag in the wrong square and thus accuses her of breaking the rules. The challenger produces the challenge by relying on a simultaneous use of speech and a hand gesture that visually represents information given in talk. Goodwin suggests that by producing such a hand gesture the challenger secures the attention of her recipient and thus ensures that the action does not fail to achieve its goal, which is to hold the recipient accountable for breaking the rules and to stop her from jumping further. However, in the example, the recipient shifts her gaze away from the challenger toward the hopscotch grid, which in turn occasions the challenger’s shift from using the hand gesture to using the grid (i.e. the environment) as a resource for further producing the challenging activity.

In another example, Goodwin (2003) shows how a graphic structure in the soil on which two archaeologists focus their attention plays a central part in the action complex in which they engage. The structure (similarly to the grid in the above example) “provides organization for the precise location, shape and trajectory of the gesture” (Goodwin 2003, 20). Thus, as Goodwin (2000, 1499) claims, “talk and gesture mutually elaborate each other within (1) a larger sequence of action and (2) an embodied participation framework constituted through mutual orientation between speaker and addressee,” even though he also maintains that we should take into account the environment. Consequently, the different modalities (language and embodied practices) and the environment elaborate each other, are mutually interdependent, and are relevant to the organization of an action complex or an activity in interaction.

The body can also be used for displaying orientation or participation in an activity. Schegloff (1998) considers how a particular stance of the body (below neck and above waist), which he calls ‘body torque’, affects the interactants’ behavior and shapes the way they produce talk. ‘Body torque’ is a postural stance in which a person’s head and gaze are oriented to a “temporary” interactional activity, i.e. facing a direction that is distinctly different (approx. 90 degrees or more) from the continuing direction of the body parts below the neck. The ‘body torque’ tends to occur in contexts in which two “competing” activities are at play. By placing oneself in the ‘body torque’ the interactant can manage a “temporary” activity (by shifting head so that it faces the place of that activity), but continue to display orientation to the “official” activity (with the body parts below neck)

82 See also the special issue of Research on Language and Social Interaction 27(3), 1994, in which the articles shed light on the question of whether gestures as part of spoken utterances can affect the recipients’ understandings of an utterance and whether recipients act upon these types of gestures.
which is projected to start or continue shortly, and which will be returned to after the “temporary” activity has come to a conclusion.

In sum, the conversation analytic work, or the ‘naturalistic framework’, provides an important methodological starting point for unraveling the role that embodied practices have in stance taking. This research focuses on those features that are “visibly consequential to the participants at any given time” (Bolden 2003, 196). As was noted above (cf. Drew 2005), this does not mean that an embodied practice has to override the action produced by talk and to be “more” consequential for the unfolding sequential activity than talk. The rationale behind this is that even though gesticulation forms a central part of interaction, gestures and other forms of embodiment themselves rarely receive separate responses. However, embodied practices, of course, still can and do communicate.

Consequently, the focus should be on the use of an embodied practice as one resource in the overall communication, which is accomplished through a combination of both embodied and spoken practices. And it is this combination that can be seen to be consequential for the ongoing sequential activity. This approach assumes that it is the interactants’ conduct as a whole, the combination of talk and gesticulation, which can be used by recipients to understand what the speaker is doing. Therefore, it also provides an apt framework for examining how interlocutors use various combinations of linguistic-embodied practices in their stance taking. This is discussed in greater detail and connected to the ‘gaze’ paper (IV) in the following section.

Finally, it is important to note that the above division of embodiment research into the two groups in sections 6.3 and 6.4 does not mean that these groups are mutually exclusive or concentrate on completely different aspects of embodiment. In fact, they share interests and their work and findings tend to overlap with each other. As Heath (1992, 115-116) quite aptly suggests, a single gesture can simultaneously form several action trajectories, i.e. an iconic representation of something, but at the same time choose a particular recipient as the primary recipient and thereby structure co-participation as the activity emerges. The above research clearly foregrounds and provides empirical evidence of the fact that the use of embodied behavior (including gaze) in interaction is not idiosyncratic to one speaker, but rather that uses of various embodied practices are part of the regularized contingencies of interaction.

6.5 Recurrent linguistic-embodied practices as resources for stance taking

In the above sections I outlined two approaches to embodiment in interaction: one that considers how embodied practices add meaning to talk (a pragmatic approach in section 6.3), and one that considers them to be used for organizing turn-taking, for designing turns, and for producing actions in interaction. The ‘gaze’ paper (original article IV)

83 It is combination that many conversation analysts refer to as “embodiment” and “embodied behavior.”
assumes that embodied practices and speech together form ‘language’ and are thus used for communication between interlocutors. It investigates particular routinized combinations of embodied and linguistic practices during the sequential production of assessments (Goodwin & Goodwin 1987; Goodwin & Goodwin 1992; Pomerantz 1984). It shows how the particular ways of using gaze and gaze direction function as a resource for not only expressing a stance, but also for others to trace and understand another’s stance and act accordingly. In other words, the basic claim of the paper is that different modalities (linguistic and embodied) can contribute to the construction of assessment sequences and thus to the interlocutors’ stance taking. These questions are then examined from two perspectives. First, the ‘gaze’ paper (IV) focuses on

1) how particular recurrent and routinized embodied practices (including gaze) together with other (e.g. linguistic) practices are used in certain sequential positions to produce particular actions and action sequences;

This aim is in fact very important. As many experimental studies have shown (see Kendon 1994a, 177-182) the accuracy of hearing or understanding an utterance is distinctly improved if it is accompanied by a relevant embodied practice (or gesture). Consequently, showing how embodied practices together with talk form routinized behavior in real interactional situations is important for understanding how human beings can use these practices in order to maximize and make sure that their stance (or an action) is going to be understood and that it is going to be understood correctly by the recipient. This does not necessarily mean that an embodied practice overrides other modalities. If we can find patterns in which an embodied practice co-occurs and is routinely used with an action or a linguistic practice, it is likely that such a combination of practices plays a communicative role in interaction.

Second, the ‘gaze’ paper investigates

2) how interlocutors co-ordinate embodied practices in order to co-participate in an interactional activity, and to negotiate and arrive at a joint and shared stance regarding the official focus of that activity.

This type of work is closely related to what Kendon (1990) calls ‘frame-attunement’ (cf. footnote 72) It is also closely connected with the above-mentioned work by Goodwin (2000; 2003) which shows how multiple ‘semiotic fields’ are used simultaneously for building an action within a larger activity. In other words, the participants are actors within a particular participation framework in which particular embodied practices are used to display the interlocutors’ orientation towards each other and, for example, to come to a shared understanding of the topical focus in their talk.

In the examples in the ‘gaze’ paper (IV), the participation framework is partly built through the visible uses of gaze in the activity of stance taking. In other words, the interlocutors use gaze as a resource for tuning in with each other, and the particular uses of gaze help the participants orient to each other and display understandings of each other’s actions. This view is also closely related to Goodwin’s idea of participation framework as a contingent accomplishment that is achieved through the participants’
mutual orientation and participation in the activity at hand (Goodwin 1996; Goodwin 2000).

The ‘gaze’ paper describes three gaze patterns that are coordinated with speech and occur in particular sequential positions in assessment sequences. The different gaze patterns are also closely related to what type of assessment is being produced and to the preference organization of the assessment sequences. These gaze patterns are called ‘congruent gaze point’, ‘mutual gaze’, and ‘cut-off gaze’.

The ‘congruent gaze point’ occurs in such assessment sequences in which the speakers produce agreeing assessments and thus display convergent stances. I show that the shared stance is constructed both with the help of the action constructed through talk (the assessment sequence) and the particular kind of use and reproduction of linguistic elements. In addition to this, the way in which the participants use gaze in the examples provides an important resource in this activity. The following patterns are identified:

1) The fact that the Stance Object is accessible to the interactants enables them to use it as congruent gaze point and, together with the linguistic referent, tie the object of talk to the stance-taking sequence; in other words, the first gaze shift (together with the accompanying linguistic indexical) is also used as a deictic expression; by using gaze in this way, the first assessor presupposes (or recognizes) that her co-participants are positioned so that they can gain visible access to the referent (cf. Goodwin 2000, 1501). In sum, in an assessment sequence the ‘congruent gaze point’ elaborates talk and is used by one interlocutor for showing her co-participants what the focus of the assessment activity is. The congruent gaze point acts as an indexical in the same way as the ‘it’-pronouns in the examples. By using talk and gaze together, a speaker can guide her interlocutors to the appropriate referent in case there is a possibility that it remains ambiguous. However, as is explained in the paper, other contingencies also apply, such as whether the referent is ‘given’ (familiar) or ‘new’ (unfamiliar) as a Stance Object.

2) The interactants’ action of looking at the ‘congruent gaze point’ is a shared gaze pattern. After the referent becomes foregrounded and noticeable as a Stance Object (verbally or non-verbally or both), the participants begin to shift their gaze toward the Stance Object and at least for a short period of time, in overlap with the verbal assessments, simultaneously engage in looking at the Stance Object;

3) Even though the gaze organization is separate from the organization of assessments, it seems that in an assessment sequence in which the participants make agreeing assessments, the assessment (first or second) is made only after the assessor has secured direct visual access to the referent (cf. Goodwin 2000, 1503).

---

84 See Goodwin and Goodwin (1987; 1992) and Pomerantz (1984) and section 4.1 for more information about assessments and assessment sequences in everyday conversation.
In sum, not only is there a connection between the ‘congruent gaze point’ and the object of the stances; the embodied behavior, the actions and the use of language all act as important resources for the interactants to engage in stance taking and co-construct a stance.

The ‘mutual gaze’ is closely related to the ‘congruent gaze point’, because both occur in contexts in which the interactants produce an agreeing stance about a Stance Object in an assessment sequence. However, these two patterns function in slightly different ways. As is claimed in the ‘gaze’ paper, in the ‘congruent gaze point’ the gaze patterns seem to occur before assessments and are maintained for a fairly long time span. However, ‘mutual gaze’ tends to occur during a second assessment or a display of a shared stance in second position. A likely explanation for this is that when the recipient hears a positively loaded first assessment, she, in order to display her agreement with it, initiates a gaze shift toward the interlocutor who uttered the first assessment. Therefore, the participants who engage in producing the assessments are likely to look at each other during the second assessment. And whereas an act of looking at a ‘congruent gaze point’ requires that the referent is present and perceivable for the participants, ‘mutual gaze’ can occur notwithstanding the presence or absence of the referent.

Finally, the ‘cut-off gaze’ is used in contexts in which an interlocutor expresses a divergent stance. The sequential position of the ‘cut-off gaze’ suggests that it is a response to a prior stance. In addition to this, the ‘cut-off gaze’ also precedes a divergent stance by the producer of the ‘cut-off gaze’. Regarding the latter point, therefore, the ‘cut-off gaze’ can also be called a ‘pre-positioned’ embodied practice or an ‘action-projector’ gesture, which occurs at a transition space and provides a preview of and projects upcoming talk that will be available for analysis for the recipient (cf. Schegloff 1984a, 267; Streeck & Hartge 1992, 14). What is important is that the ‘cut-off gaze’ differs from the organization of regular gaze trajectories in turn transition (cf. Kendon 1967), because the time frame within which it is maintained is considerably long, i.e. it spans at least one extra possible completion point. The ‘cut-off gaze’ can occur in at least two sequential positions: after the first part of an assessment sequence, in which the referent is introduced, or after the first assessment. Nevertheless, it always follows a stance that is treated as problematic by the recipient and precedes the verbal utterance which produces a divergent stance. We can also speculate whether the ‘cut-off gaze’ contains iconic features and is thus, in the sequential position it occurs, a contextualization cue that has communicative value and helps participants to share the interactional framework (cf. de Fornel 1992). It is possible that a gaze withdrawal from a co-participant and the fact that it is kept withdrawn for a relatively long time period is indexical of (or at least metaphorically connected to, cf. Streeck & Knapp 1992, 14) an emotional state or a stance of the interlocutor who produces it. In other words, the interlocutor who produces the ‘cut-off gaze’ displays divergence with the immediately previous action or stance. This also gains support from the sequential placement of the ‘cut-off gaze’, that is, before the speech unit it is related to (Streeck & Knapp 1992, 13).

---

85 See also the analysis in the ‘identity’ paper (original article V), in which an interviewee produces a cut-off gaze after an interviewer’s intervention that expresses a stance that proves to be problematic for him. After the cut-off gaze, the interviewee produces a disagreement.
As was mentioned at the beginning of this section, the aim in the ‘gaze’ paper is not to focus on the inherent “meanings” of the gaze patterns, nor is it to examine how an embodied practice occasions an action. Rather, the aim is to show that interlocutors can use different routinized embodied practices together with linguistic practices at particular and recurring sequential positions in order to “behave correctly” in a particular interactional activity and to further a particular communicative goal (in this case, the expression of stance in assessments) (cf. Kendon 1990, 231). Indeed, in the ‘gaze’ paper, the gaze patterns in the assessment sequences are not meaningful per se, nor can we be absolutely sure whether any of them add information to the assessments (although this is quite possible in ‘cut-off gaze’). However, the paper shows that the gaze patterns recur in particular sequential positions in the assessment sequences (relative to the unfolding of the action and particular linguistic practices) and thus contribute to the overall stance taking and the production of the assessment sequences. The ‘gaze’ paper also shows that interlocutors can use embodied practices (in this case gaze) for arriving at attitudinal consensus regarding the official focus of the interactional moment or activity. Indeed, finding such recurrent combinations of embodied, linguistic, and action-sequential resources can show how such consensus is achieved in interaction. However, further careful analysis of these and other embodied practices and the role they play in talk-in-interaction is clearly needed.
7 Categories as resources for intersubjective stance taking in news interviews

The ‘identity’ paper (original paper V) relies on the combination of conversation analytic methods and the theory of stance introduced in the above chapters, in order to discuss the use of category terms in a news interview and how these are used for negotiating and constructing identities for human beings. The two fundamental assumptions in the ‘identity’ paper (V) are that by selecting a categorical term a person can be claimed to have a particular identity (cf. Antaki & Widdicombe 1998b, 3), and that categories and collections of categories are occasioned and negotiated in situated discourse, and that they are recognized and understood in that particular situated moment (cf. Hester & Eglin 1997, 19). The ‘identity’ paper (V) is able to show an example of how an (Arab-American) ethnic identity becomes manifest in the interlocutors’ actions and turn taking and that the participants negotiate the ethnic identity for a third person.

In retrospect, however, there are two major concerns that can be raised about the ‘identity’ paper (V). First, it analyzes only one example, and therefore does not introduce patterned or recurrent behavior in terms of how categories are used as part of stance taking (cf. the outline of the aims of analyzing stance taking in the ‘theory’ paper (I)). There is a reason why the ‘identity’ paper (V) did not discuss patterns of using categories: it concentrates on an ethnic category, and finding patterned behavior even from a large corpus that specifically relates to a particular type of category and proves that that category is a salient element of the interlocutors’ conduct is very difficult. Second, as the name of the original article V (i.e. ‘identity’ paper) suggests, it aims to focus on how

---

86 It is also worth noting that the ‘identity’ paper (V) was written at a very early stage of this work. In fact, it was the first submitted paper of the five original papers, submitted to a conference proceedings (Haddington 2002b) in November 2002. The first submitted manuscript differs a lot from the final version. At the early stages of this work my intention was to study how interlocutors as part of their stance taking in news interviews talk about ethnic minorities and gender, and simultaneously construct and negotiate stereotypical identities for these groups or individuals in them. However, since the data did not yield many examples the focus of the dissertation was changed.
interlocutors construct ethnic ‘identities’ to the participants or other people. This opened up the Pandora’s Box, as I will explain below.

The amount of research dedicated to ‘identity’ is enormous. The construction of identities has been examined and researched for example in critical linguistics, sociolinguistics, anthropology, and cultural studies to such an extent that it is impossible to give a comprehensive and detailed overview of that work here. One important preliminary that the recent research on identities across various disciplines predominantly shares is that identities are not stable or fixed, but rather emergent and dynamic (e.g. Boxer 2002; Hall 1996; McIlvenny 2002; Moerman 1988; Moerman 1993). Consequently, identities have increasingly been studied in terms of how they are actively negotiated, renewed, and contested in situated discourse (e.g. Antaki & Widdicombe 1998b; Fant 2001; Kitzinger 2002; Piirainen-Marsh 2002). However, even this preliminary is understood in very different ways. For example, in some sociolinguistic research the findings and even data collection tend to reflect the analysts’ ideological hypotheses, presuppositions and objectives (see also Schegloff 1992b, 106). Such work often does not attempt to ask ‘what happens’ in interaction, but rather tries to explain ‘why it happens’. As a ‘because’ to the ‘why’, such work often claims that it is static sociological variables such as access to power (e.g. male over female) that contribute to the reported interactional phenomena (e.g. interrupting in conversation). Since the aim of the present work was originally to investigate identity in naturally-occurring talk with the methods of conversation analysis, the above sociolinguistic approach would not do.

Conversation analytic research on identity, on the other hand, tends to concentrate on so-called ‘discourse identities’—e.g. ‘current speaker’, ‘listener’ (Zimmerman 1998)—which are claimed to be the materials out of which social and institutional identities (e.g. ‘call-taker’ and ‘caller’ in emergency calls, or ‘interviewer’ and ‘interviewee’ in news interviews) are built (Antaki & Widdicombe 1998b, 11; Fitzgerald & Housley 2002, 583; Schegloff 1992b, 105-106; Zimmerman 1998). In similar lines, Fitzgerald and Housley (Fitzgerald & Housley 2002, 582) claim that in radio phone-ins the categorical work is closely related to the organization of the unfolding question-answer sequences. For example, the questioner not only orients to the sequence of conversation, but also to the obligation of the category ‘questioner.’

Although it is possible to use conversation analytic methods for examining so-called transportable identities, which relate to visible (physical) or oriented-to (socio-cultural) marks, such as gender and ethnicity (Zimmerman 1998, 90), some conversation analysts claim that if one investigates transportable identities, those methods should be used with caution (see especially Roth 1998, 102-103). One important reason for this is that people can be characterized by several category terms and hold several identities simultaneously. Therefore, it is a challenging task to find interactional evidence of interlocutors displaying an understanding of some aspect of identity, or to find that some identity (e.g. an ethnic identity) affects the unfolding interaction more than any other (e.g. a researcher or a doctor) (Antaki & Widdicombe 1998b, 5; Heritage & Zimmerman 2001; Schegloff 1992b, 106).87 In fact, as Antaki and Widdicombe (1998a, 2) claim, negotiation of

---

87 This was briefly discussed in section 4.1, which discussed conversation analysis. As I said there, conversation analytic work starts from the assumption that interactional practices recur freely
identities is part of the interlocutors’ daily routines. It is not a primary activity that interlocutors engage in. Consequently, it should be borne in mind that in interaction, “negotiating an identity” rarely occasions a next action but rather that a particular action can be used for negotiating an identity. A good example of this is Kitzinger’s (2002) analysis of lesbian and gay coming-out sequences in which the act of coming out is not doing ‘news-announcement’ but is produced as a list item in one TCU in a long turn. In other words, the coming out is produced as a passing reference that the interlocutors are not given a chance to comment on. Kitzinger’s finding provides an interesting pattern, but it should be noted that the same feature of turn design which in Kitzinger’s data is used for “coming out in passing” can in other interactional contexts be used for communicating any other difficult issue (regarding identity, family, etc.). Consequently, the use of a particular action or turn-design feature is not exclusively connected to identity (see also Roth 1998; Sacks 1972a).

Since the original focus of my work was on transportable identities, the use of the functional-interactional approach turned out to be difficult. This is because the functional-interactional approach adapted in the present thesis is essentially meant for analyzing sequentially unfolding activities of stance taking and the various linguistic, interactional and embodied patterns therein. An attempt to further connect and incorporate the idea of identity-construction with this approach would have required a quantum leap. Although the indexical properties of language in the examples of the ‘identity’ paper (V) together with the interactional evidence suggest that some type of (dubious) identity-construction is in play, the divide between stance taking (as it is outlined here) and identity proved to be too broad for finding interactional and linguistic tendencies of ‘identity-construction’.

In view of the above, this chapter aims to supplement the ‘identity’ paper (V) by suggesting that instead of concentrating on one particular type of category (e.g. ethnic) it is more fruitful to adopt the functional-interactional framework outlined in the previous chapters and keep one’s eyes open to all types of categories, and to seek for interactional and linguistic patterns of stance taking in which category terms are introduced, oriented to and talked about (Thornborrow 2001).

In the following sections I will do the following. In section 7.1, I claim that the analysis of how categories are made relevant and understood in news interviews greatly benefits from ‘membership categorization analysis’ (henceforth MCA), which was also used in the ‘identity’ paper (V). MCA was originally developed by Harvey Sacks and it has subsequently been worked on and developed by several scholars within ethnomethodology and conversation analysis; although surprisingly rarely with news interview data.88 I connect this with the work by Roth (1998) who investigates person description as a situated activity in news interviews, and with Fitzgerald and Housley (Fitzgerald & Housley 2002) who investigate the way in which identities are constructed across different contexts notwithstanding the larger macro-social or contextual factors (such as identities, gender and ethnicity).

88 However, see Pirinen-Marsh (2002) for memberships in multicultural broadcast interaction and Leudar and Nekvapil (2000) for category work in television debates. For studies on the use of membership categories in radio, see Housley (2002) and Fitzgerald and Housley (2002).
over the course of interaction in radio phone-in programs. Roth (1998, 81-82) explores “the interactional and institutional organization of how interviewers describe interviewees in the course of interacting with them” and how the descriptions of the interviewees’ personae become relevant and consequential for the participants. A central part of the analysis, as Roth (1998, 84) also argues, is to look at the actions that the interlocutors’ turns implement and how the use of category terms contribute to the production of those actions. However, rather than concentrating on how categories and person-descriptions are “interactionally relevant to the task of soliciting newsworthy information and opinion” (cf. Roth 1998, italics added) I suggest a way to investigate how they are used as resources for and negotiated in the course of the activity of stance taking.

By analyzing two examples I claim that categorical terms can contribute to the ‘positioning’ and ‘alignment’ activities described in the ‘theory’ paper (I) and the ‘positioning’ paper (II). Second, I show that the use of categories and category-bound activities can contribute to the meaning of a question and evoke a preferred stance. Third, I suggest that previous work on MCA and person-description together with the functional-interactional approach is one possible way to probe for the linguistic and interactional patterns by which categories are dialogically negotiated in talk-in-interaction. The examples come from BBC World’s *HardTalk* and BBC 2’s *Newsnight*.

The suggested approach differs from the majority of work in ‘membership categorization analysis’, but is similar to Roth’s (1998) work in that it pays close attention to the sequentially unfolding interaction, or the actions and activities that the interlocutors produce for negotiating categories and category-bound activities. However, the role that grammatical structure has in the negotiation of categories and category-bound activities has not been discussed to date. Finally, I suggest that the members’ assignment of categories is an important and rich resource for stance taking and that the intersubjective negotiation of categories and category-bound features can be a specific type of stance-taking activity. However, the analysis provided below is still only suggestive. More analysis is required of the use of category terms in televised news interviews and in other forms of talk-in-interaction in general. In sum, this chapter illustrates the process of development in the present research project from a macro-social starting point toward a more detailed analysis of the routinized and recurrent practices of stance taking in talk-in-interaction.

### 7.1 Evoking categories and stances for humans

Before it can be claimed that a category is relevant for interlocutors at a particular moment in talk-in-interaction, or that category-bound activities or features are being negotiated and constructed, we have to identify an interational moment where that happens. Such moments can best be identified by searching from the data moments in which interlocutors describe persons or refer to persons, and simultaneously cast a person into a particular category with associated features (Antaki & Widdicombe 1998b, 3). Roth (1998) claims that in news interviews the interviewees’ character can be described in different ways but that eventually these descriptions all relate to a particular category
that has been evoked in the introduction of the program. The use of a particular category to characterize the interviewee in a news interview establishes the interviewee’s status as, for instance, an invited expert and contributes to the newsworthiness of the interviewee (Roth 1998). It also becomes evident in the design of the interviewers’ turns (Roth 1998, 82). This is also evident in the analysis of the example in the ‘identity’ paper (V) in which the interviewee is introduced as the President of the Arab-American Institute at the beginning of the program, and in which “Arab-American identity” runs as a consistent theme throughout the interview. After a relevant category has been identified the analyst must show how the participants and, since we are dealing with news interview data, also the overhearing audience, are able to recognize that that moment is ‘doing categorizing’. Since the objective here is to study categories in view of how they are used in stance taking, the notions ‘category’ and ‘associated features’ are essentially important. This is because it is the relationship between a category and category-related features and activities that tend to be negotiated over what can be seen as a stance-taking activity. This negotiation can be done in two different ways: interlocutors can negotiate whether a particular category term indexes certain category-bound activities (i.e. these activities are the Stance Object) or they can discuss whether particular category-bound activities index a particular category (i.e. the category is the Stance Object).

Harvey Sacks was the first to recognize the significance of categorization within the ethnomethodological inquiry. He developed a framework which has later been called ‘membership categorization analysis’ (MCA) (Sacks 1972b; Sacks 1992a, 40-48, 333–340, 396-403, 590-595). He emphasized the importance of looking at the orderliness of the use of categorizations in interaction. The purpose of MCA is to describe how people use membership categories and membership categorization devices in talk-in-interaction for accomplishing their everyday activities (Hester & Eglin 1997, 3; Mazeland et al. 1995, 277). Furthermore,

MCA directs attention to the locally used, invoked and organized ‘presumed common-sense knowledge of social structures’ which members are oriented to in the conduct of their everyday affairs (Hester & Eglin 1997, 3).

One of Sacks’s first observations was indeed that humans continuously categorize others and describe persons, and they use language to do so (cf. Antaki & Widdicombe

---

89 This is due to an incident that occurred in the United States in December 2001. In this incident an Arab-American secret service agent, who was also President George W. Bush’s personal agent, was removed from an airplane. Since this incident occurred close after the terrorist attacks in September 11, 2001, it caused debate about racial profiling, and racism and prejudice against Arab-American people in the United States.

90 Although Harvey Sacks was also one of the founders of conversation analysis, membership categorization analysis and conversation analysis have different interests and have developed independently of each other (Hester & Eglin 1997, 2). However, here I attempt to combine some aspects of both types of analysis (see also Fitzgerald & Housley 2002; Hester & Eglin 1997, 2).
He noticed that by categorizing people and things into groups interlocutors can dynamically order categories into collections (membership categorization devices, i.e. MCDs), such as [man/woman], [protestant/catholic/jew], or [doctor/patient] (Schegloff 1992b, 107). The above categories can be “commonsensically” heard to go together, whereas other combinations of categories cannot, for example [protestant/patient]. However, Sacks also noted that it is possible to dynamically categorize people into collections that otherwise seem disparate (Hester & Eglin 1997, see also the analysis on the use of the categories "Secret Service agent", "crop duster" and "Arab-American" in the original paper V ). Sacks also noticed that people frequently tie particular category-bound activities or features to certain categories, so that a category may be inferred from knowledge of such an activity or a feature. One of the most important insights of this work was that categorization facilitates interactional situations and provides for sequential coherence (Mazeland et al. 1995, 272). In addition to this, Sacks (1972b, 333) introduced two rules of application that he used to describe categorization in interaction: ‘the economy rule’ states that “a single category from any membership categorization can be referentially adequate.” This means that one categorical term (e.g. the conservative party) from any collection (e.g. political parties in Britain) is sufficient to satisfactorily describe a person. Thus, no extended list of membership categories is needed. The following rule, namely the ‘consistency rule’ then states that

if some population of persons is being categorized, and if a category from some device’s collection has been used to categorize a first member of the population, then that category of other categories of the same collection may be used to categorize further members of the population (Sacks 1972b, 333)

Subsequently, Sacks’s framework has been applied in two alternative ways. The first is what Hester and Eglin (1997, 22) call a ‘decontextualized approach’, which investigates membership categorization devices as ‘natural’ or cognitively stored commonsensical knowledge, which are then drawn upon in everyday talk. The second approach, which Hester and Eglin (1997, 22) call the ‘ethnomethodological approach’, maintains that the use of categories and category devices is not natural or pre-given but always situated, and “their sense is therefore locally and temporally contingent” (Hester & Eglin 1997, 18). In other words, the meanings of membership categories and membership category devices are always negotiated in situ and there are therefore countless ways in which categories can be combined in meaningful category devices with associated features. This becomes evident in the ‘identity’ paper (V) in which the participants understand that combining the categories ‘Secret Service agent’ and ‘crop duster’ in that interactional moment in which

---

91 This does not mean that other collectivities or non-person objects cannot be classified into categories and groups (Hester & Eglin 1997, 3). However, the focus here is on the categorization of human beings.

92 Indeed, it is important to note that categorization is not necessarily dubious, hegemonic or ideological in a negative sense. Rather, categorization is a useful device that facilitates communication.
the connection is made, is intended to refer to ethnic (Arab-American) identity and to display it in a questionable light. The present work thus draws on the second, the ethnomethodological approach.

In the following I suggest that the analysis of categories and category-bound activities/features can enrich the analysis and understanding of stance taking. They can be used as resources for stance taking, for example, for establishing a relevant background or justifying one’s question or comment, or for explaining someone else’s stances. This is also shown in the following example. It comes from BBC World’s HardTalk. Tim Sebastian, who interviews Richard Ottaway, a conservative MP, invokes a category and connects it with category-bound activities that are potentially hostile and damaging for the interviewee. The interview discusses the investigations on the reasons for waging war against Iraq in 2003, and whether these reasons were justified or not. It is noteworthy that the category, ‘conservative party’, and how it relates to the interviewee, have become clear earlier in the interview, but this relationship is still constantly renewed and negotiated in the interview.

(14) BBC World, HardTalk, Aug 4, 2003: As a conservative MP
IR: Tim Sebastian, IE: Richard Ottaway (29 / 1 / 6:00)

1 IR: But I come back to this point,
2 Malcolm Rifkind,
3 raised it,
4 tum- (GLOTTAL) --
5 last ^year,
6 he’s saying that the ‘party sho[uld have 'a=sked,
7 IE: [(Hx)]
8 IR: Con]servative party should have asked,
9 <MRC>more ... probing .. question[s]</MRC>.
10 IE: [Y]es,
11 (H)[2 (H)2]
12 IR: [2 #In #a w2]ay,
13 [3it’s it’s3] % ignored its duty,
14 IE: [3(Hx)3]
15 %The[4= uh4] --
16 IR: [4as an4] op[5position5].
17 IE: [5&5]
18 %The=,
19 %I %mean,
20 %uh,
21 %frankly,
22 I thought Iain Duncan Smith,
23 u=h,
24 did the right thing,
25 to support the government,
26 when it went to war,
27 In fact,

93 This connection is first established approximately 3 minutes into the interview. See example 9 below.
I would not been .. comfortable as [as #] --

[Or just] leapt

Washington’s bandwagon.

The interviewer’s questioning turn in example (14) is interesting because of the following interrelated features. First, the questioning turn (lines 1-16) can be heard to contain a claim which is accusatory. Second, since the questioning turn does not contain a clear question, there is something in the interviewer’s claim that can be heard to require a response, which the interviewee indeed attempts to give. I will now consider these features in more detail.

The interviewee, in spite of the fact that no question has been asked, displays his readiness to take the turn and respond to the interviewee at a TRP in line 10 by producing the particle Yes, which is followed by his long audible inbreath. His hand movement
simultaneously with the inhbreath further supports his readiness to adopt speakership at this particular moment (see Figure 2 above).

In spite of the interviewee’s turn start, the interviewer continues to produce more talk in lines 12-13. At the end of line 13, there is another TRP, at which the interviewee again tries to take the turn but fails, because the interviewer still continues talking in line 15. Then in line 18, the interviewee is able to start his response. What is going on? The interviewer never produces a question that can syntactically be identified as one, nor are his turns prosodically designed so that they would clearly mark his turn as a question.\textsuperscript{94}

As Clayman and Heritage (2002, 132-138) claim, although the majority of the interviewers’ turns are designed as questions, interviewers sometimes play “devil’s advocate” in relation to the interviewee’s position and use assertions rather than questions in order to elicit a response from their guest. In this way, the question/answer organization is momentarily brought to a halt and replaced by a claim/counterclaim organization. It is precisely this that happens in example (14). The interviewee hears the interviewer’s turn (lines 6-9) to have produced a claim that requires a response, and already in line 10, he is ready to commence and respond to the interviewer’s turn. In the following I will suggest that the interviewee’s attempts to take the turn and formulate a counterclaim are occasioned by the interviewer’s use of a particular category and category-bound activities (cf. original article V). In this context, the category and the category-bound activities are used as a resource for constructing the interviewer’s accusation, and thus to intersubjectively set up a difficult position for the interviewee (cf. original article II). The interviewee orients to this position, engages with it linguistically and aligns with the stance encoded in the interviewer’s turn. But how are the interviewee and the overhearing audience able to hear that the categories are being used as a resource for producing an accusation? The MCA framework can help us with answering that question.

In example (14) above, the category and the interviewee’s and other individuals’ membership in that category is rendered noticeable and understandable in the following way. In lines 2-9 the interviewer identifies a third party, Malcolm Rifkind, and voices a stance he’s saying that the ‘party should have asked, that is attributed to that third party. The proper name is a ‘locally initial reference term’ and is used for making the referent as clear as possible for the interviewee (Schegloff 1996c) and for the audience. The use of the proper name and the third-party stance also contribute to the interviewer’s neutralistic stance (see section 5.2). In line 6 the noun the party invokes a category. The fact that it is produced with the definite article and without a clearer description of which party the interviewer is talking about suggests that he assumes the interviewee to be able to recognize the reference of this categorical term. In other words, the interviewer can assume that in this particular interactional moment the NP ‘the party’ comprises a “minimal recognitional” (Sacks & Schegloff 1979) that the interviewee can easily identify.

Although the interviewer does not directly describe Rifkind as a Tory (cf. the economy rule, Sacks 1972b, 333), the interviewer can assume that the interviewee and the audience

\textsuperscript{94} See the discussion of news interview interaction, turn-taking organization and turn-type pre-allocation in section 5.2.1.
recognize that the recognitionals the party | Conservative party and Malcolm Rifkind are co-selected so that Rifkind can be heard to be a Tory. This is possible because the category (the Conservative party) is tied to particular commonsensical category-bound activities (Hester & Eglin 1997; Sacks 1972b), namely that the Conservative party should have asked, <MRC>more ... probing ... questions</MRC> in lines 8-9 and #In #a way, its its % ignored its duty, as an opposition. in lines 12-15. As Sacks (1972b) claimed, an appropriate category for a member can be inferred from the knowledge of particular features (see also Antaki & Widdicombe 1998b, 4). Since Malcolm Rifkind is described to criticize the Conservative party for not fulfilling its duty as an opposition party, and as such criticism would be strange from a member of any other party, we can assume that he is a Tory. In other words, Rifkind’s action makes sense only in relation to the category ‘conservative party’.

As was mentioned in sections 5.2.1 and 5.2.2, news interviews are targeted for the ‘overhearing audience’. Therefore, although the interviewer can assume and does assume that the interviewer recognizes the NP the party, he cannot assume that the audience automatically recognizes what it refers to. The interviewer’s next action in line 8 displays his orientation to this potential problem. He produces a self-repair and provides a clearer recognitional, the categorical term Conservative party, to replace the potential trouble source the party, and recycles the rest of the previous utterance verbatim (see diagraph 8 below).

Diagraph (8)

6 IR: {} the party should have asked,
8 IR: Conservative party should have asked,

In other words, the fact that the interviewer makes a self-repair to produce a clearer category term is indicative of his orientation to the less-than-sufficient information that he provided with the term the party in line 6, and thus displays his orientation to the viewing audience. It is possible that this type of self-repair, which clarifies a potentially ambiguous categorical term in the interviewer’s turn, is a recurrent pattern. Consider the same pattern in example (15) and in diagraph (9) that occurs a little earlier in the same interview:

(15) BBC World, HardTalk, Aug 4, 2003: As a conservative MP
IR: Tim Sebastian, IE: Richard Ottaway (29 / 1 / 3:24)
1 IR: You call it a flaky basis,
2 but it was a flaky basis,
3 → that your party supported,
4 → the Conservative party supported,
5 IE: (H)
6 IR: this flaky case of going to war.
Diagraph (9)

3 IE: that your party supported,
4 the Conservative party supported,

It is also possible that the above self-repairs are deviant cases which support the general phenomenon that the use of an explicit category term is preferred in news interviews, and that this preference displays an orientation to the overhearing audience. Compare with example (16) below, in which a clear description of a membership category (Labour party) is given right away in line 3. 

(16) BBC2, Newsnight, Jan 19, 2004: Member of the Labour party
IR: Jeremy Paxman, IE: Andrew Fisher (36 / 1 / 5:20)

1 IR: <A>Andrew Fisher</A>,
2 (0) "you are a student,
3 you are a "member of the 'Labour Party,
4 Did you --
5 (0)<A><HI>Ne- you</HI> didn't by any chance</A> "canvass
6 ... for the two thousand and one election,
7 did y[ou]?
8 IE: [I] did % --
9 did some leafleting yes.

A more detailed analysis is also required in order to rule out the possibility that the self-repair in examples (14) and (15) is idiosyncratic to this particular interviewer.

Finally, and most importantly, in example (14) the single category terms the party and Conservative party also implicate the interviewee's membership in those categories, and thus together with the category-bound activities, describe him. As such, they also contribute to the fact that the question sets up a challenging position for the interviewee (cf. original article II). If the interviewee denied the accusation (i.e. that the Conservative party should have asked more questions and that it has ignored its duty), he would denounce a commonsensical category-bound activity, i.e. that opposition parties are supposed to challenge the policies of the ruling party. However, if he immediately agreed with the above accusation, he could be perceived to be a “traitor” to his own party. The interviewee of course recognizes this, because he knows his own affiliation. But how is the interviewer’s turn perceivable for the overhearing audience as a question that sets up a position? It is because the audience understands that since the turn is directed to and projects a response from the interviewee, the interviewer considers him as the appropriate person to respond to the criticism and the accusation. Consequently, the audience is also able to recognize 1) that the interviewee is a member of the group that is indexed by the group label, 2) that the reported failure to fulfill the category-bound activities is attached

95 Given the topical agenda, it is not relevant to give a description of what the interviewee studies, e.g. 'you are an economics student’ or ‘you are a linguistics student’. Therefore, no self-repair in this case is required. However, the membership in the Labour party is important for the current interview.
to that group, and 3) that the _accusation_ that is produced with the help of the reported failure, is also targeted at the interviewee. This of course simultaneously displays the interviewee in a rather negative light. Therefore, the _design of the question together with the use of the membership category and the associated activities incorporate a stance that is hostile, and at the same time set up a difficult position for the interviewee_ (cf. original article II).

Categorization can also be used in a less hostile context. In example (16) above the use of category terms is used for introducing a new guest and also for connecting the guest to the topical agenda and making explicit the relevance of the guest for the overhearing audience (cf. Roth 1998). Example (16) comes from an audience participation program in which Prime Minister Tony Blair (Labour party) is invited to the program to explain and defend his government’s plans to introduce a new university tuition fee system. This issue caused a lot of controversy in Britain in 2003 and 2004.

As was mentioned above persons can be described with several categories simultaneously. In example (16) this becomes evident when the interviewee is connected to two categories: a student and a member of the Labour party. These two categories are not likely to be heard commonsensically to go in the same categorization device (see section 7.1). It is also important for understanding the rationale behind the question that the interviewer does not foreground for example the guest’s main subject of study instead of his political affiliation, by introducing him for example as ‘a member of a political party and a linguistics student’. Significantly, the interviewer introduces him _both_ as a student _and_ as a member of the Labour party, and connects these two categories in a minimal but very efficient way in two consecutive intonation units that have a parallel syntactic structure (see diagraph (10) below).

Diagraph (10)

2 IR: ^you are a student,
3 you are a ^member of the 'Labour Party,

The two category terms in the consecutive intonation units and in parallel syntactic positions establish a paradigmatic relationship in which they resonate with each other and invoke through their indexicality an emergent and situated identity for the guest in this particular interactional context (cf. dialogic syntax in section 4.2.2). This situated identity is important because without it, the following question would not be understandable in terms of the stance that it evokes. The question is designed as statement + tag, _you\(<HIP> \textit{didn’t by any chance}\textit{ canvass ... for the two thousand and one election, did you?}_{..}_ It voices a typical category-bound activity of active party members before political elections. By designing the question in this way the interviewer addresses the incompatible relationship between the two categories and the conflicting situation this dual “membership” presents _in this context_. In other words, the party the guest belongs to and which he faithfully helped during the previous election is now making his life difficult. The two categories used in the interviewer’s turn can be seen to and could in fact set up a difficult position for the guest because they raise together with the description of the category-bound activities the issue that the party has “betrayed” one of
its faithful members. However, the guest does not treat the question as setting up a position for him. Rather, as we will see in section 7.2 this question in this particular context in fact helps the interviewee in doing his next action. The following section considers how categories and category-bound activities are negotiated as part of stance taking.

### 7.2 Negotiating categories and category-bound activities as part of stance taking

The analysis in section 7.1 discusses one aspect of categories: the way in which they are invoked, how they can be understood both by the recipient and the overhearing audience as categories within particular membership category devices. In example (14) they are doing ‘positioning’. The analysis of that example also shows that the interviewee understands the interviewer’s turn to be doing precisely those things. In example (16), on the other hand, the categories are used to introduce the interviewer and to establish a background for the subsequent question. I now turn to a more detailed discussion of how the categories and category-bound activities are oriented to and negotiated over the question/answer sequence and how the stance-taking framework figures in this equation. The type of analysis undertaken in the ‘identity’ paper (V) and further discussed below is not dissimilar to Sacks’s work on MCA or to subsequent work on the occasionedness and situatedness of categorization in talk-in-interaction. However, I provide a closer view to the role of language, especially recycling of words and syntactic structures, and the ways in which the interviewee engages with the interviewer’s question and with the activity or action in it.

Categories and category-bound activities can be seen to comprise an important part of stance taking in two ways. First, the interlocutors can negotiate an appropriate category to describe the participants (Subject1 and Subject2) or the Stance Object (see example (2) in the ‘identity’ paper (V)). Second, the interlocutors can negotiate whether the activities that an individual (Subject1, Subject2, or a third party) has engaged in are appropriate for the category she represents. In this case, the category-bound activities are the Stance Object. The latter in fact happens in examples (14) and (16). As we saw in example (14), the interviewee’s attempt to respond occurs at the first TRP after the interviewer has invoked the category term Conservative party and the connected category-bound activities should have asked, <MRC> more ... probing .. question[s]</MRC>, (in lines 8-9). In other words, the interviewee’s immediate attempt to take the next turn after the interviewer has uttered the category-bound activities shows that he understands and orients to the interviewer’s TCU as doing an accusation, targeted toward him as a member of the category, which requires a response. Nevertheless, the interviewee is not able to take the floor, because the interviewer produces two more TCUs in lines 12-13.

96 In the ‘positioning’ paper (original article II, 29) I claim that “it is [...] important to notice that positioning becomes clearly visible only if the interviewee treats the questions as doing positioning.”
and 15, in which he reports another category-bound activity that the conservative party has failed to accomplish, it’s it’s % ignored its duty, [as an] op[2position2]. and which therefore also contributes to the positioning activity. According to Berg (2003a, 172), in news interviews the preferred next action after an interviewer accusation is a denial or a rebuttal.97 As I show in the ‘denial’ paper (III), denials in news interviews deny a presupposition or an assumption in the question or the adequacy of the question as a whole. The interviewee’s response in example (14) is doing none of these. The reason for this seems to be the following. First, the interviewee cannot deny a commonsensical category-bound activity, i.e. what the opposition parties generally do in politics. This may suggest that although it is possible to deny a presumption, a presupposition or even an accusation, it is difficult to deny any of these if they are related to a category-bound activity that evokes “true” common-sensical knowledge. Second, denying the relevance of the question would be highly inappropriate in this context. Rather, the interviewee carefully aligns with98 the position that has been set up in the interviewer’s question.

The alignment is perceivable in the way in which the interviewee invokes a new member of the category, engages linguistically with the interviewer’s turn, and designs his turn. First, the interviewee aligns with that part of the position that is related to the commonsensical category-bound activity. He takes a stance which is constructed with the help of the membership categorization device invoked in the interviewer’s turn, and invokes another member of the category Conservative party, namely the (now former) leader of the party, Iain Duncan Smith. Smith can be heard to be a member of that category precisely because of the activity the interviewee attributes to him in lines 24-25; namely supporting the government (i.e. not being in the government). The interviewee designs his stance as a personalized stance (note the stance markers, %I %mean, in line 19 and I thought in line 22)99 and thereby does not concede that the accusation of whether the party has ignored its duty as an opposition is right.

Second, the importance of the category and the category-bound activities is further supported by the affective stance the interviewee takes in line 28 and repeats in lines 33-36: or as a Conversative MP=, % I would not (GLOTTAL) be comfortable, seeing British troops going to war, with us opposing it.. Rather than agreeing or disagreeing with the

97 Berg thus seems to refer to preference organization in terms of how it relates to the action the questioning turn does and what the preferred way to answer it is. In other words, in order to save one’s face it is preferred to deny an accusation than to admit it. Preference can also refer to question design. For example, certain kinds of yes/no questions can be seen to prefer particular types of answers (Raymond 2000). Finally, the format of an answer is also closely related to preference. That is, a quick answer is preferred format and a slow and hedged answer is a dispreferred one.

98 As was mentioned in section 4.2.1, ‘alignment’ refers to the range of possible types of convergent and divergent positions that interlocutors take in relation to each other. As was also mentioned, alignment can be managed linguistically and through the design of a turn.

99 Cf. Kärkkäinen (2003, 171) who claims that the stance marker ‘I think’ is routinely used for bringing in the speaker’s personalized perspective that provides a new and differing perspective to that which has been said in the previous turn. Here the stance markers are used after an accusation and frame the interviewee’s response that indeed provides his personal perspective to the issue.
accusation, this TCU in lines 34-36 incorporates a personalized stance that reports a reason why the Conservative party did not fulfill its category-bound activities. The relevance of the category and the category-term is established through the engagement between the interviewer’s and the interviewee’s turns (see diagraph (11) below), by which the interviewee displays his membership (line 33) in the membership categorization device in line 8).

Diagraph (11)

8 IR: Conservative party should have asked, 
33 IE: or as a Conservative MP,

Third, in lines 37-39 the interviewee states (H) ... Of course we have questions about it, Of course we have our doubts, and uncertainties, and thereby makes a partial concession to the accusation that relates to the category-bound activity of ‘asking questions’, but only after he has voiced his support to the party leader and his personal stance that reports a reason for supporting the government. It is noteworthy that whereas the first part of the interviewee’s turn is presented as a personalized stance, the second part acknowledges the relevance of the category-bound activities of the party by using the first person plural pronouns (the accusative ‘us’ and the nominative ‘we’) in lines 36-38, which index the Conservative party and communicate membership in it. In sum, rather than simply denying or agreeing that the party has failed to meet with its duty as an opposition party, the interviewee constructs a stance in which he supports the party leader and his actions but still concedes that the party has questions about the war in Iraq. In other words, the Stance Object, i.e. the category-bound activities, and how they relate to the question at hand are negotiated and understood locally in a short segment of time within a question/answer adjacency pair.

Further support for the significance of categories and category-bound activities in stance taking can be seen in the question/answer adjacency pair and the fact that in the answer the categorical terms are recycled and used as a resource from the previous turn. This is in line with the finding made by Jayyusi (1984 as quoted in Eglin and Hester 1992, 263). As Eglin and Hester (1992, 262) say, adjacency pairs are “a sequential device by which the parties may secure joint orientation to, and mutual understanding of, the categorizations they are making.” As is further shown in the ‘identity’ paper (V), a dialogic and paradigmatic approach to grammatical structure provided by ‘dialogic syntax’ (see section 4.2.2) yields interesting results in terms of how two apparently disparate category terms (‘a secret service agent’ and ‘crop duster’, see example 16 and diagraph 12 below) are paralleled with the help of similar syntactic structures, and used to invoke through their indexicality a new culturally marked membership category (‘Arab-Americans’) with connected features of untrustworthiness. Moreover, in the ‘identity’ paper (V), the recycled category term is syntactically positioned in the object position (i.e. it is a predicate nominal in a one-place intransitive clause and therefore the
semantic center of the clause, cf. Payne 1997, 111), which coincides with the fact that the
category term is the Stance Object (see example (17) and diagraph (12) below).¹⁰⁰

(17)  CNN, Crossfire, Dec 27, 2001: Racial profiling
IR: Tucker Carlson, IE: James Zogby (003 / 4 / 1:21)

91  →  IE:  (H) (0.8) But a ^Secret 'Service 'agent,
92     [A ^guy who] --
93   IR:  [Wait,  
94     'wait a] --
95 IE:  [2who has2] 'passed --
96 IR:  [2'<A>No no</A>2].
97   .. No,  
98   (0) <A>you're ^missing the 'distinction. 
99   This is interes% -- </A>
100 →  He ^said he was a 'Secret Service agent. 
101 →  "Mohamed 'Atta said he was a ^crop duster.

Diagraph (12)

91. IE: But a ^Secret 'Service 'agent 
100 IR: he said he was a 'Secret Service agent  
101 IR: "Mohamed Atta said he was a ^crop duster.

Note also that the action—‘claiming that someone claims to be someone’¹⁰¹—that the
interviewer attaches to the two different referents (he, the Arab-American Secret Service
agent and Mohamed Atta, who was one of the terrorists who attacked the World Trade
Center on September 11, 2001) in example (17) above is an action that is particularly
relevant for attaching a category to an individual. In addition to this, the copula ‘be’ in the
complement clause is likely to recur in this context, since it characterizes the subject NP
he in lines 100 and 101 (cf. diagraph 10).

In contrast to example (17) and diagraph (12), the recycled category term in example
(14) above is in the subject position, which coincides with the fact that the Stance Object,

¹⁰⁰ Note that a similar activity of recycling a category term can be seen in diagraph (10) above.
However, in that example the recycling is done within a turn, whereas in example (16) and diagraph
(12) the recycling occurs between two interlocutors. This reflects the different actions that these
two similar types of recycling are doing. In diagraph (10) the recycling is used by the interviewer to
introduce the guest and to construct a particular situated identity for the guest which is relevant for
the subsequent question. In diagraph (12) the recycling is used for negotiating a particular identity
for the individual being talked about. Note that both of these examples are syntactically similar in
that they contain the ‘be’ verb in an intransitive clause in which the category term is a predicate
nominal. In view of the above, the clause structure provides an interesting working environment for
investigating how apparently different categories are connected to each other in order to invoke
new meanings and identities. More work is of course required.

¹⁰¹ The ‘claiming that someone claims to be someone’ is done with a ‘X said’ + ‘intransitive clause’
structure.
which is under negotiation, consists of the category-bound activities (see diagram (13) below).

**Diagram (13)**

6 IR: {} the party should have asked,
8 IR: Conservative party should have asked,
33 IE: or as a Conservative MP,
34 IE: I would not be comfortable,

As we saw above, the recycling of the category term in line 33 and the co-indexical first person pronoun in line 34 are sufficient for establishing a connection between the interviewer’s and the interviewee’s turns, and for showing that the category term is indeed a central part of how the interviewee’s stance is constructed. The activities that are connected to that category are in the interviewer’s turn, should have asked, *more ... probing .. questions*<MRC>, in lines 8-9 and it’s it’s % ignored its duty, in line 13. In the interviewer’s turn, the transitive verbs ‘ask’ and ‘ignore’ are lexical verbs which together with arguments in the object, namely ‘questions’ and ‘duty’ describe an activity. In other words, the combination of the meanings of the predicate and the argument (i.e. predication) creates a meaning that describes a category. These semantic connections are central elements of the description of the category-bound activity that is attached to the category (cf. the ‘be’ verb in diagram (12) in which such a semantic connection does not occur). It is possible that for negotiating category-bound activities, the use of full lexical verbs is preferred.

Similar phenomena regarding the relationship between categories, category-bound activities and stance taking can also be seen in example (16) and diagram (14) below.

**Diagram (14)**

5-6 IR: {} you didn’t by any chance ^canvass | for the two thousand and
5-6 IR: one election,
7 [you] did?
8 IE: I did --
9 did some leafleting yes.

As was mentioned above, example (16) and diagram (14) come from an audience participation program. Thornborrow (2001) has studied lay participants’ conduct in radio phone-in programs and talk shows. She has noted that after the interviewers’ selection and introduction of the next speaker the lay speakers make salient some aspect of their own personal status and identity, before going on to state their opinion, ask their question or say whatever it is they have to say as a contribution to the talk.

They lay participants do this, Thornborrow (2001) further argues, in order to “warrant the relevance of their public participation at that moment.” They lay participants’ identity is connect to what they are going to say and thus also constructed *in situ*. Example (16) is, however, an interesting deviant case, because the interviewer already constructs an
identity for the next speaker, which, as we will see, is slightly contested by the guest. Also Thornborrow (2001) discusses these “host-attributed identities” briefly and claims that when interviewers construct or attribute an identity to the guest, "the binary division between participants in terms of expert and lay begins to collapse from an interactional point of view, and a more discursively differentiated speaker status comes into play.” This is exactly what happens in example (16), when the interviewer invokes the student’s membership in the Conservative party and in thus brings a new perspective to the student’s question.

Conversation analytic work has considered the above type of questions *you didn’t by any chance ‘canvass ... for the 2001 election, did you?’* in terms of how they are designed to invite or prefer particular answers (see section 5.2.2). As we can see in line 5-7 the interviewer’s question is a statement + tag question. Clayman and Heritage (2002, 210) claim that such questions strongly invite a confirmation of the statement, i.e. ‘I did canvass for the 2001 election’. The question also contains a negative polarity item (*any in by any chance*) (see section 5.2.2) which generally, if incorporated in a question, embodies preference for a “no” answer (2002, 211). However, as we can see in lines 8-9, the interviewee neither confirms the statement nor gives a “no” answer. In addition to this, he does not design his answer as a dispreferred answer (with hesitation markers, delays, cut-offs, etc.), which suggests that he does not treat the question as preferring a “no” answer. One reason for this could be that the “double negative” in line 5 (*didn’t* and *by any chance*) turns the preference organization upside-down so that the question in fact prefers a “yes” answer. In order to prove this a sufficient collection of such occurrences of the “double negative” would be required.

If the “double negative” indeed invites a “yes” answer, it means that the interviewer has designed the whole questioning turn very carefully for the recipient’s benefit by taking into account who he is and why he is in the program (cf. forward-type intersubjectivity). So rather than setting up a position to the interviewee, the design of the question helps the interviewee to amplify his critical stance toward the government’s actions.102 In other words, criticism coming from a student in this context is self-evident, 102 Consider the guest’s critical question to Tony Blair in the following turn.

(18) BBC2, Newsnight, Jan 19, 2004: Member of the Labour party
IR: Jeremy Paxman, IE: Andrew Fisher, IE 2: Tony Blair (36 / 1 / 5:26)

8 IE: [I] did % --
9 did some leafleting yes.
10 ..
11 [GLOTTAL]
12 IR: [Yeah.]
13 IE: Uhmm,
14 <A>I mean the my 'question to Mr 'Blair would 'be=</A>,
15 .. you 'said,
16 .. <A>two and a half years ago</A>,
17 (0) that's when the .. 'manifesto was 'published,
18 (H) that you were 'against top-up fees,
19 and you had 'legislated to 'prevent them.
20 You are 'now ta=liking about people,
21 who 'agree with that 'policy 'now,
but by introducing the guest also as a member of the Labour party, and possibly an active member, evokes a stance that the party has “betrayed” the guest. And as we can see, the guest answers “yes” and thus confirms the fact that he is member of the party and has in fact helped it during the previous election. However, it is also important to note that even though he confirms the statement, he does not fully agree with it. This is closely related to the negotiation of category-bound activities. In other words, in this case the negotiation of category-bound activities (canvassing for the election) is particularly interesting because, as we saw above, the interviewee answers “yes,” but at the same time designs his turn so that the stance in it carefully aligns with the question and displays a slightly divergent stance from it. As we can see in diagram (14) above, the verb leafleting in line 9 is parallel with the verb canvass in line 5. The category-bound activity ‘canvassing for election’ is perhaps appropriate if a speaker wants to communicate a stance that she is very active in her party. However, the activity of ‘leafletng’, which is modified by the pronoun ‘some’, suggests a less active role in the party. In other words, although the interviewee answers “yes”, he simultaneously communicates a stance differential: he is not as active in the party as the verb canvass suggests. This stance differential is managed and becomes evident in the linguistic resonance that rises from the engagement between the verbs ‘canvass’ and ‘leaflet’. It seems therefore that the interactional moment in an audience participation program in which an interviewer selects the next speaker by introducing her and attributing an identity for her is a potential place for identity negotiation.

In sum, the categories and the category-bound activities play a central role in how the stances are carefully oriented to and negotiated in news interviews. They are used for such obvious tasks as introducing guests and giving relevant background information for a subsequent question. However, categories and category-bound activities are also used for setting up positions for interviewees and they can also be negotiated as part of the interlocutors’ stance taking.
7.3 Summary

The ‘identity’ paper (V) and the above examples (14) and (16) show that categories do matter. It is quite probable that they are constantly used for accomplishing actions and activities, and that together with category-bound activities and features they are negotiated as part of the interlocutors’ stance taking. In this chapter I have analyzed two examples and showed that membership categories and category-bound activities are used as a resource for stance taking in news interviews. I suggested that in order to examine how they are used one should first identify an interactional moment during which an interlocutor refers to a person or a group by using for example a lexical categorical term that does referencing to that person or group. A central part of this analytic stage is to see how the interlocutors are able to recognize and understand a particular (socially and culturally loaded) categorical term to be doing such referencing, especially if it is not explicit. I also showed how the actions (an accusation and a question) and the activities (positioning and alignment) that the interlocutors produce are closely tied to the category and category-bound activities. I claimed that these actions and activities are rendered noticeable and understandable precisely because of their close connection to the categories and the category-bound activities (see examples (14) and (16)). In this way, the use of the category terms also displays an orientation to the overhearing audience. Finally, on the basis of both the ‘identity’ paper (V) and the analysis of the above examples, I have shown how interlocutors can either negotiate an appropriate category to particular category-bound activities, as happens in the ‘identity’ paper (V) in which case the category is the Stance Object, or the interlocutors can negotiate what activities should be considered appropriate for a particular category, in which case the category-bound activities are the Stance Object.
8 Conclusions

This work contributes to stance-taking research by introducing a functional-interactional vantage point which combines the benefits of conversation analysis, functional linguistics (the ‘theory of stance’) and interactional linguistics. These approaches are important because they share the methodology of careful and close analysis of naturally-occurring talk and the fundamental idea of interaction as intersubjective. Nevertheless, they differ in terms of their analytic foci and therefore none alone can provide the tools for analyzing stance taking as it has been approached here. In this conclusion I briefly discuss these three approaches and their role in the present project.

First, conversation analysis focuses on the question of how different practices contribute to the social order of interaction. It looks at various interactional organizations and how the interlocutors’ behavior is intersubjectively organized into sequentially connected practices, actions and activities. Conversation analysis provides for the present work the methodology, together with a large amount of relevant findings, for analyzing how the various practices, actions and activities that are used as resources for stance taking are organized sequentially. This is important because stances are taken, understood and negotiated, and come to have a meaning in specific interactional contexts, via the actions and activities interlocutors produce. This is something to which functional linguistics and some interactional linguistic work gives only scant attention.

Second, functional linguistics focuses on the use of grammar in discourse and provides the important understanding that linguistic elements have functional purposes in interaction. Functional linguistics often examines grammar from a cross-linguistic and typological vantage point (cf. Preferred Argument Structure). Although it does not aim to consider the social organization of interaction, the functional linguistic ‘theory of stance’ acknowledges the relevance this has on stance taking (Du Bois 2003d). In addition to this, the ‘theory of stance’ also casts the internal and subjective view of stance held in traditional linguistic research in a new light by claiming that stance is not an individual experience or an inner act, but indeed a shared intersubjective activity accomplished publicly in interaction. An important element of intersubjective stance taking is dialogic syntax, which describes the ways in which interlocutors recycle their co-participants language in order to engage and align with stances made by them. It also provides a
theoretical framework and a concrete method of representation for the analysis of intersubjective stance taking.

Indeed, what I hope to have shown in the present work is that in order to investigate stance taking in talk-in-interaction, a mere analysis of individual lexical and semantic stance expressions studied even in their contexts of use is inadequate. In other words, this work does not try to explain how interlocutors choose a stance from a reservoir of predetermined internal, subjective or psychological “attitudes” and “beliefs.” Rather, it has looked at how stances are taken and negotiated, and importantly they come to have a meaning in specific interactional contexts through this intersubjective negotiation. Most importantly for this work, the ‘theory of stance’ and dialogic syntax also provide the specific understandings of the notions ‘alignment’ and ‘engagement’. These terms are used for operationalizing the notion of intersubjectivity. ‘Alignment’ does not refer to agreeing and disagreeing per se, but rather to the different types of convergent and divergent alignment between two interlocutors’ stances. ‘Alignment’ is managed through syntactic, lexical and semantic ‘engagement’ in which linguistic forms, meanings and functions engage and resonate with each other. The term ‘engagement’ is used in dialogic syntax to refer to the ways in which various linguistic elements are recycled and reused in interaction. Indeed, almost all the phenomena described in this work include examples of engagement which contribute to the ways in which interlocutors align with each others’ stances (the ‘positioning’ and ‘alignment’ activities (II), the production of assessments (IV), the \texttt{NEG} + \texttt{POS} pattern that is used for producing two action combinations (III), and the negotiation of categories and category-bound activities (V and section 7.2 above)).

Third, interactional linguistics can be seen as a combination of the conversation analytic and the functional linguistic work. Although it is naturally open to the idea of intersubjectivity of interaction, the work in it has with some exceptions (cf. research on collaborative completions and turn increments) concentrated primarily on single linguistic elements (e.g. words and grammatical structures) within individual TCUs, and described them relative to the sequential position and the action in which they occur, or how their use differs from the ways in which they are traditionally believed to be used in discourse.

Since the focus in conversation analysis is on social order as it is sequentially managed in and through talk-in-interaction, and in functional and interactional linguistics on morphosyntactic aspects of language, none of them alone can provide the tools for investigating stance taking, which is essentially an intersubjective, sequential and situated activity and a sum of various interactional, linguistic and embodied practices. Consequently, the payoff of combining these approaches is that the analyst can investigate, on the one hand, the role of turn design, actions, and sequential organization, and, on the other, the actual use of grammar in interaction and provide evidence of how they together can contribute to and function in the interlocutors’ intersubjective and dialogical stance taking. As I further suggest in the ‘theory’ paper (I), one should pay attention to cultural resources (e.g. membership categories) that are used for constructing a stance if it can be shown that such a resource is relevant for the interlocutors in the interactional moment that is being analyzed. Considering cultural resources has proved important in investigating the use of categories and category-bound activities and in
understanding the ways in which interviewers’ questions set up positions for the interviewees in news interviews.

As I have noted frequently in this study, when one looks at actual data it is easy to see that “stance is everywhere” and “everything is stance.” In order to get some rigor to the examination of stance taking, I have suggested that the best and most fruitful way to approach it is to collect examples of routinized and recurrent practices of stance taking. In the present work good candidates of such recurrent practices and practice combinations are the following:

1) The production of denials and claims for insufficient knowledge, and the linguistic pattern used to implement them (III).
2) The use of some gaze patterns in assessment sequences (IV).
3) The use of category terms and associated category-bound activities in questions and answers in news interviews (V and chapter 7 above).

In addition to this I have shown that it is important to consider how stances are constructed on a turn level. These are the following:

4) The ‘positioning’ activity, in which the interviewer’s turn as a whole builds up a preferred stance or sets up a difficult position for the interviewee (I and II).
5) The ‘alignment’ activity, i.e. the way in which the interviewee aligns with a difficult question (I and II).

Consequently, even though stance and stance taking can be seen to be complex issues, this work has been able to provide a small contribution to ‘stance’ research by collecting and analyzing some routinely used linguistic, interactional and embodied practices of stance taking.

This work has contributed to stance-taking research by studying news interview data (apart from the ‘gaze’ paper IV). Although stance taking has earlier been studied to some extent in everyday talk, it has not been studied in news interviews. This part of the present work also supplements the conversation analytic work on news interviews, which has considered such stance-related phenomena as interviewer neutralism, hostility and adversarialness, and interviewee evasion and equivocation. However, it has given fairly little attention to how these phenomena impact a co-participant’s behavior. Thus, the aim here has been to give more emphasis to the intersubjective relationship between the interviewer’s questioning turn and the interviewee’s answer. The implications of investigating stance taking in news interview data are significant. The stances that especially political figures take in television will accrue over time and thus impact their publicly displayed persona and contribute to the social construction of their personhood (Du Bois p.c.). The ‘positioning’ and ‘alignment’ activities studied in the ‘positioning’ paper (II) are an important element of this social construction. As I argued in the ‘positioning’ paper (II), these activities indicate that stance taking in news interviews is not as emergent as it is in everyday talk. It seems that prior beliefs, contextual, political and cultural values play an important role in news interviews, although they are of course not “fixed,” but emergent, and created and negotiated in interaction. The ‘positioning’
and ‘alignment’ activities are used for constructing a particular type of persona particularly for the interviewees, but also for the interviewers (consider the interviewing styles of such famous interviewers as for example Sam Donaldson, Larry King, Jeremy Paxman, and Peter Jennings). Stance taking in news interviews is about “who you are,” “what you say” and “how you say it.” Moreover, by investigating recurrent practices of stance taking we can get a better understanding of the rhetoric that interviewers and interviewees use and which they continuously rely on and are without a doubt very good at.

The ‘gaze’ paper (IV) stands out among the original papers, because it discusses embodiment and embodied practices in everyday talk. There is already a fair amount of conversation analytic research on embodied practices. However, there is not much systematic analysis of routinized combinations of embodied and other practices, nor how these practices are used as resources in stance taking. Consequently, in spite of the apparently different status of the ‘gaze’ paper in this work, it addresses these important questions. The relationship between stance taking and embodiment requires more research in the future.

In sum, stance taking is an essential element of interaction. Interlocutors orient to each others’ stances and design their stances on the basis of those in previous talk. They also use a wide variety of different practices and combinations of practices to express their stances and to intersubjectively engage and align with each others’ stances.
Bibliography

Institutional interaction: Conversation analytic studies. Helsinki, Suomalaisen Kirjallisuuden Seura, 121-137.


President Candidates by the News Program Zeit im Bild. Weiner Linguistische Gazette 38-39: 63-76.


Raymond G. (2000) The Structure of Responding: Type-conforming and Nonconforming Responses to Yes/No Type Interrogatives.


Appendix 1: Transcription conventions


**SPEECH**

**UNITS**

Intonation unit
Truncated intonation unit
Truncated word

**TRANSITIONAL CONTINUITY**

Final
Continuing
Appeal (seeking a validating response from listener)
Combinations of appeal

**SPEAKERS**

Speech overlap
(numbers inside brackets index overlaps)
Name/identity/address is pseudo
Name/identity/address is real

**ACCENT AND LENGTHENING**

Primary accent (prominent pitch movement carrying intonational meaning)
Secondary accent
Unaccented
Lengthening
High booster

PAUSE

Long pause (0.7 seconds or longer) ...(N)
Medium pause (0.3 – 0.6 s) ...
Short (brief break in speech rhythm)(0.2 or less) ..
Latching (0)

VOCAL NOISES

Alveolar click (TSK)
Glottal stop (GLOTTAL)
Exhalation (Hx)
Inhalation (H)
Laughter (one pulse) @
Laughter during speech (1-5 words) @ (e.g. @two @words)
Laughter during speech (+6 words) @ (e.g. <@>six words </@>)
Others (SNIFF), (DRINK), etc.

QUALITY

Special voice quality <VOX>two words</VOX>

Loudness
Forte: loud <F> </F>
Piano: soft <P> </P>

Pitch
Higher pitch level <HI> </HI>
Lowered pitch level <LO> </LO>
Parenthetical prosody <PAR> </PAR>

Tempo and rhythm
Allegro: rapid speech <A> </A>
Lento: slow speech <L> </L>
Marcato: each word distinct and emphasized <MRC> </MRC>
Rhythmic: stresses in a beatable rhythm <RH> </RH>
Arrhythmic: halting speech <ARH> </ARH>

Voice quality
Whispered <WH> </WH>
Breathy <BR> </BR>
Creak %
Creak during speech % (e.g. %two %words)
Crying <CRY> </CRY>
Yawning <YWN></YWN>
Quotative <QUOT></QUOT>
Smiley voice <SM> </SM>

PHONETICS

Phonetic/phonemic transcriptions (/ /)

TRANSCRIBER’S PERSPECTIVE

Uncertain hearing # (e.g. #two #words)
Researcher’s comment (( ))
Indecipherable syllable #

SPECIALIZED NOTATIONS

Duration (in seconds) <D:1.2>
Intonation unit continued &
Restart {Capital initial}

SOME SPELLINGS AND GLOSSES

uh, unh, um hesitation (filled pause)
m, hm awareness, wonder, backchannel
huh, hunh awareness, wonder, backchannel
mhm, unhhunh, uhuh backchannel or affirmative
response (final syllable stressed)
unh-unh negative response (initial syllable stressed)
uh-oh alarm cry

EMBODIMENT

Loosely based on Goodwin (1980) and Lerner (2003)

GAZE SHIFTS AND GAZE DIRECTIONS

Consider the notations together with the following example

19  ABIE:
20  MAUREEN: Oh my *go=esM__As_ [...] ...[1.4] 'Ten years *o-id:]
22  TERRY: [I was so im^pressed]...[I mean this *kid:]
24  TERRY: [I was so im^pressed]...[I mean this *kid:]

150
Capital letter followed by small letter indicates a participant looking at a co-participant, e.g.

Abbie looking at Terry
Terry looking at Maureen

- e.g. in the extract from example (3b) below, in line 20, Abbie begins to look at Terry and in line 22, Terry begins to look at Maureen in line 24.

Consecutive commas indicate gaze shift
Gaze shift from one point to another
- e.g. in line 24, Terry moves her gaze slightly before looking at Maureen.

Consecutive colons indicate the phenomenon being described
Phenomenon being described
- e.g. the mutual gaze between Maureen and Terry in lines 22 and 24.

The capital X stands for a participant looking at a gaze point
Participant looking at gaze point
- e.g. Terry looking at a picture in line 24 after slight head movement.

Consecutive underscore signs indicate a continued gaze direction
Continuing gaze at something or someone
- e.g. Abbie looks at Terry in line 20 and stops when Terry has uttered the 'impressed'.

Absence of line indicates that interactant looks neither at a co-participant nor toward a shared gaze point
Participant not looking at participant or shared gaze point
- e.g. Abbie after looking at Terry in line 20.