# Gaze behavior in face-to-face interaction

# Gaze behavior in face-to-face interaction

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. mr. S.C.J.J. Kortmann,
volgens besluit van het college van decanen
in het openbaar te verdedigen op vrijdag 5 oktober 2012
om 10.30 uur precies

door

**Federico Rossano**
geboren op 27 mei 1979
te Bologna (Italië)
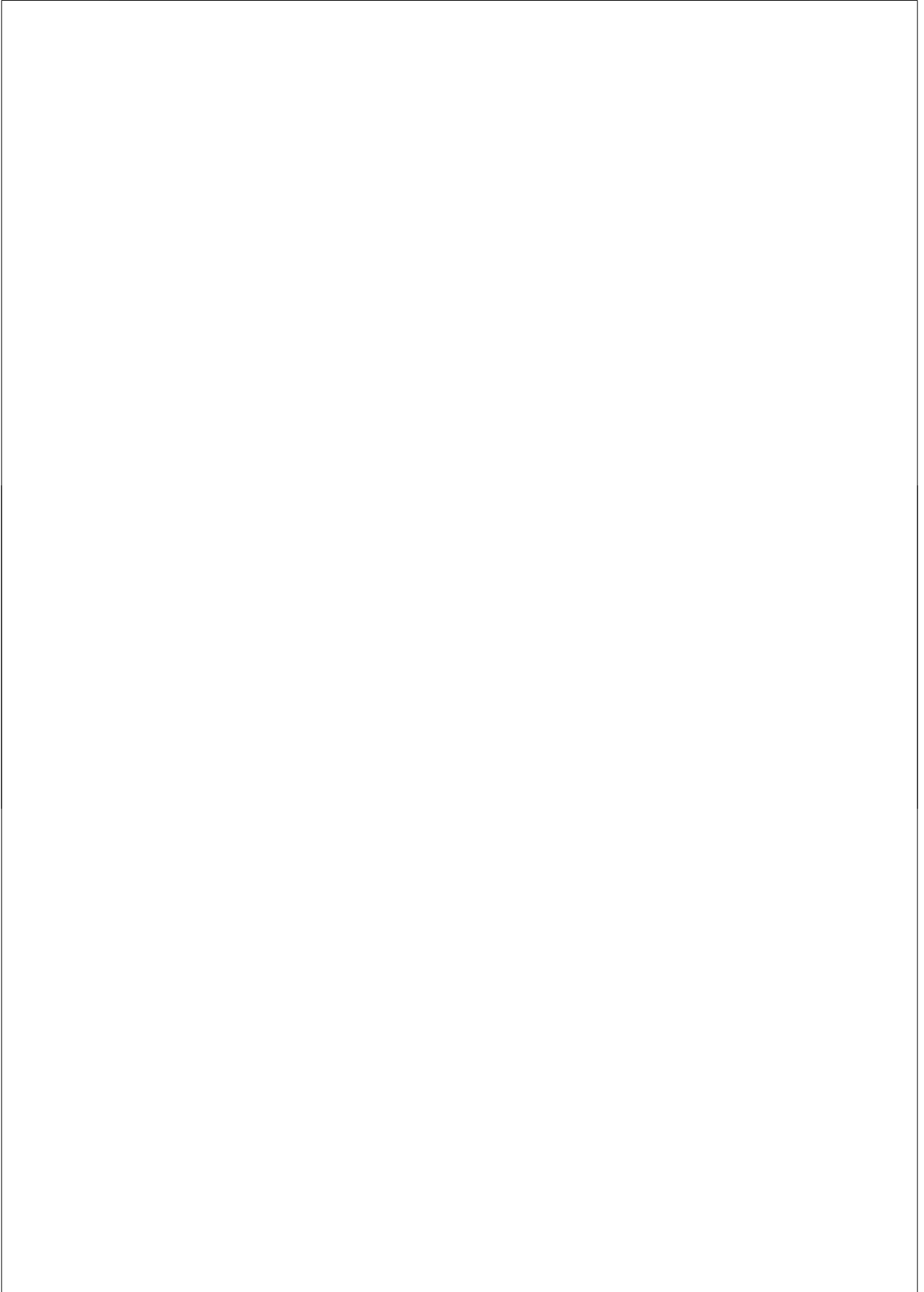
## Acknowledgments

"It was the best of times, it was the worst of times, it was the age of wisdom, it was the age of foolishness." It was my life as a Ph.D. student. Exciting, challenging, unpredictable.

Some people say that, when they were young, they already knew what they wanted to be as grown ups. As children, they experienced a moment that made clear to them what they wanted to pursue in life and everything would naturally follow from that crucial moment in their childhood. I knew it too. As a child, I wanted to compete in the Olympic games as a swimmer. I didn't want to win the Olympic games, I just wanted to participate. In my eyes as a young child, being amongst the top swimmers in the world was like the maximum realization for a human being. What else could one want in life? It took me many years of hard training and poor performances in the pool to realize that I was not talented enough to make it to the Olympics. But the desire remained.

Throughout my (still young) academic life, I have been fortunate to meet and learn from many incredible scholars, but only when Steve Levinson offered me a Ph.D. fellowship to work in his group, I thought that I had almost realized my childhood dream. Becoming a researcher at a Max Planck Institute is, to my eyes, the closest thing to participating in the Olympics that I will ever experience, for it is hard to imagine a research center with as high a concentration of young, bright international scholars with so much talent. Thank you, Steve, for thinking that I could ever deserve to be part of such an impressive research family. I have made Steve's life as a supervisor incredibly difficult and I have resisted his advice in many ways (the "age of foolishness" of the incipit). Despite these difficulties, each time I went to his office to discuss my work-in-progress, I always left with the feeling that his comments had suddenly unlocked some secret about my data that I had been incapable of seeing until that very moment. When I am trying to see the big picture in anything I study, I ask myself, "What would Steve say about this?"

If I have been difficult with Steve, words cannot begin to describe my behavior toward my supervisor Tanya Stivers. No doubt she never expected so much stubbornness from a student. She taught me how to make sense of my data, how to give a good presentation and how to write a good paper, and she helped improve my English the hard way: via countless corrections in

i

early drafts of this manuscript. She has been the toughest supervisor I could have ever imagined, but if I am now capable of effectively participating in academic discourse, I owe it mostly to her.

Speaking of tough supervisors, it is difficult to put to words just what having Manny Schegloff as a supervisor has meant to me. He embodies Conversation Analysis and it has been the biggest honor and privilege to be able to discuss ideas with him and to have him commenting on my work with his famously merciless red pencil. He taught me more than a few sentences could ever capture. Most importantly, he has been supportive of me every step of the way, even when we strongly disagreed. I have always been so impressed by his charisma, his intellectual integrity and his analytical mind that when in 2002 he asked me about my plans for the future, I confess having answered: "in the future, I want to be you". Although I will never succeed, just like my desire to be in the Olympic games, I still hope to become Manny Schegloff one day.

There are four other people who have not played a supervisory role in my studies, but who have nonetheless inspired me in unimaginable ways: Adam Kendon, Charles Goodwin, Jack Sidnell and Gail Jefferson. I have had the privilege of interacting with and learning from Adam and Chuck for many years and it is no surprise that my dissertation is on gaze in interaction. Whenever I disagreed with these Masters about the organization or function of gaze in interaction, that disagreement stemmed from a desire to take their lessons seriously: be honest with your data and never take previous claims for granted. Jack has been a friend and a model for how a young academic should be. He had the patience to engage in endless discussions with me on any topic and to bear, as book editor, with my tendency to see deadlines as suggestions, rather than obligations. I met Gail in the Netherlands in 2004 and I had the privilege of participating in monthly data sessions with her until her untimely departure. I will never forget her energy, her sharp comments and her directness. I wish I could look at data with her one more time. I miss you, Gail! All of them remain, together with my supervisors, the giants and I am simply an apprentice.

The members of the Language and Cognition Group at MPI have been teachers, colleagues and friends to me, and their talent continues to amaze me. I have been fortunate in finding friends in all of the groups at the MPI and I felt supported by all of them, not just academically. Thank you Mandana, Marianne, Elizabeth, Malte, Lara, Paul, Lilla, Tilman, Kaoru, Clair, Gertie, Mark, Giovanni, Keren, Sarah, Jan Peter, Laura, Annelie. Two of them hold a very special place in my heart and I cannot thank them enough for agreeing to be my

# Contents

iv

*List of Abbreviations*

APBS: Adjacency-Pair-Based Sequence

BS: Base Sequence

CA: Conversation Analysis

ETS: Extended-Telling Sequence

FPP: First Pair Part

Post Ex. BS: Post-Expansion of the Base Sequence

SCT: Sequence Closing Third

SPP: Second Pair Part

STS: Superior Temporal Sulcus

TCU: Turn Constructional Unit

TRP: Transition Relevance Place

**In transcripts:**

Cl.: Clitic

Throw.2s: Verb Throw 2nd person singular

# 1   Introduction

**One might have thought of sight, but who could think of what it sees?**
Wallace Stevens *Esthetique du Mal* (1944)
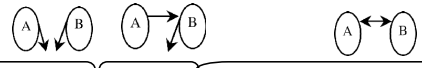
## *1.1   The Problem*

What does it mean to be able to see and to be seen by others?  What do we do with our eyes when we talk to one other? This dissertation investigates how people use their eyes and their body during face-to-face interactions and outlines the orderliness of their behavior. In particular, it shows how individuals from a specific culture (Italians from Emilia-Romagna, a region in Northern Italy) deploy their gaze during dyadic naturally occurring interactions in their homes. The aim is to outline some of the practices and norms that those individuals are sensitive to, with the goal of providing a detailed description of the orderliness of gaze behavior in human interaction and refining our understanding of participation in social situations. Investigating gaze in face-to-face interaction helps us better understand how participants manage rights and obligations in interaction, specifically with respect to cooperation, agency and the accountability of social behavior. It also highlights the fine-grained calibration and organization of turns-at-talk and sequences of social actions. A more specific goal of the analysis is to demonstrate that gaze behavior may differ when all other aspects of the interactional scene but the sequential context remain the same. The analysis focuses on how participants use their gaze at the beginning, middle and end of sequences of talk. The findings are situated in a comparative and evolutionary context, contributing to a more general framework for the understanding of how human beings use their eyes in interaction. Additionally, they provide important information for computer scientists and experts in artificial intelligence, in terms of the predictability of gaze behavior in humans and its potential reproducibility in robots or avatars.

First, it is useful to take a closer look at the phenomena that will be discussed here. Example 1.1 is an extract from a conversation between two young women (a university student, A, and a recent graduate, B, both in their early twenties). This conversation occurs in the kitchen of the student's house. The two participants are good friends, have known each

1
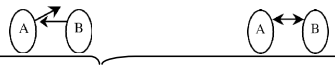
other for several years and are sitting at a table, at a 90-degree angle from each other. When the conversation begins they both have an empty glass and a can of soft drink in front of them. For the moment, the focus should be on how they use their eyes while they are talking to each other, considering what kinds of social action they are performing during their conversation. The conventions used in the transcript and in particular the meaning of the oval symbols are explained in Appendix B and in § 1.5.1, but basically, the arrows show whether the interlocutors are looking at each other, up, down, or at specific objects in the surrounding environment.
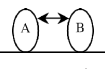
**(1.1) 2GGOSS-stasera 00:15[1]**
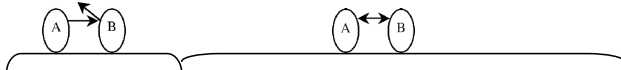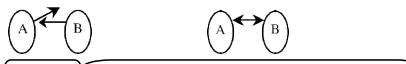
```
01 B:  .hh A(h)l l o r a stas(h)era cos' e' che  fate
            So            tonight    what is that do.2s
       .hh S(h)o t(h)onight what are you doing

02 A:  Eh    andiamo a  Villa Chiara=
       Eh    go.1p   to Villa Chiara
       Eh    we go to Villa Chiara=

03 B:  =Ma  a  che   ora  vi  incontrate
        But at which hour you meet.2p
       =But at what time do you meet

04 A:  Vado  alle nove e  mezza dalla Gloria.
       Go.1s at   nine and half  to   Gloria
       I go to Gloria's (house) at nine thirty.

05     ((B annuisce))
       ((B nods))
```

---

[1] See Appendix A for more information concerning this and all the other interactions analyzed in this dissertation.

2

06   ((some turns with side sequence omitted))

07 B:   *[Nove e   mezza ma] andate subito    a  Villa*
        Nine and half  but go.2p  immediately to Villa
        **[Nine thirty but] do you go immediately to Villa**

08      *Chiara alle nove e   mezza?=*
        Chiara at   nine and half
        **Chiara at nine thirty?=**

09 A:   *=Con- le  dieci con  gli altri.*
         With the ten   with the others
        **=With- at ten with the others.**

10      (0.4)

11 A:   Ci  *incontriamo.*
        Cl. meet.1p
        **We meet.**

12      (0.6)

13 A:   *Ci [v i e n i?*
        Cl. come.2s
        **Will [you come? ((to Villa Chiara))**

14 B:   *[Io   esco      alle nove_ (0.5) ah io pero' se arrivo*
         I    get.1s out at   nine        oh I  but  if arrive.1s
        **[I    get out (of work) at nine_ (0.5) oh but if I come**

15    *arrivo    a mezzanotte eh*
      arrive.1s at midnight eh
      **I arrive at midnight eh**



16    (0.6) ((A displays shocked facial expression see Fig. 1.1))



Figure 1.1. A's facial expression during line 16.



17 A: *Alle nove esci?*
      At    nine get.2s out
      **At nine you get out (of work)?**



18 B: *E[h*
      Eh
      **Y[es ((confirming something already said))**



19 A:   *[Cazzo* ((Poi A annuisce, guardando in basso))
          Dick
         **[Shit ((and then A nods, looking down))**



Figure 1.2          Figure 1.3

20    (0.6) (Fig. 1.2)    (1.2)  (Fig. 1.3)

**(1.8)**

21 A:  *Domani   sera    che   fai?*
       Tomorrow evening what do.2s
       **Tomorrow evening what do you do?**

22     (0.5)
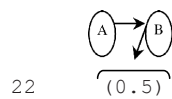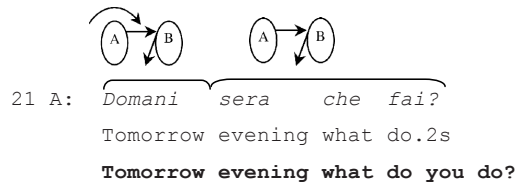
The first things to note concerning this interaction are the setting and the spatial configurations that make this social situation possible. This interaction occurs in the kitchen room of a house, with the two participants sitting at a table with soft drinks available (see Figures 1.2 and 1.3 in the transcript). It is known that the place where a conversation takes place, the seating arrangements, the possible occurrence of drinking and eating, the socio-economic status of the participants and the kind of relationship that exists between them are all factors that might affect their conversation. Many studies have shown, for example, that participants tend to arrange themselves according to the interactions they plan to have (Ekman & Friesen, 1974; Kendon, 1977; Sommer, 1959, 1962). While it is true that the physical structure of the environment constrains interactional positioning (Goffman, 1971; Goodwin, 1981), these physical arrangements tend to directly reflect preferred cultural patterns. It is clear, indeed, that a number of these aspects might be typical of specific cultures such as Western ones (e.g., the fact that when two friends want to have an extended conversation they tend to sit at a table and have food or drinks available). Yet, what might be culturally specific about the environment and the modalities through which this meeting is arranged should not hide a more generic fact, which is that when participants want to have a conversation, they arrange their bodies close to each other with a specific function, beyond hearing better what the other might be saying.  Usually, they position themselves in "an eye-to-eye ecological huddle, which tends to be carefully maintained, maximizing the opportunity to monitor one another's mutual perceivings" (Goffman, 1963: 95), so that each participant can see the other's face, but also what the other might be looking at and doing. All these aspects will be further discussed later.

5

A second thing to note in Example 1.1 is that A and B do not follow simple gaze patterns. It is not the case that every time A or B starts speaking they look down or away nor that they tend to look toward the addressee only when they approach the end of their turns of talk (see Kendon, 1967 for a claim that might suggest the opposite). Instead, they direct their eyes in many directions. Mostly they look at each other, sometimes they look away, but when they look away, it is often to look at objects that will be picked up later in the interaction. Sometimes they look up toward one side, displaying a thinking face (M. H. Goodwin & Goodwin, 1986). It also often happens that one looks at the other while the other is not looking back, yet this seems to be often a transitory phase, frequently followed by engagement of mutual gaze. In other words, gaze directions can be tracked and possibly explained, but the patterns they follow are not obvious. On the other hand, one thing that remains remarkable is how "ordinary" (Sacks, 1984b) their gaze behavior appears to be for the other participant and for anybody observing them. The two participants here are focused on the content of the conversation and the social actions performed in it (in this specific case, the main project is figuring out whether the two participants can go out together that evening or not), while the way each one uses her eyes is probably less central to their concerns. They certainly do not mention it at any point during the 45 minutes of conversation that follows. And this is what we would expect, because presenting our gaze behavior as ordinary is indeed what we wish to achieve during our daily interactions: we do not want participants to find our gaze behavior marked or deviant. There might be occasions in which this is not the case (e.g., when we are trying to flirt with someone or threaten someone), however, most of the time we attempt to avoid a "non-ordinary" use of our eyes.

Like gestures, gaze is difficult to describe and yet appears to have a prominent role in our daily interactions. As Sapir (1963 [1927]: 556) wrote with respect to gestures, which we can easily extend to gaze:

> *Gestures are hard to classify and it is difficult to make a conscious separation between that in gesture which is of merely individual origin and that which is referable to the habits of the group as a whole. In spite of these difficulties of conscious analysis, we respond to gestures with an extreme alertness and, one might almost say, in accordance with an elaborate and secret code that is written nowhere, known by none, and understood by all.*

6

Cracking the gaze code, however, is not an easy task. The task of this dissertation is to make explicit how we come to produce "ordinary", not marked gaze behavior in conversation. This means that by reaching the end of this work, it should be easier to detect and explain the regularities present in example 1.1. The findings of the next 3 chapters will provide a detailed account of what is going on in terms of gaze behavior and courses of action in example 1.1 (this example will be presented again as 3.3 in chapter 3 and 4.13 in chapter 4, to show how the different gaze practices under examination get implemented in the production of this extract).

In perfect agreement with what Birdwhistell (1970: xi) claimed 40 years ago, this dissertation posits that "body motion is a learned form of communication, which is patterned within a culture and which can be broken down into an ordered system of isolable elements." Yet, as it will be explained later, any body motion has to deal with the physiological limits of our bodies, which have evolved in a specific way for clear adaptive purposes and are common to all human beings (e.g., we have approximately 200° and not 360° visual range). Keeping in mind that our gaze behavior is partly affected by our biology and partly "a learned form of communication patterned within a culture", it is important to specify some of the specific assumptions—grounded in previous empirical observations—that this dissertation is based on.

The first assumption is that human beings did not evolve eyes to be able to communicate better. We have eyes to perceive the world around us, facilitate our movements and the detection of prey and more generally to coordinate body movements within the environment. As Gould and Lewontin (1979) explained with the metaphor of "the spandrels of San Marco", one could think of the decorations in the spandrels (see Figure 1.4)—the approximately triangular spaces between two arches and a horizontal molding or cornice above them—as the main goal of the architect and the arches that create the spandrels as a consequence of that initial artistic goal. Alternatively, and more reasonably, one could assume that the architect wanted to create some arches, particularly functional and aesthetically pleasing in the architecture of the building. But once the arches had been built, the spandrels were there and allowed for further decorative opportunities. Similarly, humans, as primates, evolved eyes mainly for perceptual purposes. But once we had eyes, evolution facilitated the possibility of also deploying them for other purposes, such as communicative ones. The fact, for example, that our sclera (the white of our eyes) is not pigmented—being

white rather than dark, as in every other primate species—appears to be an adaptation to facilitate the detection of our eye movements by conspecifics (see § 1.2.2 for more details).
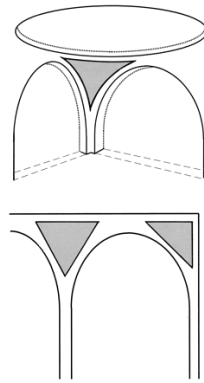


Figure 1.4. Schematic representation of a three-dimensional and a two-dimensional spandrel.[2]

The second assumption is that looking at somebody's face is usually more interactionally relevant than looking anywhere else during a conversation (at least in Western cultures). Studies based mainly on American English and European languages (e.g. Argyle & Cook, 1976; Argyle & Dean, 1965; Argyle & Graham, 1976; Exline, 1963; Goodwin, 1981; Gullberg & Holmqvist, 2006; Kendon, 1967; Nielsen, 1962) show that participants spend a significant amount of time looking at each other's faces when they interact. Argyle and Graham (1976: 6) emphasized that "in our thinking about the processes underlying social interaction we have come to believe that gaze at other persons is of central importance". However, in dyadic interactions in laboratory settings they also showed that while background stimuli had an unreliable effect on gaze, objects relevant for the subject's task attracted gaze for large amounts of time and therefore reduced gaze at the other person. Moreover, Goodwin (e.g., 1981, 1984) has claimed that looking at objects in the nearby environment used for accomplishing activities competing with the talk (e.g., drinking, smoking, eating) is less problematic than looking away in general, as long as it remains a brief disengagement from looking at the other participant. Nonetheless, the claim that, in general, looking at each other's face during a conversation should be more relevant than looking at any other element in the environment not directly relevant to the task-at-hand has been challenged cross-culturally. For example, the work of Rossano et al. (2009) shows how members of a Mayan community living in Mexico and speaking Tzeltal tend *not* to look at

---

[2] Image from Gould (1997: 10751).

8

speakers while listening to them, so that *not* looking at anything in particular while listening to somebody speaking is the default home position for the eyes, and looking at some object in particular might be considered a sign of disattending the conversation.

The third assumption is that participants' gaze behavior in interaction is related to the fact that the participants are talking to each other. In particular, the claim is that looking at each other is part of and contributes to the development of interactants' conversation. There is no doubt that individuals monitor the environment around them regularly; however, when two people are co-present mutual gaze creates an opportunity for them to initiate conversation. In general, individuals tend to avoid looking at each other during silences unless the silence constitutes just a pause or a gap (Sacks, Schegloff, & Jefferson, 1974: 715) in the conversation, rather than a full disengagement. This assumption of a correlation between talk and gaze behavior is the basis of Kendon (1967) and Goodwin's (e.g., 1979, 1980, 1981) work on the organization of gaze in interaction.

Finally, the fourth underlying assumption of this work, again shared by Kendon and Goodwin and previously emphasized at the beginning of this introduction, is that gazing at someone is also a social act and not merely a perceptual tool to better understand speech. Nonetheless, if we do not understand how our eyes can also be used as a perceptual organ, we are bound to lose a powerful tool to understand what might be going on during a social interaction.

The general claim of this dissertation, then, is that gaze (in Italian conversations, at least) can be understood in terms of action formation and a sequential analysis of actions in interaction (see, e.g., Schegloff, 2007b), and not simply in terms of turn-taking or displays of engagement and attentiveness as has been proposed (e.g., Goodwin, 1981). This work will show that gaze is used, for example, to coordinate the development and closure of sequences and courses of action, to pressure for responses and pursue them and to indicate special states of recipiency. Moreover, it will show that each individual deploys specific gaze behaviors according to her/his role as speaker or recipient but also in relation to what s/he is trying to achieve during a conversation. Mutual gaze, in this sense, will be interpreted as an interactional achievement that results from the behavior of each participant performing specific gaze patterns, rather than as a goal toward which participants are normatively driven. So speaker gaze toward the recipient might have very different motivations than recipient gaze toward speaker. And yet if both perform such behaviors, they will end up looking at each other's eyes and thus engaging in mutual gaze.

The news here is not simply that to understand gaze in interaction one needs to take into account what the social situation is and what people are trying to do, in terms of physical and social actions, at a micro level. The news consists in suggesting that there are multiple levels of order that play out at the same time and that can be comprehended through a careful and detailed analysis of people's use of their eyes during conversations. Gaze, indeed, is used to perceive the world, to control the accuracy of our body movements and other people's movements in the environment, to display attentiveness and engagement, but also for structural communicative purposes as well. Too often two different functions of gaze behavior are emphasized in scientific investigations: gaze behavior as *responsive* to environmental demands (e.g., attention) vs. gaze behavior as *shaping* the interactional environment and the actions accomplishable (e.g., projecting next action). To simplify the dichotomy, one could say that gaze is usually either studied as "a measure of" something or "as a means for" something. Yet it is clear that in a single stretch of interaction, the same individual will be confronted with the need to pay attention to the environment and still use her/his eyes for her/his communicative needs. A more fruitful analysis of gaze behavior in social interaction, then, should be aware of the importance of each component, informed of their limits and biases and somehow make sense of both functions by extracting some basic rules concerning gaze deployment in social interaction. Therefore, the logic followed in this work requires trying to unravel the gaze machinery, the system of norms, biases and habits that guide people's eyes and actions during a conversation. Once the system is known, we may be able to focus on what the specific participants in a specific interaction are trying to do. As Sacks (1987: 67) said:

> *You cannot find what [people] are trying to do until you find the kind of things they work with. If the system had a built-in bias for agreement, and you did not know about it, then you'd be counting a whole bunch of things as agreements that might well be accounted for in other ways. [...] Now, what they* wanted *to do is another question, and it is only, I think when we begin to have some considerable idea about the sort of things they are operating with (like a preference for agreement), and the sorts of ways they operate with those things, that we have much of an idea about such matters as 'what they're trying to do'.*

In other words, an investigation cannot start from trying to interpret what people are trying to do with their eyes, but rather in documenting the patterns and regularities of their behavior,

10

including whether and to what extent their behavior is associated with what we know participants are doing through their talk and bodies.

As Lorenz (1966: 273) suggests, while proposing the use of a biological approach to the study of behavior:

> *Biological approach and method can successfully be applied in the study of behaviour, exactly as Darwin had done in his book,* The expression of the emotions in man and animals*. What, then, are these good old Darwinian procedures?*
>
> *I would name the most obvious one first, as being most important. It is the unbiased observation of the organic system and the inventorizing of its component parts. In all natural sciences description has to precede systematization, and both together are the prerequisite for abstracting the natural laws prevailing in the operations of the whole.*

Observation and description always have to precede systematization. As will be shown in every chapter, any initial description of a possible practice will then be followed by a systematic search for more instances of the practice and an attempt to provide falsifiable hypotheses about its functions, to be tested via quantification. Finally, and not in every instance, a question can be raised concerning how participants can detect the current unfolding of a specific practice or how that practice can come to be; that is, how that practice might evolve.

An example of how the latter point could be approached, and of the importance of taking into account more than one level of organization, comes, for example, from ethology, in the words of Lorenz (1966: 276):

> *A phylogenetically adapted motor pattern which originally served the species in dealing with some environmental necessities, acquires a new function, that of communication. The primary function may persist [...]. In many cases, however, the primary function recedes into the background or disappears altogether so that a complete change of function is achieved. Out of communication, two new, equally important functions may arise [...] the first of these functions is the canalization of aggression in a manner permitting its discharge without damaging fellow members of the species [...] the second is the formation of a bond which keeps together two or more individuals. This is achieved by most so-called greeting ceremonies which an animal can perform only with a certain, individually known partner, whose presence, for this reason, becomes an indispensable need in the animal's life. It is quite erroneous to say that such ceremonies are "the expression of" a bond; indeed they themselves constitute it.*

*The second characteristic of ritualized motor patterns is a change of form which the unritualized prototype underwent in the service of its new communicative function and which quite obviously was brought about by the selection pressure exerted by the survival value of communication.*

This quote suggests that there are multiple elements to take into account to understand the communicative system we use, how it came about and how it works. From an evolutionary perspective, the emergence of a communicative pattern presumably requires some level of ritualization of a behavior that originally had some specific function other than communication becomes a communicative signal by getting reinforced and modified in its meaning via repeated interactions with conspecifics. Lorenz suggests that among the new functions of a communicative act, the canalization of aggression and the development of a bond may occur through the occurrence of this ritualized act. Additionally, the act is not (just) "an expression of" the relationship between individuals, but rather it "constitutes it". We do not act in a certain way simply because we are friends, for example, but rather, by acting in a certain way we reinforce our friendship and display to one another the kind of relationship we think we have or want to have. That is, we are "expressing" and "constituting" friendship in every act we perform toward a potential or current friend. The two additional functions mentioned above are particularly relevant for a work on gaze in interaction because looking at someone in the eyes is a highly arousing act, e.g., skin conductance increases when gaze is directed toward the observer (Hietanen, Leppanen, Peltola, Linna-Aho, & Ruuhiala, 2008; Nichols & Champness, 1971) and people often look at each other's eyes to threaten each other (aggression) and to flirt (create intimacy) (Eibl-Eibesfeldt, 1989). Many authors have claimed a relationship between gaze and emotional arousal in humans, in particular claiming that gaze aversion is a way to reduce emotional arousal, both in adults (e.g., Ellsworth & Ludwig, 1972; Kendon, 1967) and in young infants (e.g., Brazelton, Koslowski, & Main, 1974; Field, 1979; Stern, 1974), and showing, for example, that in 4 month olds, gaze aversion is usually preceded by a heart rate increase and their heart rate decreases while the gaze is averted from the adult's face (Field, 1981). Yet, aside from using our eyes for threatening or creating intimacy, we also tend to look at each other when we talk, and this is not just to better parse the sounds we produce, but also to display attentiveness or commitment, to pressure for response, to select next speaker, to monitor each other and so on and so forth. Mutual gaze, then, may indeed be one of the motor patterns originally deployed with a specific function that then obtained new functions, by

canalizing aggression, helping in establishing a bond and more generally allowing human beings to do other things with it, making their interactions more effective.

It was suggested earlier that the study of gaze in interaction needs to take into account multiple levels of organization, multiple systems that affect the final deployment of people's eyes, and ought to develop a more fine grained model of when and how people move their eyes, with the hope of understanding why as well. Considering some of the current theories, as far as levels of organization needed to understand human behavior in social interaction is concerned, three different positions stand out. One insists that there is mainly one core level of organization, the *interactional* level that shapes every other level (see, e.g., Schegloff, 2006). A second one suggests the existence of three different systems that interact at the same time, the *socio-cultural* system, the *interactional* system and the *language* system (see, e.g., Levinson, 2005). Finally, a third one also focuses on three levels of organization, *cognition*, *interaction* and *culture*, which are usually investigated by three different disciplines, but are necessarily intertwined, and, as such, require a more complex kind of investigation (see, e.g., Enfield & Levinson, 2006b; Sperber, 2006). In line with the above ethological concerns, I propose an approach that requires a basic understanding of four different levels of organization:

1. Individual
2. Micro-social (dialogical/conversational)
3. Macro-social/Cultural
4. Species

To understand what happens in any single interaction, one should take into account the culture a participants belongs to, his/her socio-economic status and relationships, the rules of conversation or of politeness, idiosyncratic behaviors, cognitive abilities and competences of the single individual, as well as her or his own desires and beliefs. That said, one should also realize that we are biological creatures, that we have bodies with limits and specific capacities that necessarily affect what we can and cannot do when faced with a problem or a specific situation. For example, we are endowed with a specific visual system that allows us to perceive the world, act in it and move through it. We are endowed with two eyes positioned frontally, rather than laterally. The limits of our bodies are the limits of our natural repertoire of possible movements. All of these biological developments came about through evolution, in adaptation to the contingencies of a constantly changing environment. If we

neglect these aspects of human biology and physiology, we may provide accurate descriptions of what we can see people doing, without necessarily knowing why it is so, how it came about and most importantly, what the other possibilities may be, whether other movements or other actions could easily emerge. In addition, we might interpret acts that occur only because of the limits of our bodies as intentionally communicative. For example, having the capacity to only see 15-30° of the visual range with proper visual acuity, out of a 200° range (see, e.g., Bruce, Green, & Georgeson, 2003), means that we need to move our eyes often if we want to ensure we accurately perceive what is happening in front of us, or even just for reading this written page. But these eye movements obviously do not occur in order to intentionally communicate to other participants which part of this page we are reading. Rather, it is a necessary requirement for our eyes to be able to perform this task. It can be informative for a recipient, but it is not designed for the recipient, as will be briefly outlined in what follows.

What is proposed here, though, is not an embodied theory of social interaction. What I am suggesting is an approach that recognizes that the machinery, the repertoire that has emerged and has been reinforced through individual ritualization and social imitation — making a specific behavior a cultural praxis through historical sedimentation — has a biological foundation that continues to exert pressure on what the limits of those movements are and what the limits of our interpretation of those movements should be. Moreover, if we find that some practice is shared by all cultures and all human communities, it may be because this practice might have some strong adaptive function and be evolutionarily the best adaptation for our species, and not just for our specific community. To summarize, if we want to be able to better understand what humans do when they interact with each other, we need to keep in mind the following four points:

1. We have individual cognition and individual memories of our previous
   interactions.
2. We use the interactional machinery we believe we share with other interactants and
   we recipient-design our communications and actions, because of what we know about
   others.
3. Our actions are informed by what we deem to be socially and culturally
   appropriate.
4. We are biological organisms with specific limits and capacities, and our bodies
   have evolved for specific adaptive purposes.

To achieve an analysis that keeps these four points in mind, a broad literature has to be introduced and discussed.[3] This will constitute the background against which the empirical chapters will be built. It will also provide a glimpse of the theoretical underpinnings of this investigation. To understand how a plane can fly, it is not enough to look at it, and yet without looking at it, we would not be able to explain how that specific plane can fly while another might not. In the same way, in order to understand how we use our eyes when we interact with each other, aside from describing the orientation of the gaze of the participants, it is useful to know something about the hardware (i.e., our eyes), about vision and about human sociality to make sense of how and when we direct our eyes during a conversation. To begin, then, a first question is: what difference does it make that human beings have eyes?

## 1.2   To See: An Evolutionary Perspective on Sight

In 1859 while outlining his evolutionary theory about the origin of species by means of natural selection, Darwin raised a distinct and serious issue for his own theory: the existence of the eye. In a section specifically dedicated to the eye called, "organs of extreme perfection and complication," he wrote:

> *To suppose that the eye, with all its inimitable contrivances for adjusting the focus to different distances, for admitting different amounts of light, and for the correction of spherical and chromatic aberration, could have been formed by natural selection, seems, I freely confess, absurd in the highest possible degree. [...] If it could be*

---

[3] I follow here the model outlined in the volume Roots of Human Sociality (Enfield & Levinson, 2006a) where contributors and contributions from a range of disciplines were combined to address a single, broad question: what are the roots of human sociality?

> *demonstrated that any complex organ existed, which could not possibly have been formed by numerous, successive, slight modifications, my theory would absolutely break down. (1859: 186-189)*

He also suggested, however, what kind of evidence would support his theory, even if it was difficult to envision rationally at the time:

> *Yet reason tells me, that if numerous gradations from a perfect and complex eye to one very imperfect and simple, each grade being useful to its possessor, can be shown to exist; if further, the eye does vary ever so slightly, and the variations be inherited, which is certainly the case; and if any variation or modification in the organ be ever useful to an animal under changing conditions of life, then the difficulty of believing that a perfect and complex eye could be formed by natural selection, though insuperable by our imagination, can hardly be considered real. (1859: 186)*

What Darwin was looking for was a range of organs less sophisticated than the human eye that could still be used for vision and of any use for the animal possessing it, maintaining that the modifications could be inherited. Recent evidence seems to have solved this Darwinian puzzle.

The first eyes that appeared on Earth belonged to the trilobite, a voracious predator already known at the time of Darwin's writing. It can be dated to about 543 million years ago. This is not to say that this is the first mechanism ever developed to perceive light, as there were already single-celled organisms like amoebae that could detect sunlight and were using this information to distinguish up from down in the water. Moreover, there were also multi-celled animals that had light sensitive organs, called *ocelli*, which were similarly used to detect light direction. The difference, however, is that those organs were not capable of forming images; they could only detect light direction (like plants). Vision is certainly the most sophisticated mechanism to detect light direction, but in addition to it, it allows us to detect images through the retina. The retina is a multilayered tissue made of nerve cells that convert light signals into neural signals and through the optic nerve communicate directly with the brain. With the help of some apparatus such as a lens, it is possible to detect with some precision any image that is projected onto the retina. This allows the perception of shapes and therefore helps in the detection of other animals moving around the environment.

It has been claimed (Parker, 2003) that the development of an organ such as the eye changed the face of the planet by allowing some animals to perceive others much better and allowing them to move better and faster through the environment. In this sense, it becomes easier for the predator to hunt and the prey must develop other means to defend itself. Noticeably, apart from the presence of outer defensive hard parts such as shells and skeletons, animals also developed another outstanding resource: color.[4] Indeed, once light and images can be perceived with some acuity, color perception becomes possible for animals, and nature evolved in accordance with this new capacity. This is of great interest if we think about the amazing variety of colors present in nature and the fact that developing and changing colors is costly for any organism. What would be the point of, for example, the mimetic behavior of the chameleon or the colorful plumage of so many birds or even the amazing variety of colors of so many fruits and flowers, if these colors could not be perceived? The existence of all of the above shows how a great deal of the external appearance of animals and plants on our planet has developed as a result of the driving force of the sense of sight and the presence of sunlight.

Sight might help, but how advantageous is it? Only 6 phyla of animals have eyes out of the extant 33 phyla, so it might look like there was not much need for their evolution. On the other hand, about one-third of those 33 phyla have specialized organs for detecting light (Fernald, 2006; Land & Nilsson, 2002) and the 6 phyla that have eyes constitute about 96% of the existing animal species on Earth (Land and Nilsson, 2002). In other words, they have been by far the most successful ones in terms of natural selection and differentiation.

---

[4] Notice that color is a perceptual byproduct of light constructed by the brain and not a property of objects. It is the wavelength that the eye can perceive once light is reflected by a specific material surface (that partly absorbs and partly reflects the different wavelengths that compose light). So a blue object appears as such because it absorbs all of the wavelengths of light apart from the ones that our brain categorizes as blue. Different animals can perceive very different colors depending on the color receptors they have. So many mammals have monochromatic or dichromatic vision, whereas most primates, including humans, have trichromatic vision, and many other animals (e.g. many birds, see Bowmaker, 1991) can also perceive wavelengths in the ultraviolet spectrum, which human beings cannot perceive at all.

## 1.2.1  The Evolution of the Eye

How did the eye evolve? As mentioned previously, the capacity of detecting light direction is not sufficient to consider an organ an eye. There must be some photoreceptors that transform light into some sort of neural signal that allows the brain to perceive the image projected on the retina usually through the help of a lens. There are, however, functional eyes that lack a lens (nautilus) but allow the animal to form an image on the retina (and therefore "see"), by using the principle deployed by the ancient Chinese and described by Leonardo da Vinci as the "camera obscura".[5] The nautilus is indeed provided with pinhole eyes. Moreover, there are eyes lacking a pupil or lacking the cornea, but every eye must have, by definition, "specialized phototransducing cells" (Fernald, 2000). All vertebrates (e.g. mammals) have a single type of eye called "camera eye"[6] (because it works a bit like a camera), and human eyes consist of lenses, iris, cornea and pupils.

Nilsson and Pelger (1994) made a "pessimistic estimate" of how long it would take for a light sensitive spot to evolve all the way to a fully developed lens eye by focusing on the development of spatial resolution (i.e. visual acuity). They modeled a sequence of evolutionary stages of the eye that can be found among animals existing today and they found that the whole evolution would require approximately half a million years. This is particularly important for Parker's model, as outlined above, as it confirms that the development of the organ "eye" could indeed have happened rather quickly, at the beginning of the Cambrian and since then have tremendously affected animal evolution.

Out of the Cambrian emerged eight different types of functional eyes (Fernald, 2000; Land & Fernald, 1992; Land & Nilsson, 2002); that is, eight optically distinct ways of forming an image on the retina which results in seeing something. This variety has been even larger in terms of different types of eye organizations, according to others (e.g., Salvini-Plawen & Mayr, 1977).

---

[5] Light penetrates a dark environment through a tiny hole and forms an image on the dark surface opposite to the hole.

[6] Interestingly also octopuses have camera eyes but they developed independently from the vertebrate ones and are actually functionally better because they do not have a blind spot which is present in the vertebrate eye. This is often used as an example of "convergent evolution" because the ancestors they had in common did not have a camera eye.

The development of different mechanisms for processing light is particularly relevant for this investigation. Indeed, what vision does is extract information about the environment by using light and two of its properties, in particular: intensity and different wavelengths (frequency). Different animals have developed different solutions to the problem of processing light information and the solutions emerged through natural selection and therefore adaptation to environmental conditions. What matters here is the awareness that behaviors and practices are not the sole product of social negotiations or historical sedimentation, but rather the adaptive point of equilibrium for many needs and constraints in terms of coping with the world the individuals inhabit. This relationship between problems, evolution and development of specific behaviors can be easily understood by focusing for a moment on the fact that human beings have both eyes on the same (frontal) side of our head.

What difference does it make that we have both eyes at the front of our head rather than on the sides, like rabbits? According to some scholars (e.g., Cartmill, 1974; Parker, 2003), the position of the eyes on the head tells us something about the position of the animal in the food chain. Specifically, Parker (2003) claims that animals with eyes on the sides of the head are mainly prey, while animals with frontal eyes are mainly predators. Clearly, the generality of this claim is problematic because not all predators have eyes that point in the same direction (see Cartmill, 1992 on this), and therefore other factors must come into play. However, some of the reasons given for this distinction warrant consideration. For example, having lateral eyes means having a wider angle of vision, which allows for improved detection of movements, although the perception of depth and distance becomes more difficult. In this sense, seeing more and being more sensitive to movement might allow prey to react faster, and the depth of vision becomes less crucial for survival. On the other hand, having eyes on the front of our head, in particular having the angles of vision crossing each other, enables stereopsis (i.e. the perception of binocular depth). This is particularly important in terms of determining distance, for example, in pinpointing where an object or prey might be. According to some (e.g., Le Gros Clark, 1959; Smith, 1924), primates, together with other arboreal animals, developed binocular stereoscopic vision because it helped them in determining distances between branches of the trees. However, other explanations for the frontal position of the eyes exist. Noticing that the closeness of the eyes in the middle of the face reduces parallax and therefore the distance at which stereoscopic vision can work (i.e., one can perceive depth and distances well only at close range), Cartmill (1974) suggested that the convergence of the eyes must have evolved to allow better vision at close range, just like cats and other predators. The development of good sight, the loss of

claws and the development of grasping extremities would point toward insect eating and predation with the important aid of vision. Allman (1977) confirmed that the frontal position of eyes with a spherical lens favors visual acuity even more than stereopsis (animals with eyes on the side have more blurred vision), and this is particularly important at night, in low light. In other words, the predators that favored frontal eyes are mainly nocturnal predators, because for diurnal ones the lateral position of the eyes is less of a problem.

Alternatively, Sussman (1991) suggests that what would have favored certain evolutionary adaptation in primates is fruit foraging (better perception of edges, shapes and distance would help seeing them). Moreover, contrary to most other mammals, primates have trichromatic vision and this appears to have developed in relation to fruit foraging because it helped recognizing mature fruits through leaves and vegetation (see, e.g., Jacobs, 1995; Regan et al., 2001).

If binocular vision and stereopsis contribute to 3-dimensional vision and allow a better perception of shapes and distances, then it is no surprise to find out that human beings look at objects anticipatorily, before reaching for them. They do so in order to calibrate their movements (see, e.g., Hayhoe & Ballard, 2005; Johansson, Westling, Backstrom, & Flanagan, 2001; Land, 2006), and also to guide their movements while approaching an object to grab it (see, e.g., Jeannerod, 1986 on how finger posturing anticipates the real grasp, occurs during the transportation of the hand and partly relies on visual feedback). Looking at objects helps establishing distances and therefore facilitates the motor control of our hands and arms when we reach for things. Whenever we are involved in a task such as preparing a cup of tea (Land, Mennie, & Rusted, 1999) or a sandwich (Hayhoe, 2000), our eyes mainly perform four monitoring functions: i) *locating* objects used later in the process; ii) *directing* the hand or object in the hand to a new location; iii) *guiding* the approach of one object to another; and, iv) *checking* the state of some variable relevant for the task. Usually the eyes reach the object that we reach for slightly less than a second before manipulating it, and usually our eyes only visit objects that are relevant for the task-at-hand, independently of how conspicuous other objects in the environment might be. The eyes, then, help us to plan our movements and provide feedback on our motor performance; feedback that becomes less salient the more routinized the action becomes.

All of the above implies that when we observe individuals in interaction looking at their glass or a cookie before they reach for it, we are simply observing a behavior that has evolved because of the specificity of primate eyes and the advantages and limits of their frontal position. We will see how this phenomenon plays a role in everyday interactions in the empirical chapters that follow.

## 1.2.2  Eye Morphology in Primates and Eye Movements

Human gaze behavior is a highly evolved system: unique amongst primate species, the human orbit has evolved to display the sclera, in humans also referred to as "the white of the eye", and our eye outline is extraordinarily elongated in the horizontal direction. Human beings are the only primate species with a sclera without pigmentation (see Figure 1.5). The function of these physiological differences seems to be to make gaze direction discernable to others at a distance, as we have lost the "gaze-camouflaging" function common to other primate eyes and have enhanced the "gaze signaling" function (Kobayashi, 1997, 2001; Kobayashi & Hashiya, 2011; Morris, 1985). Some empirical evidence supporting this claim comes from an experiment run by Tomasello et al. (2007) in which human infants and other great apes were compared in terms of their tendency to rely on eyes or head direction in following the gaze of a human experimenter. They found that while human infants rely only on eyes, great apes rely mainly on head direction.  This suggests that the presence of a white sclera co-evolved with human cooperative interaction, making it easier for others to perceive what we are looking at (also at some distance), contrary to other great apes who mainly interact on a competitive basis.[7]

---

[7] In the seminal paper "On the human interaction engine", Levinson (2006) claims that human beings have a unique propensity to interact and their interactions have a distinctive structure, when compared to other species.

Figure 1.5. Human eyes vs. chimpanzee eyes.

Moreover, it has long been established that humans can judge the direction of other humans' gaze to within a few degrees of arc (Gibson & Pick, 1963). This capacity has been claimed to be crucial for the development of joint attention and human social cognition (see, e.g., Butterworth & Cochran, 1980; Scaife & Bruner, 1975; Tomasello, 1995, 1999), and for language acquisition (see, e.g., Bruner, 1983; Tomasello & Farrar, 1986; Tomasello & Todd, 1983). The capacity to follow gaze, which human infants already develop as early 3 months of age (D'Entremont, Hains, & Muir, 1997) and enables the possibility of establishing joint attention with another individual, is particularly important from a developmental point of view because it helps in understanding what others are thinking or intending to do (e.g., Baron-Cohen, 1995; Frith & Frith, 2001). Moreover, correct gaze following at 10 months predicts language comprehension and gesture production at 18 months, therefore suggesting a strong positive correlation between gaze following and language acquisition (Brooks & Meltzoff, 2005). Individuals with autism appear to have a different way of engaging in gaze following and also an impairment of joint attention (Carpenter & Tomasello, 2000; Charman, 2004; Dawson et al., 2004; Loveland & Landry, 1986; Mundy, 1995; Mundy & Newell, 2007). By the end of the first month of life, human infants show the emergence of "effortful" concentration on the mother's face (Lavelli & Fogel, 2005) and at 6 months of age, infants are sensitive to the social communicative value of gaze, as they follow gaze only if the gaze shift is preceded by an ostensive signal, such as gaze contact (Senju & Csibra, 2008). Further evidence of human infants' predisposition to detect gaze behavior as a socially relevant cue comes from the work of Farroni and colleagues (Farroni, Csibra, Simion, & Johnson, 2002; Farroni, Menon, & Johnson, 2006; Farroni, Pividori, Simion, Massaccesi, & Johnson, 2004). This research shows that newborns have a preferential attention toward faces with direct gaze toward the observer, even when these faces are only schematic ones. Interestingly, even newborn chimpanzees have a preference for human faces with gaze directed toward them

22

(Myowa-Yamakoshi, Tomonaga, Tanaka, & Matsuzawa, 2003). Thus, all of these findings demonstrate that there is an important relation between being able to detect where somebody else's eyes are directed and becoming social beings.

Focusing on the interaction between eyes and development, one of the most notable features of a human infant face is the size of her/his eyes: they are particularly large compared to the rest of their body. At birth, a human eye has a diameter that is approximately 70% of the size of an adult eye: 17 mm at birth vs. 24 mm as an adult (Gordon & Donzis, 1985). In this respect, the eye is probably the human organ that grows the least during development. Two kinds of vertebrate tend to have large eyes compared to their body size: nocturnal predators and herbivore mammals such as cows and rabbits (see, e.g., Howland, Merola, & Basarab, 2004 on vertebrate eye sizes). In the first case, big eyes facilitate the detection of prey at a distance; in the latter they help the detection of predators. Clearly, a newborn human is anything but a great predator and has little capacity to escape if attacked. Moreover, if one of the advantages of having big eyes is that they improve visual acuity (and therefore facilitate the detection of prey at a distance), this cannot apply to human newborns because they can only see at a distance of about 8-15 inches in the first few weeks. This distance is important because it is approximately the distance between the faces of a baby and its caregiver when the latter holds the baby in her/his arms. So, we are not born with highly developed eyes that can easily scan the environment at a distance and detect things, but newborn eyes are particularly sensitive to human faces, and to black and white forms, much more than to colors (see, e.g., Fantz, 1963). A possible hypothesis concerning the evolutionary advantage of infant eye size and perceptual capacities at birth is that the range of visual acuity and large white sclera can facilitate mother-infant communication. Thus, they may have evolved this way for social/communicative purposes (see Ross & Kirk, 2007 for a review on eye size in primates).

### 1.2.3 Eyes and Vision

Our retina, the light sensitive layer at the back of our eye, is made of two basic photosensitive cells: rods and cones. Rods are photoreceptors that work particularly well in conditions of low light and allow scotopic vision; that is, the monochromatic vision of the eye in conditions of low luminance (i.e., when it is dark). Given that cone cells cannot function in conditions of low luminance, only rod cells are active in those conditions and this means that in situations of darkness there is no color perception. Rods are particularly important for night vision, motion detection and peripheral vision.

On the other hand, cones are the cells responsible for color vision,[8] the ones that provide us with the highest visual acuity. There are no rods in the *foveal pit,* the center of the fovea, the part of the retina responsible for sharp vision, while cones are mostly concentrated in the fovea. This distribution of these two photoreceptors allows us to see over a wider range. The area rich in cones (the fovea) makes us see with high visual acuity during bright light, while the area rich in rods (i.e., the rest of the retina) provides high sensitivity vision in dim light. Such a distribution allows for a good balance between visual acuity and visual sensitivity. The fovea represents less than 1% of the retina and we can see sharply only through the fovea. This explains why we can see sharply only about 15-30 degrees of arc, while we have a basic perception of about 200 degrees (Bruce et al., 2003). This biological configuration also explains why we have to move our eyes to scan the environment and perceive the shapes of objects in a room with a decent level of visual acuity. It also explains why we move our eyes to look at a cup before we reach for it: excluding for a moment the role of the lens in visual acuity, if we had cones more evenly distributed throughout the entire retina, we would be able to perceive the position of the cup and its shape with high definition[9] while still looking at the face of our interlocutor. However, given that our retinas do not have this configuration, our eyes have to be redirected toward the object and moved away from the other participant to properly control our hand movement.

---

[8] Most mammals are dichromats. Primates are the only mammals having trichromatic vision (Jacobs, 1993, 1995). Visual acuity of humans is not really remarkable, in comparison to other species. Indeed, most birds that hunt can see better than us (e.g. the eagle) and many diurnal birds can see ultraviolet light.

[9] Note that visual acuity is not only dependent on the presence of cones in the retina, but other factors affect it as well, among which how we can modify the thickness of our lenses to bring nearby objects into focus, through a process called accommodation (see, e.g., Bruce et al., 2003).

In the second half of the 20th century developmental and psycholinguistic researchers exploited the correlation between eye movements and foveal vision with cognition and attention. This happened in particular as a development of the initial investigations in the 1950s and 1960s made by Robert Fantz with infants and by Alfred L. Yarbus on scene perception and saccadic eye movements in adults. Fixation duration, direction of first look and frequency of switching between co-present objects have been interpreted as reflections of underlying cognitive processing. A detailed investigation of the history and findings of looking-time research in developmental psychology and of eye movement research in reading, during scene perception and language production (see, e.g., the success of the Visual World Paradigm in psycholinguistics) is beyond the scope of this dissertation. Some comprehensive reviews and methodological discussions of the above approaches can be found, however, in Aslin (2007); Griffin (2004); Rayner (1998), Spelke (1985) and Tanenhaus et al. (2000) and an extensive overview of recent works on eyes and perception can be found in Liversedge et al. (2011).

Given the focus here on the use of eye-gaze during social interaction, no detailed review of the visual perception of humans will be provided here. However, it is worth reporting how different models of visual perception have strongly influenced our understanding of the function of our eyes and how they interact with the environment that surrounds us. A behaviorist approach to perception, such as the one proposed by Watson (1913, 1924), attempted to reduce perception to basic sensations excluding introspection. This was countered by the Gestalt psychology approach, which pushed for a more phenomenological and nativist approach to vision, with a strong focus on introspection. In recent years, Gregory (1972, 1980) proposed a top-down approach to vision, where hypotheses about the world are constantly tested against fragmented evidence based on the images we perceive. This was countered by Marr (1982) with an alternative computational bottom-up approach, emerging from artificial intelligence. Marr's approach is fundamentally unconscious and computational in scope, which relies on inferences about the images perceived. It also provides humans with a certain level of awareness of the visual world by processing the information obtained via the retina at different levels by using statistical regularities. Finally, an influential approach to vision that developed in the last part of the 20th century is the ecological one, which owes its initial framework to the work of J.J. Gibson (e.g., 1979). The core intuition of this approach is that action and perception are intimately intertwined and cannot be treated as separate problems. The point of departure should not be the stimulation of the light on the retinal image, but rather the optic array provided by the

environment. The way in which light extends over space and time provides information to the observer about shapes and guides the movements of animals through space. We do not just try to perceive shapes to memorize pictures in our brain; we use vision to make sense of what we can do in and with an environment. In particular, when we perceive an object, we do not just see what it looks like, but we also see its *affordances*, the actions that we can perform with it, its possible uses. Contrary to conventional theories of vision, which postulate a view of perception that is mediated by inferences about the world that develop by transforming mental representations, Gibson suggests that perception is not mediated by any psychological process such as memory or inferences but is rather an immediate picking up of invariants in the optic array.

Recently, however, Milner and Goodale (1995) have put forward a different hypothesis about vision, one that differentiates a visual path for action from a visual path for perception. Basing their claim initially on findings from research on individuals with brain damage, they suggest that there are two different pathways in the brain, as far as our visual system is concerned: one dedicated to perception (the ventral stream) and one dedicated to action (the dorsal stream). The idea is that visual input affects our actions in different ways. The dorsal pathway controls fast movements and allows us to grasp, for example, the handle of a cup. On the other hand, through the ventral pathway, the visual input interacts with memory and allows us to construct a conscious and more accessible representation of the world, allowing us to plan for more complex courses of action. Many recent experiments conducted either on individuals with brain damage or using visual illusions seem to confirm the existence of these two different pathways (see Milner & Goodale, 2008 for a review). This model combines the insights of Marr's computational model and Gibson's ecological approach, by suggesting that both types of processing of visual information are actually occurring in our brain at the same time.

## 1.3 To Be Seen

A recent study found that by simply placing a pair of printed eyes on an honesty box used to collect money for drinks in a university coffee room, people paid three times as much for their drinks than they did when a control image was used (Bateson, Nettle & Roberts, 2006). Just seeing a pair of eyes had apparently caused people to behave quite differently, in particular being more cooperative and generous, than in the control condition. One of the accounts for this finding is that at an unconscious level, printed open eyes facing the participants induced the perception of being watched (and therefore evoked an association to their accountability to others). Another experiment by von Grünau and Anston (1995) showed that people were much faster and more precise in detecting the presence of a straight gaze directed toward them among a large number of averted gaze, rather than the other way around. This special alertness to a straight gaze stimuli has a clear adaptive explanation: staring can be a sign of hostility or anger, but eye contact can also be a sign of liking or friendliness, and as the authors explain "whether maintained stare is a sign of dislike or like, it is certainly an indication of a potential social interaction." (p. 1297)

We know that humans have neurons that respond very selectively to stimuli that look like faces and eyes (e.g., Emery, 2000; Haxby, Hoffman, & Gobbini, 2000), and we have already mentioned how human infants are sensitive to direct gaze toward them even when schematic faces are used, rather than photographs of real faces (Farroni et al., 2004). One reason why the impression of being watched might affect human behaviors could be that we are socially concerned with maintaining face (Goffman, 1955, 1963) and naturally concerned with maintaining a positive reputation (see, e.g., Alexander, 1987; Milinski, Semmann, & Krambeck, 2002). However, another option is possible as well: we are more cooperative when we think we are being watched because we are concerned about punishment. From a sociological point of view, for example, we know from the work of Foucault in *Discipline and Punish* (1979) that the control of a prisoner and his behavior can be institutionally exerted just by making the prisoner believe that he is being observed. Foucault cites in particular the plan for a perfect prison, originally suggested by Bentham, that he calls *Panopticon*, in which a single tower in the middle of a prison courtyard with a 360 degrees view of every cell's window, would immediately make prisoners be more cooperative, as they would never know when they are actually observed by a guard. So the fear of being

perceived breaking the rules and as such risking being punished would make them cooperate more.

Many writers have provided admirable descriptions of how characters in their works experience the look of somebody else toward them (see, e.g., Boll, 1994 [1963]; Sartre, 1956 [1943]).[10] Yet one aspect often neglected is how these glances are made possible during a conversation. The answer lies in the ways in which we position our bodies when we talk to each other. In the 50s and 60s Sommer developed a series of studies concerning how participants position themselves in space depending on the kind of interaction they want to have (Sommer, 1959, 1962, 1967), observations later confirmed and developed, among others, by Scheflen (1964, 1975a), Ekman and Friesen (1974) and Kendon (1977). The main aspect of Sommer's work that matters for us is the idea that participants would position themselves so that they could monitor one another and make eye-contact comfortable, if they were planning to have a conversation. This idea of positioning our bodies so that mutual monitoring becomes easier has been further developed by Goffman (1963, 1964, 1981b) and Marjorie Goodwin (e.g., 1980), who claimed it to be a crucial component of a social situation and showed us how it plays out in real time. Kendon (1977) even coined a specific new term to describe the configuration which participants assume when they have a social interaction: F-formation.

> *An F-formation arises whenever two or more people sustain a spatial and orientational relationship in which the space between them is one to which they have equal, direct, and exclusive access. Such a pattern can be seen in the circle of the free-standing conversational group. [...] The F-formation system serves as an important means of maintaining the separate identity and integrity of an interactional situation. It provides a means by which the participants can maintain differential access to one another and it facilitates the maintenance of a common focus of attention. It is thus an important part of the means by which behaviour is organized in occasions of face-to-face interaction such as conversations. (p. 209)*

The effect of being watched on people's behavior suggests that eye-gaze has a strong regulatory power, which enables participants to affect the way others behave in their presence. We'll see later how this has been investigated with respect to social interaction and conversation.

---

[10] See Tomkins (1963) for a historical review concerning how writers have written about the use of eyes in relationships.

We have already mentioned that two of the most typical situations, in the absence of talk, in which two individuals would find each other sustaining mutual eye contact are situations of threatening and flirting. Even though they seem to represent the two ends of an emotional spectrum, one having to do with menace and aggression and the other one with intimacy and desire for closeness, they have something in common: they are both actions that provoke high arousal (see Eibl-Eiblesfeldt, 1989 for a detailed account of how these arousing situations are managed and how universal and possibly innate they appear to be). Again, if we move from the behavioral level to its biological motivation, we can note how the very perception of two open eyes looking at us, when perceived, provokes a systematic increase in galvanic skin response (GSR) (see e.g., Nichols & Champness, 1971; Hietanen et al. 2008), which is to say we get aroused. Yet, the fact that people look at each other's face during a conversation regularly would suggest that we have found a way of canalizing these arousing emotions into a communicative practice, as Lorenz (1966) suggested. It should be noted here that other primates, while monitoring each other's behavior in a manner similar to the one used by human beings, tend to engage in eye contact mainly in agonistic situations, such as threatening, or to solicit sexual favors (see, e.g., Goodenough, McGuire, & Wallace, 1993; Hinde & Rowell, 1962; Kaplan & Rogers, 2002; Redican, 1975). There are, however, reports of the engagement of mutual gaze to initiate reconciliations (de Waal, 2001), and in non-aggressive and non-sexual social interactions (Yamagiwa, 1992), mainly deployed by juveniles. Mutual gaze, even in chimpanzees, occurs quite regularly in mother-infant interactions (Bard et al., 2005; Ehardt & Blount, 1984), while it rarely occurs during adulthood and there appear to be differences in the amount of eye contact that usually occurs not just across species (see, e.g., Kaplan & Rogers, 2002), but also across groups within the same species (Bard et al., 2005). Nonetheless, the length of mutual gaze that occurs during a natural conversation between two human beings clearly far exceeds what is usually observed in other non-human primates. It therefore appears that some special evolutionary adaptation occurred in the human line that allows us to sustain mutual gaze in friendly and yet not arousing situations, even beyond childhood. If we combine this with what was previously reported concerning the different color of human sclera (white vs. dark one for other primates), it should be clear that we are biologically evolved and adapted to rely on gaze direction and mutual gaze during social interaction and communication.

However, it is important to note here that when we talk to each other we tend to fixate not each other's eyes,[11] but rather the upper part of the nose, which sort of represents the center of our face, to be able to fully perceive facial expressions (see, e.g., Gullberg & Holmqvist, 2006). Nonetheless, the very possibility of looking at each other and the tendency to often engage in mutual gaze, at least in Western cultures, have been noted by Simmel as the "uniquely sociological function" of the eye and the "union and interaction of individuals [being] based upon mutual glances" (Simmel, 1969: 358). In particular he claimed that "the totality of social relations of human beings, their self-assertions and self-abnegation, their intimacies and estrangements, would be changed in unpredictable ways if there occurred no glance of eye to eye" (1969: 358). Seeing others and being seen, engaging in mutual gaze, appear to have a special significance in human interactions, which might go beyond the mere perceptual or communicative functions of the eyes.

The old saying that the eyes are the mirror of the soul, modified by Cicero who claimed that "the face is a picture of the mind as the eyes are its interpreter", reflects a belief about the relationship between eyes and emotions that has been investigated intensively in the last two centuries. In his work *On the expression of emotions in man and animals* (1872), for example, Darwin claimed that feelings such as pride, humility, guilt, conceit, slyness, suspicion and others could be detected not just by the facial expression of an individual but more simply by their eyes. Most recent research on facial expressions, in particular the ones by Ekman and colleagues (Ekman, 1992, 1993; Ekman & Friesen, 1978; Ekman & Oster, 1979) but also recent conversation analytic work (Ruusuvuori & Peräkylä, 2009), has looked at the role that eyes are said to play in the compositionality of facial expressions and has considered not only the movements of the eyes but also all the muscle movements going on around them, including the brows, to accomplish, for example, frowns. Despite the importance of the eyes in displays of emotions and in the composition of facial expressions, this dissertation will not focus on these aspects, but rather on the orientation of gaze during social interaction and its communicative and regulatory functions. In what follows I will provide some reasons for this focus.

---

[11] When we actually look at each other's eyes and not just each other's face we usually fixate one of the two eyes, as fixation must converge on a single point of focus.

### 1.3.1 Eyes, Faces and Brain

Recent research in neuroscience has focused on what has been called the social brain (e.g., Adolphs, 2009; Brothers, 1990; Dunbar, 1998; Grossmann & Johnson, 2007; Johnson et al., 2005), which is to say the neural networks involved in human social interaction and communication. Among the areas identified as part of the social brain, there are the prefrontal cortex, the superior temporal gyrus, the fusiform gyrus, the cingulate gyrus, the superior temporal sulcus (STS) and the amygdala. These areas are specialized in processing information such as face, gaze, biological motion, human action, goal directedness, theory of mind and empathy.

Before focusing on the brain regions involved in the processing of eye gaze, consider two recent findings concerning the differential processing of face and eyes. Haxby et al. (2000) found that while changeable, dynamic aspects of faces such as expressions and gaze tend to activate the STS (area involved in the detection of biological motion, Allison, Puce, & McCarthy, 2000), the static aspects of faces that enable recognition and identification activate the fusiform gyrus. More recently, Calder et al. (2007) provided an even more fine-grained finding concerning the compositional nature of face processing and the separation of gaze direction detection and facial identity: left and right gaze direction of the other are coded in the right anterior STS and in the inferior parietal lobule (involved in the orienting of attention), while facial identity is coded in the posterior fusiform gyrus. Moreover, we know that while the anterior STS is involved in the coding of different gaze directions, the posterior STS is actually sensitive to the intentionality of gaze and other biological signals (Pelphrey, Morris, & McCarthy, 2004), suggesting that a further distinction can be made between processing the movement of somebody's eyes and detecting whether they are meant to communicate something or whether they are doing more than simply scanning the environment.

Besides the right regions of the STS, there are other brain areas involved in the processing of gaze (see Grossmann & Farroni, 2009; Senju & Johnson, 2009 for more detailed overviews): the amygdala (activated in the processing of emotions or by arousing stimuli), the medial prefrontal cortex (involved in mentalizing[12] and detecting and decoding communicative intentions toward self) and the orbitofrontal cortex (involved in decision

---

[12] With the term mentalizing I am here referring to the capacity to attribute intentions to self and others, a mechanism considered to be an automatic cognitive process (Leslie, 1987).

making). Moreover, a study by Kampe et al. (2003) has shown that hearing somebody calling our name and seeing somebody's eyes directed toward us for a prolonged amount of time, while differing in modality and sensory channel, tend to activate similar brain areas, usually involved in mentalizing, namely the paracingulate cortex and the temporal lobes. The meaning of this finding is that both calling somebody's name and directing the eyes toward that person, while acting like a summoning, signal the intention of initiating a communication by the other participant and our brain registers it. This explains the common experience of being in a public place, registering that somebody is looking at us and by the time we look up toward them, the person who was staring at us swiftly lowers the gaze or looks away. When this does not happen, the alternative is usually the initiation of a conversational exchange. Prolonged gaze toward our face works fundamentally like a verbal summoning, at least as far as our brain is concerned.

All the findings reported above show that when we observe somebody's eyes (directed toward us or even averted) a series of reactions occur in our brain, that might have to do with arousal and emotional responses if the context suggests so, but more generally that involve detecting communicative intentions directed toward us (i.e. somebody wants to communicate with us) and deciding what to do next, how to respond.

The fact that gaze directed toward us indicates not only that we are the focus of somebody's attention but also an intention to communicate with us is further confirmed by recent experimental work with infants developed by Csibra and Gergely and their model of natural pedagogy (see Csibra & Gergely, 2009 for a recent summary of their model). In their model of infant social learning, gaze toward infants is one of the ostensive cues that infants use to recognize that they are the ones being addressed, to expect the occurrence of some referential act in an ostensive context and to interpret the content of the communication as kind-relevant and generalizable. We will see in the following chapters how these neurological and psychological findings contribute to our understanding of the mechanisms at play during social interaction, in particular with respect to the use and perception of eye-gaze.

## 1.4   Gaze and Social Interaction

Most of the detailed work on gaze in interaction has been conducted by social psychologists or kinesic researchers working on interaction in English or other European languages. The main focus of social psychologists has been to determine how specific factors would correlate with the amount of looking toward each other's faces during a dyadic interaction, using such variables as personality, gender, age, social status, asymmetric relationships, seating arrangements and similar (see Kleinke, 1986 for a review of research in this area). Researchers such as Ralph Exline, Michael Argyle and Phoebe Ellsworth pioneered research in this area in the 60s and 70s. Alongside the social psychological approach, there was a move from linguists and anthropologists, together with those interested in cybernetics and psychotherapy, toward including nonverbal features of the interaction in their analysis and descriptions. I have already mentioned Edward Sapir, but there were many others as well including Elliot Chapple, Edward T. Hall, Gregory Bateson, Albert Scheflen and Ray Birdwhistell. The work of Birdwhistell, founder of the kinesic approach, has been particularly inspiring for many generations of researchers (e.g., Adam Kendon among others). Birdwhistell (see, e.g., 1970) tried to provide the field with a method and theoretical tools to produce an accurate and detailed description of body movements during an interaction, assuming their compositionality, communicative import and cultural variability (see, e.g., Leach, 1972). He assumed that there was nothing in peoples' behavior that should be disregarded *a priori* as meaningless. This assumption has later been pushed even further by Conversation Analysts as summarized in Harvey Sacks' phrase "order at all points" (Sacks, 1992 [1964-72]: 483). On the other hand, the obsession with wanting to provide a fully accurate description of any body movement, the focus on single case analysis and the attempt to divide body movement in pseudo-linguistic elements, similar to morphemes, led to the development of extremely long descriptions of body behavior that could not be used in any systematic comparison or corpus study. The fact that Birdwhistell never fully completed a description of any stretch of interaction he investigated confirmed the problem of having a too fine-grained tool to analyze something that might be interpreted and understood using different, broader, categories. There are, however, a series of insights from that approach, as reported by Albert Scheflen (1975b), a close collaborator of Birdwhistell:

1)      The dichotomy according to which language is communicative, while every other visible behavior simply works as a cue for who is speaking or what is supposed to happen next is wrong;

2)      A participant does not speak, gesture, smile, hold a posture simultaneously to form a single message with redundant parts. Each modality is employed for specific purposes, some of which could be purely communicational, others might be regulatory and others again might be used to induce or sustain specific relationships between the participants in the interaction.

The program was pushed for an understanding of visible behavior as a core part of the communicative process and not simply as a mainly redundant side thing, just used to cue who has to start or who has to stop doing something. Their main insights are as follows:

1. Regulatory functions are not the only functions of visible behavior.
2. Visible behavior is organized sequentially (an insight outlined by conversation analysts, e.g., Sacks, 1992 [1964-72]; Schegloff, 1968).
3. What is managed during a conversation is not just the exchange of information and not just direct communication, but a more complex set of things, including relationships and projects that might not be part of the content of the conversational exchange.

All these concerns have informed the work contained in this dissertation and will be emphasized in the analysis of specific stretches of interaction. These insights have permeated most of the work on gaze in social interaction, in particular the work of Adam Kendon and Charles Goodwin, whose studies will be described in more detail below.

From a conversation analytic point of view, research on gaze in social interaction has focused mainly on three different dimensions: first, its relationship to participation in the conversation; second, its regulatory functions and its role in turn-taking; and third, its role in action formation. These areas, however, have received different amounts of attention, as will be clear from the following report of the main findings in each one of them.

*1.4.1 Participation Roles and Participation Frameworks, Engagement and Disengagement*

A great deal of work has been dedicated to the relationship between gaze, participant roles in conversation and the participation framework in place. The idea that gaze is closely related to participant role (speaking, or being addressed, in particular) is persistent across the literature. In particular, it has long been claimed that in dyadic interaction, people tend to look at the other participant more when they are listening than when they are speaking (Argyle & Cook, 1976; Argyle & Dean, 1965; Argyle & Graham, 1976; Bavelas, Coates, & Johnson, 2002; Duncan & Fiske, 1977; Exline, 1963; Goodwin, 1981; Kendon, 1967, 1973, 1990; Kleinke, 1986; Nielsen, 1962; Rutter, 1984). The first person to note this difference was Gerhard Nielsen in his seminal work on self-confrontation (1962). Interested in the timing of visual behavior in face-to-face interaction, he found that not only is there a "tendency to look more at the alter while listening to him than during one's own speaking periods" but that most of the dyads "show a striking regularity: looking at alter decreased when speaking increased and looking at alter increased while listening increased" (Nielsen, 1962: 141). Kendon (1967) provided a more precise description of the different patterns of speaker and hearer gaze. He claimed that hearers give speakers long looks interrupted by brief glances away, while speakers alternate looks toward and looks away from the recipient of approximately equal length.

These studies implicitly suggest that such gaze behaviors are independent of attributes such as race, culture and gender, for instance. There are, however, studies that claim racial differences with respect to the use of gaze to display engagement. Specifically, Black Americans have been shown to look at the recipient more while speaking than while listening, while White Americans follow the opposite pattern (Erickson, 1979; LaFrance, 1974; LaFrance & Mayo, 1976). Moreover, recent work by Rossano et al. (2009) shows that the amount and type of gaze interaction between speaker and recipient may vary across cultures and may also be strongly related to the social actions the participants are initiating through their talk. Indeed, they find that in the context of question-answer sequences in the three cultures investigated (speakers of Italian, speakers of Tzeltal, a Mayan population from Mexico, and speakers of Yélî Dnye, from Rossel Island, a remote island of Papua New Guinea), questioners are more likely to look at their recipients than vice versa, and they tend to do so by looking toward their recipient from the very beginning of the question without alternating looks toward and looks away.

Returning to general claims about gaze patterns in face-to-face interaction, Goodwin proposed two rules (1980: 275, 287; 1981: 57) that would account for gaze behavior in conversation, according to a case by case analysis of his data, and further supported by quantification in a small corpus:

1) "A speaker should obtain the gaze of [her] recipient during the course of a turn at talk."

2) "A recipient should be gazing at the speaker when the speaker is gazing at the hearer."

If the recipient looks most of the time, then the speaker will find her/him gazing back any time the speaker looks toward the recipient. If the recipient is not looking at the speaker, the speaker has resources (phrasal breaks, pauses, restarting the turn) to solicit the recipient's gaze. By proposing these as "rules", Goodwin claims that participants' gaze behavior is interrelated rather than independent and suggests a normative order to which participants are oriented during any turn-at-talk. In his work on gaze behavior during the course of a story telling, Goodwin (1984) emphasizes how participants' visible—and, in particular, gaze— behavior helps shape the unfolding of the telling. Moreover, he proposes that a way to understand why gazing away from the speaker during a telling is rarely sanctioned or treated as problematic is that there is a relaxation of the gaze rule due to the co-occurrence of competing activities. However, the very suggestion of a rule that can be 'relaxed' immediately calls into question the normative strength of the rule suggested and whether it would not be more appropriate to come up with a rule that would account for this behavior, rather than requiring its relaxation. This dissertation will show that gaze is mainly organized at a different level than the one proposed by Goodwin.

In addition to his seminal paper on how gaze deployment affects the construction of an utterance in progress (Goodwin, 1979), the most important work on gaze in social interaction from a conversation analytic perspective is Goodwin's (1981) *Conversational Organization: interaction between speakers and hearers*. In this book, he describes gaze behavior as a display of attention and (dis)engagement in the conversation, or more precisely, a display of the type of participation framework the participants are engaged in. From this perspective, looking away is noticeable and potentially sanctionable because it displays diminished engagement in the conversation. On the other hand, engagement in some

competing activity (e.g. eating) provides a ready account for looking away rather than at the interlocutor's face and thus makes it less sanctionable. However, this claim also implicitly suggests that participants are relatively free to remove their gaze from co-participants, provided that they direct it toward another activity in which they are engaged. This appears to be problematic from a closer look at a larger number of interactions. Indeed, as it will be later shown (chapter 2), some activities require more sustained gaze by the recipient toward the speaker (e.g., tellings) or by the speaker toward the recipient (e.g., questions) than others, suggesting that this relative freedom actually depends on the gaze expectations associated with the ongoing course of action. Moreover, participants appear to calibrate their looking toward competing activities so that they do not disrupt the progressivity of the talk; that is, for example, they might look toward a glass and pick it up to drink when they expect they won't have to produce the next turn at talk (see chapter 4), so that drinking will not create a silence when talk would be otherwise expected.

Further work on the relationship between gaze and participation has been developed by Heath (1984), who showed how a gaze and postural shift directed toward a co-participant can be used to "display recipiency". As such, it "is sequentially implicative for an action by a co-participant; it breaks the environment of continuous opportunity, and declares an interest in having some particular action occur in immediate juxtaposition with the display" (1984: 253). In this sense, a body movement can elicit speech by the other participant or it can elicit a gaze re-orientation. In this way, the person who produced the body movement in first position can obtain recipient gaze and begin to speak, making the beginning or the continuation of a conversation possible. Focusing on gaze and body posture as a way of displaying (dis)engagement in a course of action has been used in a fruitful way in the analysis of doctor-patient interactions (see, e.g., Heath, 1986; Robinson, 1998), and in broadcast interactions (Ekström, 2011).

The work reported above shows that participant roles (i.e. speaker vs. recipient) affect participant gaze behavior. It also indicates that there may be normative expectations of gaze by recipients toward speakers, at least when they are not involved in competing activities and are listening to an extended telling. Looking toward or looking away from the other participant is often a good clue in terms of participants' (dis)engagement in the conversation. However, as this dissertation will show, research on gaze and participation frameworks may have not adequately addressed differences related to social action.

## 1.4.2 Regulatory Functions of Gaze

Much research on gaze in social interaction has focused on its regulatory functions. In what follows I outline what we know about gaze in turn taking, turn allocation, mobilizing response and sequence organization.

Early studies of the regulating functions of gaze in turn taking argued that gaze works to show that speaker A is finished talking and thus is used to signal the handing over of the floor. Some researchers have focused on the importance of gaze in monitoring each other's behavior and facial expressions (e.g., M. H. Goodwin, 1980; Kendon, 1967). Kendon (1967) and Duncan and colleagues (Duncan, 1975; Duncan & Fiske, 1977; Duncan & Niederehe, 1974) have argued that speaker gaze has a 'floor apportionment' function in conversation and can function as a turn-yielding cue.[13] In particular, they claim that speakers tend to gaze away at the beginning of turns and tend to look up toward the recipient when approaching turn completion in order to signal that they are ready to turn the floor over to the other participant. In his seminal work, Kendon (1967) not only claims that speaker gaze displays the relevance of response but also affects its timing. However, Kendon deals only with the general dynamics of turn taking. He makes no distinction between sentence types (such as interrogatives or imperatives) nor between action types performed through those turns at talk (e.g. announcements, challenges, complaints), but rather writes only in terms of "long" or "short" utterances.

Beyond these early studies, subsequent work found no evidence that speakers use gaze as a turn-yielding cue. Beattie (1978, 1979) suggests that a speaker's looking away during early utterance production, and reengagement during final production, are occasioned purely by the need to reduce cognitive load and that they do not have any regulating function in terms of turn taking.[14] In trying to verify previous claims about the occurrence of looking toward a recipient approaching the end of the turn, Torres, et al. (1997) found that, of all turn endings in their data, only 16% included a look toward the recipient by the speaker and these 'look-toward' only accounted for 15% of all the speaker 'look-toward'. Most recently, De Ruiter (2005), examining task-based dialogues, confirmed the lack of a systematic

---

[13] Duncan and colleagues (Duncan, 1975; Duncan & Fiske, 1977; Duncan & Niederehe, 1974) actually refer to a shift in head direction as a turn-yielding cue, but specify that this should be taken as a proxy for 'eye direction' (1977: 211).

[14] In a direct response to Beattie's paper, Kendon (1978) argues that Beattie's data (conversations between a student and her/his supervisor) were not comparable to his (ordinary conversation between Oxford undergraduates), indicating that the kind of interactional situation participants are dealing with may well affect the deployment of gaze.

relationship between gaze and turn-taking in general, thereby reminding us of the need for a better description of gaze functions in face-to-face interaction.

Another related line of research has focused on turn allocation or speaker selection in the context of multi-party conversations, rather than on turn transition. First, Goodwin (1979: 99) claimed that, while uttering a sentence, "the gaze of a speaker should locate the party being gazed at as an addressee of his utterance," and showed that the very construction of a sentence can be affected by whether recipients return the gaze of the speaker during the uttering of the sentence, thereby allowing for the establishment of eye contact between the participants. More recently, developing Sacks, Schegloff and Jefferson's (1974) work on turn allocation, Lerner (2003) compared gaze and address terms as ways of addressing a participant and showed that a speaker's look toward a participant is an explicit form of addressing a participant, but its success is contingent on the gaze practices enacted by the other participants, in particular on whether both the addressed individual and the non-addressed ones perceive the speaker's gaze.

A third, larger stream of research regarding the regulatory functions of gaze has adopted a somewhat different tack in suggesting that gaze works to solicit response. In their work on the social organization of word searches, Goodwin & Goodwin (1986) refer to the occurrence of speaker gaze toward a recipient in two examples of word searches, and they describe its function as a way of soliciting aid from the other participant. However, no systematic evidence for this claim is presented and, in both examples, this apparent solicitation through gaze is, in fact, unsuccessful.

More recently, in an experimental setting, Bavelas, et al. (2002: 576-577) find that "the listener tended to respond when the speaker looked at her, and the speaker tended to look away soon after the listener responded. Together, speakers and listeners created and used the gaze window [mutual gaze] to coordinate their actions". Their notion of the *gaze window* describes a mutual gaze situation in which it is not *just* that the speaker's gaze elicits a response, but rather the listener's response seems to terminate the speaker's gaze. By 'listener responses' they refer mainly to *mm hm*, *uh huh* and nods, which have been labeled *back channels* in the literature (Yngve, 1970) or, within CA, *continuers* (Goodwin, 1986b; Schegloff, 1982). Bavelas et al. assert that one of the main features that distinguishes the gaze window from the pattern previously described in turn exchange (i.e. looking toward the recipient indicates that the current speaker is approaching completion of his/her turn and is ready to leave the floor to the other participant) is that there is no exchange of roles between speaker and listener.

One of the main conclusions that Bavelas, et al. (2002) draw is that the speaker does not look at the listener to monitor him/her for action but rather to solicit a response. A second conclusion is that listeners display their recognition and understanding of the action performed through gaze by responding "immediately and appropriately". At the same time, it is unclear in which way the occurrence of a response before the withholding of gaze becomes evidence that listeners are responding specifically because of the gaze. Indeed, because of the coding system adopted for this study, a participant could have been looking for 10 seconds and if s/he looks away after a response, then the response would be considered related to the sustaining of the gaze, while the actual talk is not taken into account. A further problem associated with this study is that other communicative behaviors to solicit a response were not taken into account, and, as such, the specific actions performed through talk ignored. This suggests that the claim of a relationship between speaker gaze and listener's response needs further specification, which has been provided, at least to some extent, by other researchers in more recent studies, to which I now turn.

In a paper focusing on the resources that participants deploy to mobilize a response by a recipient, Stivers and Rossano (2010) identify speaker gaze as one such resource. They show that in face-to-face interactions, when the speaker is gazing at the recipient during an assessment or other non-canonical actions, the recipient usually responds to the initial assessment, while assessments that are not responded to are usually produced without speaker gaze or other response mobilizing features. The claim is that speaker gaze plays a role in mobilizing recipient response. Additional evidence for this claim comes from my work on question-answer sequences in Italian dyadic conversations (Rossano, 2010). I focus on the resources that speakers deploy to indicate to recipients that they are producing a question (and, therefore, that they are pressuring recipients for some response). I find that when speakers produce polar (yes/no) questions that do not have a prototypical interrogative intonational contour (i.e. a rising intonation), they are significantly more likely to look toward their recipient than when they are producing questions with prototypical interrogative contours. Given that polar questions in Italian can be marked intonationally but not morpho-syntactically, the occurrence of speaker gaze toward the recipient appears to work as an additional cue that the turn should be responded to. Additionally, in work focusing on the timing of responses to yes-no questions in multiple languages, Stivers, et al. (2009) show that the occurrence of speaker gaze correlates with faster responses to participant questions. Although they do not take into account recipient gaze, and the number of participants involved in the interactions varies, the finding is nevertheless supported in nine out of the ten

typologically diverse languages investigated. Another recent cross-linguistic study that takes the sequential environment of talk and gaze into account provides additional evidence for the existence of a relationship between the occurrence of speaker gaze and the expectation and timing of a response. In their work comparing gaze behavior during questions in three unrelated languages and cultures mentioned earlier, Rossano, et al. (2009) show that, on average, speakers look at recipients during questions in 73% of cases. They also show that speaker gaze behavior during questions is similar across the three cultures, while recipient gaze behavior differs. Moreover, they find that questions are usually responded to and the ones that do not elicit response are typically produced with the recipient not looking at the speaker. Chapter 3 of this dissertation investigates further how gaze can be used to mobilize and pursue a response when the latter is missing.

Chapter 4 instead will investigate how participants orient to the relevance of gaze withdrawal at sequence completion, particularly when this coincides with the completion of a course of action (for other means to project possible completion of a sequence see, e.g., Mondada, 2006). In prior work, gaze withdrawal was described as a function of reduced engagement with a conversation (e.g. Goodwin, 1981, 1984), yet the suggestion here is that the level at which gaze works is more fine-grained than previously described. In their work on assessments, for example, Goodwin and Goodwin (1987) argue that gaze withdrawal is a way of displaying diminished participation in the activity, and discuss assessments as a resource for closing topics and mutual orientation. Interestingly, though, in all examples shown in that paper, gaze withdrawal is not followed by immediate topic closure but rather by further talk (at least an additional TCU) either by the participant who withdraws gaze (if s/he was the speaker) or by the current speaker (if the person withdrawing gaze is the recipient). As will be shown in chapter 4, gaze withdrawal is indeed a resource for making a bid for closure, or for displaying a specific understanding of the ongoing development of the course of action. By bidding for closure, participants display diminished participation in the activity.

In this section we have seen the claims about the regulatory functions of gaze being organized with respect to turn-taking, yet this dissertation will challenge those claims by suggesting that gaze behavior is organized with respect to the sequential organization of courses of action. Moreover, claims about the role of gaze in soliciting a response will be partly confirmed and partly refined, suggesting that this is the case in specific sequential environments, and not just at any point in a conversation.

## 1.4.3  Gaze in Action Formation and as a Social Act

A third line of research, largely developed within the last few years, and for the most part undertaken by Kidwell, has addressed the role of gaze in implementing social actions, depending on its context and delivery. Kidwell's work focuses, for example, on the recognizability of participants' looking practices in interaction, particularly in a childcare setting and in interactions with young infants. Kidwell (2005) investigates how very young children can differentiate two practices of looking by caregivers, termed a *mere look* and *the look*. While a mere look tends to be of short duration, alights briefly on a target and is usually produced concurrently with other activities that the caregiver is involved in, the look is relatively longer, is fixated on a target and is produced as a discrete activity. Kidwell finds that young children treat these two practices of looking differently: whereas they continue to engage in whatever they were doing when a caregiver directs a mere look at them, those same children treat the look as prefiguring sanctioning. As such, after a caregiver's 'look', children tend to inspect their own actions for its source (i.e. sanctionable behavior or activity such as harassing other children), and this often leads to a disruption of that behavior.

In a related study, Kidwell (2009) shows how, in the context of children harassing other children, a gaze shift by the harassed child may be interpreted alternatively as "looking to" the caregiver as an appeal for her assistance, as "noticing" that the caregiver is approaching or as "searching" her out. The looking to gaze shift is "one in which the peer, when he or she shifts her gaze, continues her concurrent activities, readily alights and holds on a target, and directs action to the target" (2009: 150). The "noticing" gaze shift is "one in which the peer halts concurrent activities, readily alights and holds on a target, but does not produce action for the target" (2009: 153). Finally, the "search" gaze shift is one in which "the peer may stop or continue a concurrent activity, does not readily alight and hold on a target and makes appeals for assistance" (2009: 156). Thus, each type of gaze implicates a distinct social action that makes different responses relevant next.

Focusing this time on gaze withdrawal, Kidwell (2006) shows how it can be differentially interpreted depending on whether it occurs during a conversation or when it is responsive to an imperative to comply with a directive action (e.g. an order) such as the ones a police officer might produce during his line of duty. Gaze withdrawal in this case can be taken as an act of resistance, and is thus dispreferred. Evidence is provided in that mutual gaze is pursued via both *embedded* methods (e.g. speech cut offs, gaze-tracking and embodied summonses) or via *exposed* methods (e.g. verbal commands to *Listen* and *Look*).

Finally, Haddington's work (2006) focuses on how gaze can be used, in relation to the production of assessments, to display stances toward what has been assessed. He presents three types of gaze behavior (looking together at an assessable, mutual gaze during an agreeing second stance and cut-off gaze during actions that display divergent stances) and describes how they achieve stance-taking by interacting with what the participants are saying and doing during the interaction.

While this work on gaze and its relationship to social action is relatively new, research has sometimes touched on the importance of mutual gaze or eye contact in conversation in *ad hoc* situations, while focusing on other interactional practices. For example, Egbert (1996) suggests that in German, the use of the repair initiator *Bitte?* is highly context sensitive. Germans tend to use *Bitte?* when there is no mutual gaze between the participants (including situations in which eye contact is impossible, such as phone conversations), while they tend to use other repair initiators such as *Was?* when they are co-present and have mutual visual access. Sidnell (2006) also touches on the relevance of gaze in his investigation of how recipients identify and distinguish re-enactments from the main reporting of facts during a telling. In particular, he shows that during re-enactments speakers tend to look away from the addressee and keep their eyes away from the audience during the re-enactment. By doing this, the speaker shows that s/he is doing something different than addressing the other participant (cf. Goodwin, 1979), and indeed s/he is doing a re-enactment. Finally, in an examination of interviewee laughter in broadcast news interviews, Romaniuk (2009) suggests that interviewers withdraw their gaze when interviewees are laughing in such a way as to systematically terminate the relevance of responsive laughter.

This section has described some recent studies that focus on how gaze and other embodied displays can be used by themselves to accomplish social actions, and also presented research that shows how gaze behaviors are constitutive of social actions, even ones performed mainly through talk. More generally, in agreement with Goodwin's statement (1981: 30) that gaze "is not simply a means of obtaining information, the receiving end of a communication system, but is itself a social act", this dissertation will focus on the systematic organization and deployment of eye gaze in social interaction. However, contrary to previous work, this work will show that gaze in interaction is not organized primarily by reference to turns-at-talk. Gaze behavior is mainly organized in relation to sequences of talk and the development of courses of action, to be subsumed under the general umbrella of "ongoing projects", so that most of the variation in gaze direction should be observed at the beginning or at possible completion points of courses of action accomplished through one or more

sequences of talk. Note here that with the terms "courses of action accomplished through one or more sequences of talk" I refer to the fact that, for the interactional project to be considered completed, most actions require at least the occurrence of some sort of response or reaction by the other participant, and therefore an initiating action usually starts the development of a course of action produced by more than one participant. For the project to be complete and considered accomplished, more things have to happen. For example, a request for information, the action accomplished for example by a turn at talk, can be considered accomplished only if the other participant provides the information, and an offer is completed only if it is accepted or rejected and the thing offered is provided to the person to whom it had been offered. This means that the occurrence of an utterance and the action(s) that it implements opens the possibility of (and sometimes it normatively expects) the occurrence of another set of utterances or actions that would allow the initiating action to reach socially appropriate completion. This can be achieved in two turns or may require larger structures and therefore more than one sequence, though the participants' orientation toward completing the gist of the initial action remains the same.

More details concerning previous claims about gaze behavior in social interaction and how my position differs or aligns with those claims will be further outlined in the introduction of each following chapter.

## 1.5 Data and Method

In order to analyze and comprehend how participants approach the boundaries of sequences of talk and actions, I recorded a series of videos of naturally occurring interactions, interactions in which people are having dinner together, playing cards, preparing food, having coffee or traveling together in a car, without providing them with any kind of script or instructions for the interactions that later occurred and without asking them to get together just for the recording. The recordings, then, were of interactions that would have naturally occurred independently of the presence of my camcorders and of my request to tape them and therefore their occurrence was more spontaneous and less likely to be staged. Most of these interactions were dyadic, though some triadic interactions have been used as well to develop basic observations about the organization of gaze behavior. Details about the participants, the situation, the location and the main activity performed during each interaction recorded are presented in Appendix A. Table 1.1 shows the amount of data collected, transcribed, annotated and analyzed for the work reported in this dissertation.

Table 1.1. Amount of data collected, transcribed, annotated and analyzed.

| What | How Many |
|---|---|
| Data collected | Approx. 18 hours |
| Data transcribed | Approx. 5 hours and 30 minutes |
| Data fully annotated[15] | Approx. 3 hours and 30 minutes |
| Data analyzed qualitatively | Approx. 12 hours and 30 minutes |
| Data analyzed quantitatively | Approx. 1 hour and 40 minutes |
| Total number of participants in data analyzed | 28 |

Using this kind of data is crucial for the kind of investigation that I wanted to develop. Only by looking at what people really do in interaction was it possible to observe some recurrent practices for doing things with gaze. The need for this qualitative analysis, already emphasized in the quote by Lorenz reported above, is strongly related to something often observed by conversation analysts and remarked 40 years ago by Harvey Sacks (1992: 240,

---

[15] A large part of the data has been annotated for any shift in eye direction, for any shift in head direction and any postural shift and any nod by each participant. Multiple other annotations have been added with respect to the organization of the talk and social actions in those conversations, in relation to the specific practices under investigation, as it will be outlined in the following chapters.

Vol. II): "a base for using close looking at the world for theorizing about it is that from close looking at the world you can find things that we couldn't, by imagination, assert were there."

Given the focus on the systematic organization of gaze behavior in relation to social action, mainly performed through talk-in-interaction, I needed a methodology that would allow me to dissect and identify the conversational structures independently of what my working hypotheses concerning gaze were. Conversation Analysis (henceforth CA) provided this framework. This methodology focuses on action and interaction rather than on language and linguistic categories. CA allows a researcher to enter the granularity of interactions and to analyze details without losing touch with such macro structures as "sequence organization" and the "overall structural organization" of the whole event. Moreover, the focus on participants' orientations and displays of understanding of the actions of the other co-present participants, prevents an *ad hoc* use of categorical labels and analysis that cannot be supported by what can be tracked back in the data. The "proof procedure" of participants' uptake, central to the entire conversation analytic enterprise, has been fully adopted in this work: I will not simply describe when people move their eyes, but try to document what kind of interactional consequences a shift in gaze orientation can reliably produce. I will be using Conversation Analysis because it aims to investigate empirically how talk becomes action, how action is structurally organized through talk and how participants understand the sense of what they are doing while they interact.

However, the data reported here have not been analyzed only qualitatively. Insights from other disciplines will also be used in the general analysis of the data. As previously stated, I started from a qualitative analysis of single stretches of interactions, comparing ones in which some specific practice appeared to be deployed. After outlining the basic interactional features of each practice, I went back to the data, coded systematically what was happening in a specific amount of minutes (usually 10 minutes, as is indicated in each chapter) randomly picked from each interaction. The coding was done on the variables that had been previously identified as crucial for the practice investigated. I then ran statistical analysis of the coded data (mainly logistic regressions)[16] to identify significant correlations between variables, most prominent predictors of specific outcomes, and more generally to

---

[16] Logistic regression is a statistical model used to predict outcome variables that are categorical (e.g. yes or no). In this case, if a participation role (e.g. being a speaker) predicts significantly better than the other (being a recipient) whether an individual will be looking at the other participant, then we can say that being a speaker rather than a recipient significantly affects the likelihood of the occurrence of gaze toward the other participant. In other words, participants look at each other differently depending on which participation role they have during the social action under investigation.

identify and quantify the systematicity of a practice by trying to falsify the empirical hypothesis concerning the function that a specific gaze practice might have, hypothesis developed while qualitatively analyzing the participants' gaze in my corpus (Popper, 1959 [1934]). In other words, I tried to put to practice the method indicated by Darwin (1872) and further summarized by Lorenz (1966) as "a biological approach to behaviour" in the quote reported above.

The use of a statistical approach fully informed by CA findings is not alien to recent CA research. As Heritage and Maynard (2006a) put it:

> *To extract robust outcome-based conclusions about how physicians (or patients) should conduct themselves in specific moments in the flow of the medical encounter, it is important to find a meeting point between the two methodologies of coding and microanalysis [...] In other words, beyond the intrinsic worth of an analytical framework responsive to very granular, individual moments in the physician-patient encounter, we need one that simultaneously supports coding at a broader level of granularity sufficient to reach beyond individual cases to generate findings at a statistical evidential standard." (Heritage & Maynard, 2006a: 8)*

In recent years, several CA works usually in institutional settings, have looked at an interactional phenomenon and used quantitative methods to assess the association between that phenomenon and some exogenous variable whether that be time (e.g. in Clayman, Heritage, Elliott, & McDonald, 2007), prescribing outcomes/provider perception (e.g. in Mangione-Smith, Stivers, Elliott, McDonald, & Heritage, 2003) or race and class (e.g. in Stivers & Majid, 2007).

One reason for adopting a statistical approach was to understand whether gaze practices could be systematically identified across situations and individuals and to see whether the predictions about their generality, using qualitative methods, would survive an attempt to put them to test, in particular once applied to a larger number of instances, across different speakers and recipients, where the actions and context are in some sense comparable. Further details about how comparable actions and contexts have been identified for quantification purposes will be reported in each chapter.

To annotate the details of talk and visible behavior from each interaction, I used the software ELAN (http://www.lat-mpi.eu/tools/elan/), developed at the Max Planck Institute for Psycholinguistics to allow researchers to view and simultaneously annotate multiple recordings (video and audio) of the same interaction (see Fig. 1.6 below).
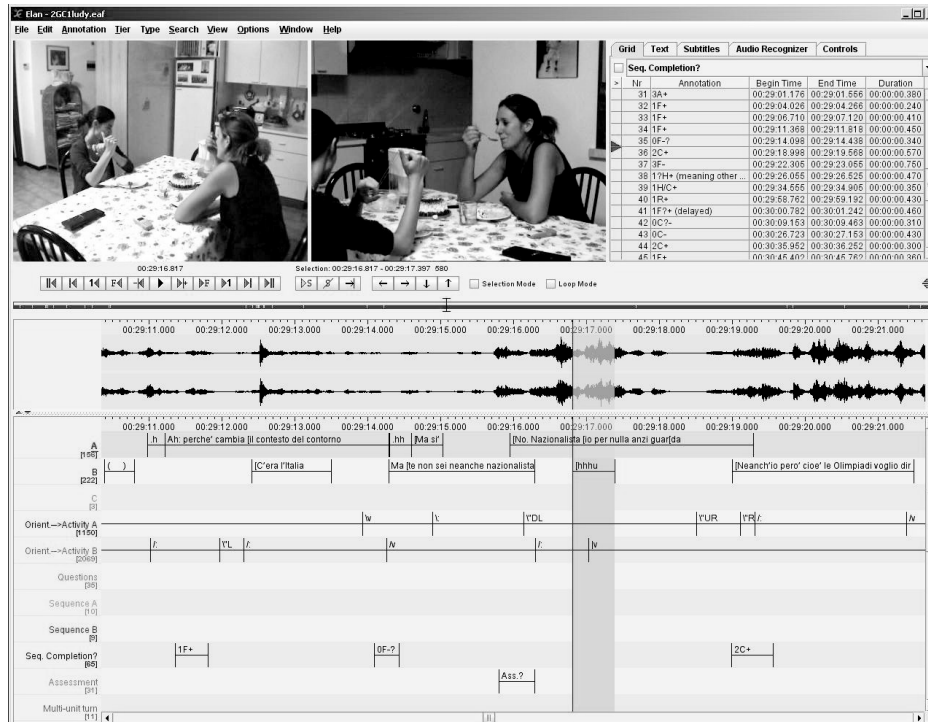
Fig. 1.6 Screenshot of recordings of an interaction annotated with ELAN

The playback and search functions proved extremely functional both for qualitative analysis and for the analysis of a large database. The layout of the software, that shows the annotations on a time scale, facilitates the identification of silences and overlaps. When this is combined with annotations in multiple layers, it allows the researcher to immediately identify whether specific visible behaviors were deployed, e.g., before or during the occurrence of a silence. Using the playback functions then the researcher can control for the correctness of the observation, as it is then possible to play just the specific bit of interaction under examination and verify in more than one video what happened exactly there. This software allows independent coding for any category we might be interested in, but at the same time it shows its role in context, as other annotations and layers can be visible at the same time.

For the statistical analysis I used the software SPSS and the software STATA.

## 1.5.1 Transcription System

The transcript of example 1.1 contains features that I will now explain. Talk is transcribed using the Jeffersonian system (Atkinson & Heritage, 1984; Jefferson, 2004a) usually adopted by Conversation Analysis. Given the use of Italian data, I adopted a 3 lines type of transcript, in which the first line is the text in Italian (the language in which the interactions occurred), the second line is a word by word translation from Italian to English and the third line is an English 'free translation', meant to convey the main gist of the utterance, both semantically and style wise.

In addition to the content and style of the talk, additional symbols are present in those transcripts and they represent a new transcription system for eye-gaze in face-to-face interaction that I developed and that can be found in Appendix B. A published example of it can be found in Rossano et al. (2009). It differs from the previous transcription system used within Conversation Analysis and invented by Goodwin (e.g., 1979, 1980, 1981). In what follows I outline some of the basic features of my system and some of the differences between Goodwin's and mine.

1.    It is ideal for dyadic interaction, though it could be used for multi-party conversations as well.

2.    It is more iconic, through the use of ovals representing faces and arrows representing gaze direction, than any other system currently used.

3.    It is combinatorial. Each participant's gaze behavior can vary independently from the gaze behavior of the other. The number of possible symbols depends on the number of individual gaze orientation we deem worth distinguishing. In other words, the system can easily be increased if more features are considered worth reporting in a transcript.

4.    It does not simply report whether a participant is looking at the other's face or not and whether there is a gaze shift. It adds many additional gaze configurations (e.g., looking at an object, closing eyes, looking away, looking up or down, etc.) that are interactionally relevant and meaningfully different from each other.

5.    By using the directionality of arrows, it partly communicates not just the direction of the eyes but also head movements (for example in the symbols

indicating gaze shifts, like turning or raising eyes up, or in symbols such as looking away or looking down).

6.      It does not represent individuals' gaze behavior as separate from the gaze of the other participant. By having both participants' gaze behavior represented in the same symbol, the system claims that the combination of both participants' gaze orientation is an interactionally significantly different configuration from any other. So if the speaker looks at the recipient but the latter has her eyes closed, this configuration has very different interactional consequences than a configuration in which they are both looking at each other. Facial expressions, or more simply a direct stare toward the recipient will not be perceivable by the latter and as such will affect what the speaker might expect from the other participants in terms of uptake. This also means, for example, that if the gaze configurations change during a silence, then that silence could be decomposed into different phases, to take into account the different potential moves that could occur at each point and to record the occurrence of specific interactional moves performed through gaze (chapter 3 will show examples of gaze pursuits during silence, gaze movements that perform social actions).

7.      By being located on top of brackets that have a beginning and end point (either on top of talk, or on symbols indicating the length of silence), each symbol indicates the stable gaze configuration in place during that stretch of talk or silence. The claim is that the length of a specific configuration might have interactional consequences (e.g., looking at each other during silence for an extended amount of time). As far as gaze is concerned, what matters is not just position and composition of the signal (Schegloff, 2007b) but also duration. This observation matches the "typical intensity" of a communicative signal as outlined in Morris (1957). Moreover, by indicating when it starts and when it ends, we can have a better grasp of the possible impact it might have on what gets said or done before or after that shift, knowing for example that the shift occurs on top of a specific word or after a specific amount of silence (see chapter 3 for a reflection on the length of silence and occurrence of pursuits).

8.      The transcription system takes its metric from the talk, given that the focus of this work is to see the systematicity of gaze behavior with respect to social actions in conversation.

The meaning of each gaze symbol is described in Appendix B.

## 1.6  Structure of the Dissertation

This dissertation investigates how participants use their eyes at the beginning, in the middle and at the end of sequences of talk in interaction. Each empirical chapter specifically investigates one of these three contexts. Chapter 2 investigates how gaze behavior is deployed at the beginning of specific conversational units such as extended tellings and adjacency pair based sequences. I show how recipients use gaze to systematically differentiate between the two conversational environments. The chapter introduces the relevance of investigating the structure of social actions performed through talk to better understand how gaze behavior is deployed during a conversation. Moreover, it investigates how recipients manage to enact the practice reported here so quickly and so reliably.

Chapter 3 investigates how speakers behave in the middle of a sequence of talk and how they can use gaze to pressure for responses. I focus on situations in which speakers who have produced a sequence initiating action that did not obtain a response pursue a response not simply verbally but also through using gaze toward the recipient. I first show how speaker and recipient gaze is deployed during sequence-initiating actions and how it affects the occurrence and the timing of a response. Then I move to instances in which a response appears to be missing and it is pursued through gaze. Finally I show how the practice of soliciting response/uptake via gaze can be deployed in contexts other than a transition relevance place.

Chapter 4 investigates how people close conversational units such as courses of action and sequences of talk. In particular, it shows how withdrawing gaze from the other participant approaching sequence possible completion can work as a bid for closure and if aligned by the other participant can lead to the actual closure of a sequence. Sustaining gaze toward the other person when a possible completion of the sequence is reached, indicates the intention to expand or the understanding that the sequence is not complete as such and would need to be expanded.

Finally, chapter 5 provides the conclusions of the dissertation. It summarizes the main findings placing them in the broader framing partly introduced in chapter 1. Finally, it discusses a model that describes how gaze behavior in interaction is deployed, what are the driving forces in place and which kind of factors might constitute exceptions to its otherwise predictable and systematic deployment.

This dissertation also provides a methodological framework that can be used to extend the findings reported here to related conversational environments or interactions occurring in other cultures or among other species. The findings should provide a useful improvement of our current knowledge about eye movement and gaze behavior during social interaction and may be of particular use for researchers interested in implementations for artificial intelligence or the study of cross-cultural interactions.

## 2 The Organization of Gaze Behavior in Two Different Sequential Environments

**How we depict any segment of the sequence is related to our conception of the whole.**
Edward M. Bruner *Ethnography as Narrative* (1986)

### 2.1 Introduction

When do people look at others and when do they look away? What are they doing by looking at another person and by looking away? How do people decide when to look at and when to look away? In this chapter I present work that contributes to answering these questions. I show that gaze behavior in face-to-face interaction is orderly, specifically in its relationship to the sequential environment of the talk and the activities that the participants are implementing through their turns. In particular, I contrast recipient gaze patterns in two sequential environments: during the first turn constructional unit[1] (henceforth TCU) of an extended-telling sequence (e.g., a story telling, a report) and during the first TCU of an adjacency-pair-based sequence (e.g., requests or offers).

I first demonstrate that there is a systematic difference between the gaze behavior of a participant listening to the first TCU of an extended-telling sequence (henceforth ETS) and the behavior of the same participant listening to the first TCU of an adjacency-pair-based sequence (henceforth APBS). Whereas recipients regularly look toward speakers as soon as an ETS is projectably underway and sustain gaze toward speakers throughout the telling, they usually do not do so while listening to a first pair part[2] turn. This suggests that being the recipient of a telling requires a different type of behavioral participation when compared to listening to a simple question. As such the findings presented here will provide a more fine-

---

[1] Turn constructional units (TCU) are the building blocks of turns at talk. A TCU is the most minimal utterance produced by one participant that, in context, can be considered a comprehensible contribution to the interaction and a complete one. They "can by themselves constitute possibly complete turns and on their possible completion, transition to a next speaker becomes relevant (although not necessarily accomplished) (Schegloff, 1996c: 55). See also § 2.2.4 for further clarification of this concept.

[2] A first pair part turn is a turn that makes conditionally relevant the occurrence of a second turn and its absence a noticeable absence. For example, an invitation makes relevant an acceptance or a rejection, a greeting makes relevant the return of the greeting by the other participant (see Schegloff, 2007b for a systematic description of the types of first pair parts deployed in conversation).

grained deconstruction of the hearer role than the one outlined by Goffman in Footing (1981a). I then examine the mechanisms that allow recipients to project which type of sequence a particular TCU is in the service of. To do that, I introduce the turn design features that tend to occur at the beginning of these different sequential environments and I propose that these features are crucial for the orderly occurrence of recipient's orientation toward the speaker. Finally, I suggest some implications of the findings for our understanding of how human cognition works in interaction, and I address the issue of the proper structure of heuristics for human interaction.

## *2.2 Speaker and Recipient Gaze*

In this section I first review what existing studies have shown about the difference between speaker and recipient gaze in interactions. I then challenge the general claim about their differences on the basis of a qualitative analysis of dyadic interactions. In order to determine an appropriate way of investigating the level of order of the new patterns identified, I propose further refining the domain of investigation. Finally, I provide an overview of what is known about ETSs in conversation as a way into the comparison of gaze behavior in different sequential environments.

### *2.2.1 Background*

It has long been claimed that in dyadic interaction people tend to look at the other participant more when they are listening than when they are speaking (Argyle & Cook, 1976; Argyle & Dean, 1965; Argyle & Graham, 1976; Bavelas, Coates, & Johnson, 2002; Duncan & Fiske, 1977; Exline, 1963; C. Goodwin, 1980, 1981; Kendon, 1967, 1973, 1990; Kleinke, 1986; Nielsen, 1962; Rutter, 1984).

As said in chapter 1, the first person to note this difference was Gerhard Nielsen in his work on self-confrontation (1962). It was, however, Adam Kendon (1967: 27) who provided a more precise description of the different patterns of speaker and hearer gaze:

> *during listening, p [the speaker] looks at q [the hearer] with fairly long gazes, broken by very brief gazes away, whereas during speaking he alternates between gazes to the recipient and gazes away of more*

*equal length, the gazes away being longer than those that occur during listening. (1967: 27)*

Charles Goodwin suggested that quantitative measures of overall frequencies of gaze "may not be an appropriate way to study the organization of gaze as an activity in its own right" (1981: 94) and relying on a case by case analysis of the data he proposed two rules (1980: 275, 287; 1981: 57; 1984: 230) to account for participants' gaze behavior:

1. "A speaker should obtain the gaze of his recipient during the course of a turn at talk"
2. "A recipient should be gazing at the speaker when the speaker is gazing at the hearer"

According to Goodwin, if the recipient looks most of the time, then the speaker will find her/him gazing back any time the speaker looks toward the recipient. On the other hand, if s/he is not looking at the speaker, the latter has resources (phrasal breaks, pauses, restarting the turn) to solicit and obtain the recipient's gaze. By proposing these as rules[3] Goodwin claims that participants' gaze behavior is interrelated rather than independent and suggests normative measures to maintain that level of order possibly for every turn of talk, giving Kendon's observations a normative force. Indeed, if the recipient looks all the time and the speaker glances back and forth then there should be no need to deploy procedures to obtain recipient's gaze orientation toward the speaker. Therefore Kendon's gaze patterns should be the default for every turn of talk.[4]

Kendon's remarks suggest something intriguing from both an interactional and a semiotic point of view: a careful observation of just the eyes of two participants in interaction would allow us to distinguish who is speaking from who is listening and when they exchange these roles. "Kendon's gaze patterns" have been apparently replicated and re-stated in all the subsequent studies mentioned above.

---

[3] In a footnote Goodwin (1981: 57) admits that even though he proposes these as rules applying to turns in general, this pattern is not found in every turn of talk.
[4] Kendon (1967) does actually indicate that this is a generalization and not every single turn of talk necessarily follows the pattern he is suggesting. The goal here is to take seriously the pattern proposed and at the same time the possibility of deviating from this pattern, only to suggest a different level of organization for gaze behavior that does not treat deviation from the pattern as "deviant behavior", but rather normative gaze behavior, only limited to a specific sequential environment.

The literature suggests that there are few doubts about the causality of this phenomenon. Indeed, the general idea is that the different gaze patterns of the participants are due to the fact that somebody is speaking and somebody else is listening and the social roles of speaker and listener associated with them. The claim usually is that the occurrence of talk in a face-to-face interaction leads to the occurrence of these gaze patterns because of a conventionalized way[5] of looking associated with a specific interactional role. The evidence for this claim comes from the fact that the participant role (of speaker and listener) is usually taken as the only independent variable and the occurrence of gaze as the dependent one. In other words, a participant would look at the other in a certain way and for a certain amount of time because s/he is speaking or because s/he is listening to the other participant, independently of the actions performed through the talk.

Regulatory functions, in particular in relation to turn taking, have been claimed for the deployment of gaze (see chapters 1 and 3 for a review). For example looking toward a recipient or looking away could be used as a cue that a speaker is or is not going to select that person to speak next (see, e.g., Lerner, 2003). This would suggest that gaze also affects the exchange of speaker/hearer roles. As such, a straightforward causal chain that links how gaze is deployed to the simple occurrence of talk should be questioned. This leads to an empirical question: is the order of gaze behavior simply due to the participation role of speaker and listener, or does it depend on the details of what participants in those roles do and have to do in terms of social actions?

If gaze is organized in relation to the participation role of speaker and listener, then speakership and recipiency would matter. There should be no difference if any distinction is made within the stretches of talk, in particular in terms of the sequential environment in which they are uttered and the social actions that are performed through their talk. If gaze is organized in relation to social action and what speakers and recipients do in each specific social action, then it should be possible to observe some systematic difference in people's gaze patterns in relation to what they are doing with their talk. In what follows I test these hypotheses through a systematic analysis of empirical data.

---

[5] Some evidence that specific gaze behavior might be conventionalized comes from recent work (Rossano, Brown, & Levinson, 2009) comparing speaker and recipient gaze behavior during questions in three different cultures that displays that there is some cultural variability, in particular in terms of how recipiency is enacted and sustained. However this work does not distinguish speaker and recipient in general, but only during question turns and therefore focuses on more specific action types than Kendon's work.

## 2.2.2 Contrastive Evidence for Kendon's Pattern

A qualitative analysis of twelve hours of recordings of naturally occurring dyadic Italian data suggests caution in embracing Kendon's gaze patterns as they stand. Although it is true that on average listeners look at speakers more than speakers look at listeners, they do not do so invariably. A systematic analysis of video data shows that the length and the patterns of gaze toward the other participant vary widely during a single conversation. At times a recipient does not look at all at the speaker, and a speaker sometimes holds her/his gaze toward a recipient even when the latter is not looking back or during gaps.[6] This means that there could be a substantial amount of time in which Kendon's patterns would not accurately describe the data and thus undermine prior claims (e.g. by Goodwin) that the recipient looking toward the speaker is a norm-governed phenomenon. In order to show some of the differences between these situations, I will present four extracts taken from two different interactions (two supporting Kendon's patterns and two that contrast with his claim). I then suggest that what distinguishes them and what matters for the participants' organization of gaze behavior is not only who is speaking and who is listening but also the sequential organization of talk.

### 2.2.2.1 Supportive Evidence

The two following examples represent supportive evidence for the gaze patterns described by Kendon. They are both dyadic interactions in which the participants are female undergraduate students and long time friends.

In example 2.1, A and B are two women in their 20s. They are sitting on two couches drinking a cup of tea and chatting about their babysitting jobs. Before this extract, B told A that she does not know how to ask to be paid more by the family she is babysitting for because they are also friends of hers. A says that she is a friend of the main family she works for too but on the other hand it is also fair to be paid given that it is a job (see line 1). The gaze behavior of both participants strongly resembles the one proposed in Kendon's pattern, in which the speaker alternates looking toward and looking away from the other participant while the recipient tends to look toward the speaker.

---

[6] See chapters 3 and 4 and Rossano et al. (2009) for further evidence supporting this claim.

57

**(2.1) 2GSOFA_aumentato 23:32**

```
01 A:  .hh  e'   giusto anche che  me li  diano  .hh
            be.3s right  also that me cl. give
       .hh it is also fair that they give them to me .hh
```

```
02  pero' in effetti non saprei  neanch'io co-<↑una volta
    but   in effect  not know.1s neither I  wh- one time
    but actually I would not even know wh-<↑once
```

```
03  io ho   aumentato
    I  have raised
    I have raised ((the price))
```

```
04   (0.5)   (0.2)
          (0.7)
```

```
05 A:  mi ricordo che:  .hh ho      aumentato e >ho     detto tipo<
       me remember that     have.1s raised    and have.1s said like
       I remember tha:t .hh I raised ((the price)) and >I said like<
```

```
06  hh c(h)ioe' non mi r(h)icordo bene le mod(h)alita'
              that is  not me remember  well the modalities
    hh I m(h)ean I don't r(h)emember well the det(h)ails
```

```
07  pero' .h  tipo uhm:: (0.6) era stato magari .h dopo  l'estate
    but       like uhm        was.3s      maybe    after the summer
    but  .h  like uhm:: (0.6)  it was maybe .h after summer
```

```
08  (0.4) quindi ero    tornata (0.3) e  >tipo ho     detto<
          so     was.1s come back     and like have.1s said
    (0.4) so I had come back ((from holidays)) (0.3) and >like I said<
```

```
09    beh   dai    da  quest'anno faccio otto   euro<una cosa cosi'
      well give.2s from this year make.1s eight euros a   thing so
      well come on from this year it is eight euros<something like that
10    (0.8)
11 B: Ah con loro?
      Oh with them
      Oh with them?
12    (.)
13 A: ((Annuendo)) Si'
                   Yes
      ((Nodding)) Yes
```

When this extract starts, participant A is holding a cup of tea in her hands and puts it down while producing lines 2 and 3 (the cup reaches the table in front of her while she is producing the first person pronoun "io" at line 3). A then reaches for a biscuit on the table and holds it in her hands until the completion of line 9, when she brings it to her lips and starts eating it. Participant B, on the other hand, is just sitting on the couch oriented toward A and neither eats nor drinks during this stretch of interaction. A produces most of the talk during this example, and she is the only speaker between line 1 and 9, where the direction of both participants' gaze is transcribed.

If we look at the speaker's gaze (A's gaze) we can see that it alternates between looking at objects (lines 2-5 when she looks down toward the table and the biscuits, and line 9, when she glances at the biscuit before putting it into her mouth),[7] looking into the middle distance (lines 6, 7 and 9), looking at the other participant (lines 5, 7, 8 and 9) and simply closing her eyes (lines 5-6). If we look at the recipient's gaze (B's gaze) we see that she looks at the speaker's face most of the time (lines 1-9) (with only a couple of exceptions when she is looking at the table where A is looking [lines 2-3 and lines 4-5]).

Example 2.2 shows a similar pattern. Here the two participants are friends who attend the same courses at the university. B is visiting A for the holidays, and they alternate

---

[7] Notice here how the relationship between vision and the motor control of parts of our body, in particular the hand, outlined in chapter 1 and claimed to be an evolutionary adaptation of primate vision, can be observed in any instance of human interaction. A looks at the biscuit before reaching for it on the table and then looks at it when she has it on her hand before putting it into her mouth to calibrate the distance, the precision of the grasping, the directionality of the movement and the physical effort to be applied to accomplish this apparently trivial task.
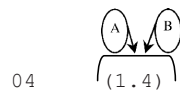
studying and chatting. A and B are sitting facing each other, A on a bed and B on a chair with a table on her right side. A few minutes before the beginning of this extract, B has made a phone call to her mother to let her know when she will return home. Line 1 completes a comment by A in which she says that she is glad that even engineers sometimes do not feel like working. The implicit reference here is to some friends of hers that the two women visited the night before. In what follows B reports the exchange with her mother on the phone. Mark is B's boyfriend who is, at that moment, in another country and did not join her for the holidays. The two friends are clearly playing with the fact that while her boyfriend is away she can still go out and meet nice young men. Later in the interaction B will stress that she was just joking about this and this is also clear from the smiles during this exchange.
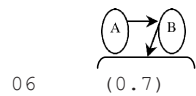
**(2.2) 2GSTUDYING_mamma 29:36**

```
01 A:   anche gli ingegneri
        also  the engineers
        the engineers too
02      (0.3)
03 B:   °Infatti°
         Indeed
        °Indeed°

04      (1.4)

05 B:   Mi fa   mia mamma hai    conosciuto dei  bei      ragazzi? Si'
        Me does my  mum    have.2s known       some handsome boys    Yes
        My mum says to me "have you met some handsome boys?" "Yes"

06      (0.7)
        (0.7)

07 A:   A:[::h
        Oh
        O:[::h
```
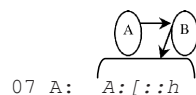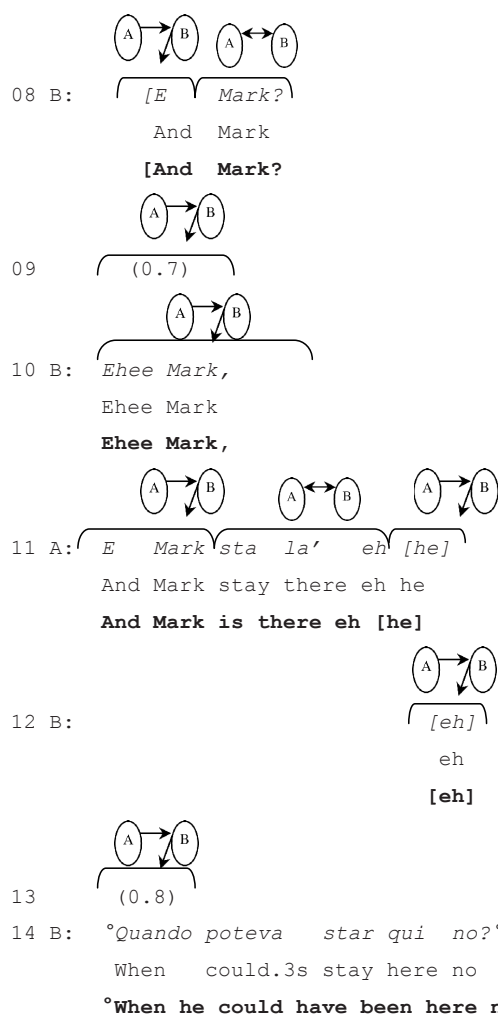
08 B:     *[E    Mark?*
          And  Mark
       **[And  Mark?**

09      (0.7)

10 B:    *Ehee Mark,*
          Ehee Mark
       **Ehee Mark,**

11 A:   *E    Mark sta la'  eh [he]*
        And Mark stay there eh he
      **And Mark is there eh [he]**

12 B:                *[eh]*
                  eh
                **[eh]**

13      (0.8)

14 B:    *°Quando poteva  star qui  no?°*
        When    could.3s stay here no
     **°When he could have been here no?°**

If we look at the turns in which B is the speaker we see that most of the time she is looking down, but she looks at the recipient twice (at lines 8 and 12). In contrast, A as a recipient looks at B most of the time (lines 5-13). We can also note that A looks at B while speaking (lines 7 and 11), while B looks at A only briefly (line 11) while being a recipient. This example shows a bit more variety than the previous one in terms of speaker-recipient gaze behavior, but we can still say that if we consider B as speaker and A as recipient their behavior supports Kendon's patterns.

What we have seen in these two examples suggests that speaker and recipient indeed display different patterns in terms of their gaze behavior and these patterns strongly resemble what Kendon suggested more than 40 years ago. On the other hand, example 2.2 has already

shown two turns of talk in which the pattern seems to be reversed, with the speaker looking all the time and the recipient looking up only briefly. This is not an isolated case, as will be shown in the following section.

### 2.2.2.2 Counter Evidence

Although the pattern described by Kendon is common, there are numerous instances in a conversation in which the gaze behavior of the participants does not follow this pattern. The following two examples illustrate this. They are taken from the same two conversations, so the participants, the reason for meeting, the general ongoing activity and all the other contextual features remain the same. If we can still see a difference and this seems to be systematic, then a different hypothesis should be provided for the organization of their gaze behavior.

Example 2.3 is taken from the same interaction as 2.1. Before the beginning of this extract the two women were commenting about the drink they had and whether it was tea or something else. A then suggested that the smell of it resembled henna, a product used to paint one's body with a short lasting tattoo (see line 1). In line 3 she begins a completely new topic. Also in this fragment, A is holding a cup in her right hand, which she puts down on the table during the turn at line 7. B is not holding anything in her hands until the completion of line 5, when she moves closer to the table and during the silence at line 6 picks up a biscuit which she puts into her mouth after completing the turn at line 7. The ongoing competing activities are therefore the same as seen during example 2.1.

**(2.3) 2GSOFA_come  25:16**

```
01 A:   Questa e' henne p(h)ura [hh
        This   is henne pure
       This is p(h)ure henna [hh
02 B:                          [hhu h[h


03 A:                          [Ieri     ho     visto il  Milan
                                Yesterday have.1s seen  the Milan
                               [Yesterday I saw Milan ((playing))
```
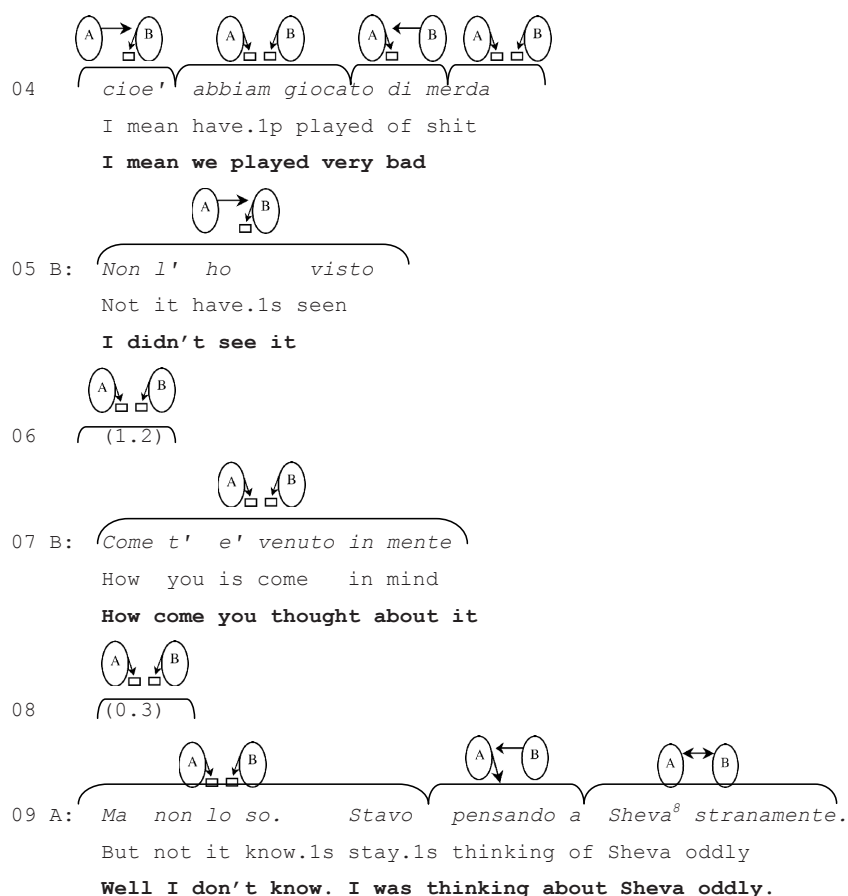
04     *cioe'* | *abbiam giocato di merda*
        I mean have.1p played of shit
        **I mean we played very bad**

05 B:  *Non l' ho    visto*
        Not it have.1s seen
        **I didn't see it**

06     (1.2)

07 B:  *Come t' e' venuto in mente*
        How  you is come   in mind
        **How come you thought about it**

08     (0.3)

09 A:  *Ma  non lo so.    Stavo   pensando a   Sheva[8] stranamente.*
        But not it know.1s stay.1s thinking of Sheva oddly
        **Well I don't know. I was thinking about Sheva oddly.**

Let's look at gaze behavior from a speaker-recipient perspective. When A is the speaker, she alternates looking at the recipient (lines 3-4 and 9) and looking down toward either her cup (lines 3-4) or the table (line 9). B, as a speaker, does not look at A but rather orients toward the table and the activity of picking up a biscuit from the table (lines 5 and 7). On the other hand, she does not look much at A even while acting as a recipient. Indeed she looks only very briefly during line 4 and for about half of her turn at line 9. While acting as a recipient A looks at B during line 5 but not during line 7. The speaker pattern is quite similar to what we saw in the previous examples while the recipient's gaze pattern is quite different. There is no systematic sustained gaze at the speaker at any point. The generalization proposed by Kendon cannot account for this piece of data. It could be claimed that the difference is due only to the timing of the competing activities (picking up a cookie and putting down a cup).

---

[8] Sheva is the nickname of a soccer player who was playing for the Italian football club Milan at the time of the recording.

The competing visual involvement of the participants could be taken as a reasonable account for their not looking toward the other participant (see C. Goodwin, 1981). However, as will be shown later in the chapter, participants can and often do disengage from looking at other objects or being involved in competing activities because of what is happening in the ongoing conversation and what the latter requires in terms of visual engagement and display of participation. It is always a choice and at least a possible alternative. The question is rather whether in some cases abandoning the competing visual engagement has priority over its continuation.

In order to show that example 2.3 is not an isolated case, let's look at another example (2.4) taken from the same interaction as example 2.2. A and B are preparing for an exam and A just said that she was trying to show B on the piece of paper she is holding what B had to say, like a TV anchorman who reads the news. They laugh about it (see lines 1 to 6).

**(2.4) 2GSTUDYING_testa   29:07**

```
01 B:  hh [ha ha]
02 A:     [Q(h)uello che  devi    leg]gere
           That      that must.2s read
          [W(h)at you have to re]ad
03     (0.4)
04 A:  hh hu
05     (0.9)
06 B:  .hhh
07     (0.7)
```
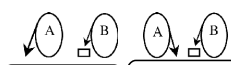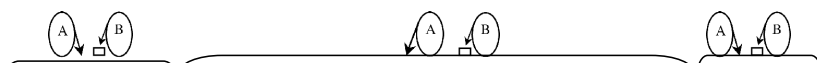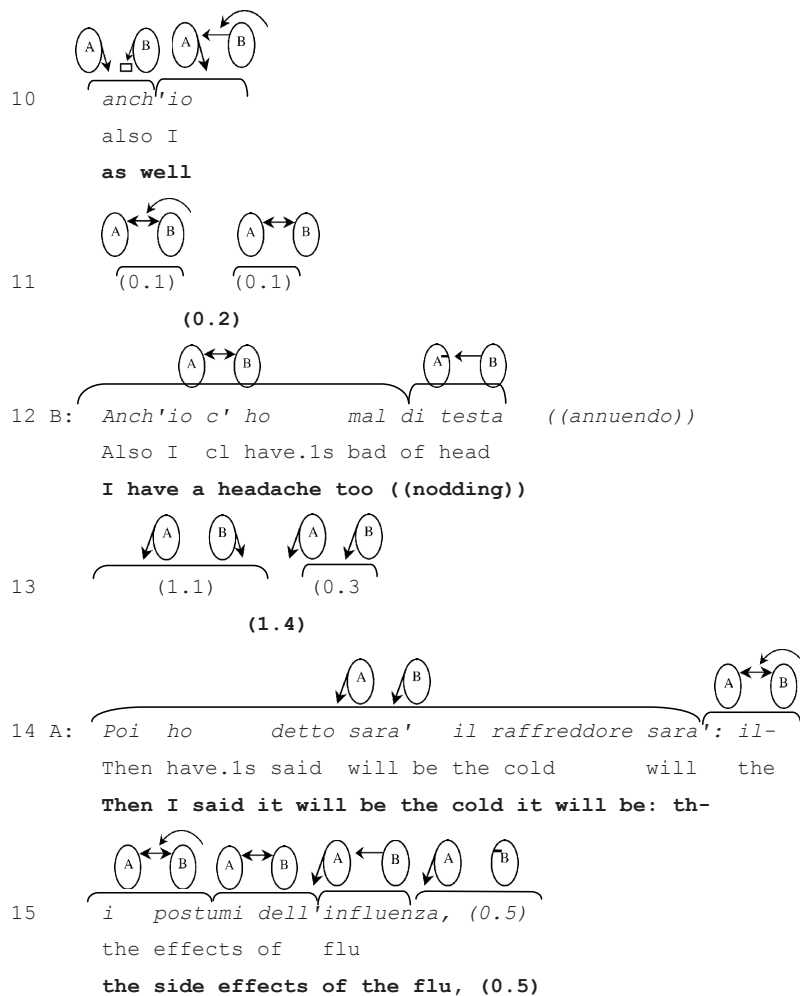


```
08 B:  Okay quindi    (0.3)
       Okay therefore
       Okay therefore (0.3)
```



```
09 A:  .h No io veramente son scarburatissima ho    un mal di testa
          No I  really   am  out of steam   have.1s a  bad of head
       .h No I am really out of steam I have a headache
```

10     *anch'io*
      also I
      **as well**

11     (0.1)    (0.1)
      **(0.2)**

12 B:  *Anch'io c' ho    mal di testa   ((annuendo))*
      Also I  cl have.1s bad of head
      **I have a headache too ((nodding))**

13     (1.1)    (0.3
      **(1.4)**

14 A:  *Poi ho    detto sara'  il raffreddore sara': il-*
      Then have.1s said will be the cold    will  the
      **Then I said it will be the cold it will be: th-**

15     *i   postumi dell'influenza, (0.5)*
      the effects of   flu
      **the side effects of the flu, (0.5)**

B turns toward some sheets on the table on her right while producing line 8 and continues looking at them during lines 9-10, while A alternates looking toward the desk and looking down during those turns. If we focus on the speaker's gaze behavior we see that B does not look at the recipient during line 8 while she does at line 11. A does not look at B at all during lines 9-10 and for part of lines 13-15, but during the turn at lines 14-15 she looks up toward B's face. This speaker behavior again resembles that seen in the previous examples. On the other hand, if we look at the recipient's gaze behavior we can observe that A does not look at B during line 8, while she does at line 11. And B does not look at the speaker during lines 9-10 and at the beginning of line 13, while she turns toward A mid turn and keeps looking

during line 14. Thus the recipient's gaze behavior more closely resembles the speaker's gaze behavior than the behavior described by Kendon.

These latter examples show an interesting diverging pattern in terms of gaze behavior when compared to the first two examples. This difference does not apply to both roles but rather to the recipient's. The speaker's gaze behavior is similar to that observed in the first two examples. However the recipient gaze behavior is radically different. If Kendon's patterns were correct, and if they were normatively governed, the behavior observed in the latter example should be unusual or treated as deviant, but it actually occurs quite often in these data. If we aim for a model that accounts for every instance of gaze behavior, then an alternative hypothesis in terms of the organization of recipient's gaze is needed - one that can embrace and at the same time explicate the different observable patterns and one that can suggest how participants know where to direct their eyes during conversation.

### 2.2.3 Two Sequential Environments

What accounts for the two patterns? Examples 2.1 and 2.3 are from the same interaction, just like 2.2 and 2.4. Therefore, the differences observed between the first batch of examples (2.1 and 2.2) and the second batch of examples (2.3 and 2.4) cannot be due to gender, general context of the interaction or whether the participants know each other. I propose that the differences are due to what participants are doing in the interaction at a micro level. In particular, I propose that the decisive factor affecting recipient gaze behavior is the sequential environment of the talk, e.g. whether the recipient is listening to the first TCU of an extended telling or of a first pair part action, such as a request for information or a complaint. In fact, the sequential environment of the talk is different in the two pairs of fragments. If we look back at the previous examples and we look at the sequential organization of the talk, we see that while the first two are extended reports, the last two examples are adjacency pairs.[9]

In examples 2.1 and 2.2 the recipient looks up when the production of a telling is projectable (lines 3-5 in example 2.1 and line 5 in example 2.2) and sustains her gaze throughout the telling (which incidentally is a report). These cases support the argument that gaze is organized also with respect to sequential environment rather than just participation

---

[9] It could be said that another difference is the length of participant's turns of talk, but this is an outcome of the kinds of actions the participants are performing. Moreover, if we look at examples 2.1 and 2.2 we can see that even among tellings the length can vary widely.

role. In both examples we see that recipients' gaze orients toward the speaker at the beginning of a projectable ETS and is sustained throughout it. The telling may start abruptly (as in ex. 2.1) or after some silence following a turn that could constitute sequence completion (as in ex. 2.2). In both cases the recipient is not looking at the speaker before the first TCU of the projectable extended telling but looks at the speaker before the completion of that TCU.

In examples 2.3 and 2.4 the sequential environment differs from the first two examples and so does the recipient's gaze behavior. In both cases the recipient does not look at the speaker during the sequence initiating turn (lines 3-4 in example 2.3 and lines 9-10 in example 2.4) while she does during the responsive turn. Note that in examples 2.3 and 2.4 the recipient of the initiating turn and the recipient of the responsive one are two different individuals, but the pattern observed in example 2.4 is the same as the one described for example 2.3: the speaker of the first TCU of an APBS keeps looking through the responsive turn, while the recipient of the first TCU of an APBS does not look up (or she does so only minimally during the second TCU as in example 2.3) even though two TCUs are produced.

These data extracts therefore suggest that participation status is not sufficient to explain the organization of gaze behavior. These data further suggest that the sequential organization of talk-in-interaction serves an important organizing function. The gaze behavior observable in extended tellings seems to correspond to Kendon's pattern, while in APBSs the pattern is different and the recipient often does not look at the speaker. In what follows I explain what the term ETS entails and what is already known about the organization of extended tellings and in particular of their beginnings.

## 2.2.4 Extended Tellings and Their Beginnings

Sacks, Schegloff and Jefferson (1974) defined as turn constructional unit (TCU) the smallest conversational unit that can constitute a turn of talk. Turns may be made up of only one TCU or of many TCUs. In outlining the organization of turn taking Sacks et al. show that:

> In starting the construction of a turn's talk, the speaker is initially entitled [...] to one such unit. The first possible completion of a first such unit constitutes an initial transition-relevance place. Transfer of speakership is coordinated by reference to such transition relevance places, which any unit-type instance will reach. (1974: 703)

Critical here is that when speakers get the floor, they are not guaranteed more than one TCU. Thus, producing more than one TCU is contingent and may require the deployment of some specific interactional practice. Schegloff (1982) describes some of the practices speakers use when attempting to produce a multi-unit turn:

a) The potential speaker may indicate, from the beginning of the turn, an interest in producing a more than one unit turn (for example projecting the production of a list by using a list-initiating marker such as 'first of all')

b) There can be a turn devoted to projecting the production of a multi-unit turn (see e.g. Sacks, 1974 on story prefaces)[10]

c) In positions different from turn beginning speakers may deploy other devices such as a 'rush through' to achieve a turn extension by speeding up the pace of talk approaching a possible completion of a turn constructional unit (see also Local & Walker, 2004)

d) After the possible completion of a one unit turn, it may be that no other co-participant starts a next turn so that if the prior speaker starts talking again, s/he may add a new turn constructional unit to the first one and thereby make it a multi-unit turn (Sacks et al., 1974; Schegloff, 1982)

The first two are ways in which the speaker projects from the beginning that her/his turn will be a multi-unit one. By contrast, the last two facilitate the occurrence of a multi-unit turn during the course of the turn, in a way that is much less predictable for the recipient.

In this chapter I have already identified a specific interest in recipient gaze behavior in ETSs (extended-telling sequences) and APBSs (adjacency-pair-based sequences). Extended tellings might occur either as a multi-unit turn or as a sequence that starts with an adjacency pair projecting that an extended telling will follow and in general both can be interrupted by repair sequences. While example 2.1 represents an example of a telling achieved through a multi-unit turn, example 2.2 represents an example of a telling sequence, where the recipient repeatedly intervenes and facilitates the production of the telling. In both cases, however, recipients[11] usually not only sustain the gaze throughout the entire telling, but they also tend to look up toward the speaker very early, before the completion of the first

---

[10] See Sidnell (2011: 174-196) for a recent summary of the main ways in which storytellings are launched in conversation.

[11] In example 2.2 I am labeling as "recipient" the person who has not initiated the telling, even though in two moments of the telling sequence she is actually speaking.

TCU that projects the occurrence of the telling. The hypothesis here is the existence of a strategy that could be described in these terms:

"To signal recognition and acceptance of an ETS, look at the Speaker"

To provide evidence supporting this hypothesis, the focus of the remainder of this chapter will be on projectable extended tellings. Indeed, the recipients have to parse every bit of talk and body behavior online to figure out what action will be relevant for them to make next. If a specific gaze pattern is regularly produced, then it is reasonable to assume that this behavior (looking up toward the speaker) is one of the relevant next actions that a recipient produces when an extended telling is recognizably in progress. To be able to do that, there must at least be a point in which the recipient is able to project that a telling is under way. Indeed, if it were otherwise, there would be no regular recipient gaze pattern so early in the extended telling sequence.

For this coordination to be successful, the collaboration of the co-participant is not only important but necessary, as s/he could block the production of an extended telling by simply taking the floor at the first transition relevance place. This implies that the co-participant must be able to recognize that, at least approaching the transition relevance place, the current TCU is not going to be the last one of the telling but rather there will be further ones. It is, indeed, at that place that the recipient is asked to withhold a full turn of talk or to produce minimal feedback (see e.g. Bavelas, Coates, & Johnson, 2000; Schegloff, 1982; Yngve, 1970). The latter usually occurs around the possible completion of most of the TCUs that constitute a multi-unit turn; this shows that each completion of a TCU represents an opportunity for the recipient to confirm her/his recipiency or to initiate actions that can be easily integrated within a multi-unit turn (such as initiating repair sequences) or to derail it [such as starting oblique sequences by asking questions (Koenig, 2007)]. Nods, head shakes and assessments can be appropriate responses to the development of a multi-unit turn and display an understanding of the ongoing talk (C. Goodwin, 1986b; C. Goodwin & Goodwin, 1987, 1992; Kendon, 2002) and can be used to convey affiliation with the teller's stance toward the events (Stivers, 2008).

The transition relevance place of each TCU is therefore a point where speakers can see if the other is playing their role as a recipient. For these reasons, this work could focus on what happens to gaze during the transition relevance place that follows each TCU of a multi-unit turn. However, in examples 2.1 and 2.2, the recipient is already gazing and continues

gazing to the speaker if what is ongoing is a telling, and this is confirmed in many other examples in the corpus. If the recipient is already gazing at the speaker and nothing really changes, it is difficult to make the case that it is at that moment that the recipient understands that the other participant is producing a telling. Moreover, there is interactional evidence that the recipient must arrive at this analysis much earlier. If this were not the case we would often observe overlapping talk after the first TCU of a multi-unit turn, as the recipient would start a turn and take the floor only to realize that an extended telling was in progress. We rarely observe this sort of interactional collision. Rather, the timing of continuers, assessments and nods suggests that recipients understand early in a TCU whether it is projecting the beginning of a multi-unit turn, and behave as recipients, at least by not taking the floor.

All this points to the TCU that projects the extended telling, rather than the first transition relevance place, as a locus of order for further investigation. Moreover, we have already noted that it is very early in extended tellings that the recipients look up in examples 2.1 and 2.2. Therefore the focus should be on gaze behavior at the beginning of extended tellings to see whether the regularities in gaze behavior can be accounted for.

In sections 2.2.2 and 2.2.3 I provisionally identified, through qualitative analysis, different recipient gaze behavior projectable extended tellings vs. adjacency-pair-based sequences. I have then further refined the domain of investigation suggesting that the focus for ETS should be on their beginnings, and in particular in the first TCU. This seems to be the place in which recipients start looking up at the speaker (if they were not already looking) and when they start to sustain gaze toward her/him.

As can be seen in Figure 2.1 and 2.2, the completion of the TCU that projects an extended telling vs. the completion of a first pair part of an adjacency-pair based sequence have different interactional implications. At the completion of the first TCU of a multi-unit turn (to pick one of the ways in which an extended telling can be produced), other participants withhold taking the floor, orienting toward the bid by the speaker to produce further talk. By contrast, the completion of the TCU that initiates an adjacency pair sequence constitutes a transition relevance place at which talk by another participant is not only allowed but typically due.
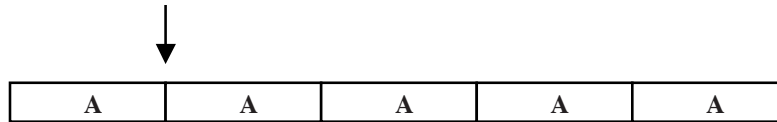
**Extended telling**



*Figure 2.1* Allocation of TCUs after a TCU projecting an ETS
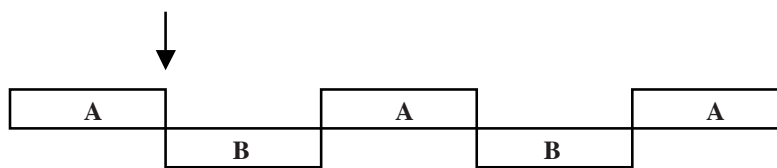
**Adjacency-pair-based sequence**



*Figure 2.2* Allocation of TCUs after the first pair part of an APBS

In the following section, I provide further qualitative evidence of how participants distinguish these interactional places through their gaze and body behavior. While further illustrating the empirical evidence of the claim here presented, the following examples also provide some initial clues in terms of which features participants rely on to be able to produce these behaviors systematically.

## 2.2.5 Further Qualitative Evidence

This section uses two extended examples to illustrate how recipients' gaze behavior functions as a display of their understanding of what kind of talk will follow (ETS or APBS). The recipient is the same in both instances and in both cases the interactants are looking at pictures together. The goal is to further demonstrate that gaze behavior differs when everything else but the sequential context remains the same.

Example 2.5 shows the beginning of an extended telling (a story, in this case) in which the recipient looks up toward the speaker and shifts his posture from a prior orientation to a competing activity to being oriented toward the teller. This example is from an interaction between two male friends who are looking at pictures together. B (blue and white shirt) is visiting A (white shirt) to help A fix some problems on his computer. A has just returned from his summer holidays and he is showing B some of his pictures. Before the

beginning of this extract B told a short story triggered by a picture. He described his shock and concern when he woke up and heard an alarm coming from a big chemical factory. Line 1 represents the end of that telling and line 3 A's reaction to line 1.

**(2.5) 2PCOMP_rosso 10:59**

```
01 B:   Si' pero' era un falso allarme
        Yes but   was a  false alarm
        Yes but it was a false alarm
02      (0.5)
03 A:   A:h
        Oh
        O:h
```

04 B:   `.h E   quel giorno  che: (0.3)`
            And that day       that
        **.h And the day in whi:ch (0.3)**

   **\*=** B raises head from pictures toward A but looks mid distance
      and not to A

```
05 B:   [c'era  tutto rosso quando mi son svegliato la mattina
         cl.was all   red   when   me woke.1s up    the morning
        [It was all red when I woke up in the morning
```



```
06      [
```

```
07 B:   Ci siam presi paura perche' pensavamo  che
        cl.   got.1p fear  because thought.1p that
        we got scared because we thought that
```

```
                              (A)↔(B)

08 B: ⎛fosse successo qualcosa=nube  [tossica ⎞
      was    happened something cloud toxic
      something had happened=[toxic cloud

                                    (A)↔(B)

09 A:                               ⎛[atomica
                                     atomic
                                     [atomic
```
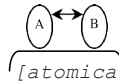
After the completion of the previous telling and the receipt of the last bit by A (line 1-3), B starts another telling. He does it with a short inbreath, the use of the conjunctive "e" (and) and a temporal reference "quel giorno che" (the day in which). After a short pause at line 4 he adds a description "c'era tutto rosso" (it was all red) and a more specific time in which this event occurred "quando mi sono svegliato la mattina" (when I woke up in the morning) which becomes the beginning of a telling. The telling continues beyond line 9. Here I represent only the very beginning.[12] The focus is what happens in terms of gaze at line 4. At line 3 both participants are oriented toward the pictures in front of them. However, at line 4, B starts a turn of talk and after the word "quel" (that) re-orients his head up more toward A, although his gaze is in the middle distance and not on A. Soon after the beginning of B's head movement, on the last syllable of the word "giorno" (day) A moves his head and gaze up toward B and then sustains his gaze toward B throughout the telling. Thus, B appears to be orienting to A's turn as initiating a new multi-unit turn. Additional evidence that he is oriented not just to the fact that B is talking but to what B is doing with his talk comes from the shift in posture performed by A during the silence at the end of line 4 and at the beginning of line 5. He goes from being oriented toward the pictures to being completely oriented toward B: he leans his head toward his right hand, while his right elbow is on the table (see picture at line 6). This shift in posture occurs at the same moment that B shifts his gaze toward A but also at a point in which A's recipiency has only been displayed through gaze and not through any nodding or verbal continuers. As Schegloff (1998), has shown, the orientation of our head and torso is usually related to the direction of our attention and in particular it becomes a display of our willingness to engage in other activities. A's shift in body posture at that point becomes an index of his disengagement from the activity of looking at pictures and his engagement in recipiency of B's ETS. Not only that, by shifting

---

[12] In all the following examples of extended tellings, only the beginning part will be shown.

his posture entirely rather than torquing his body, A displays his understanding that this new activity (listening to B's talk) will not be short and therefore requires a rearrangement of his body in relation to the projected action. My claim is that the recipient of the telling (A) finds sufficient evidence in ".h e quel giorno" in line 4 to project that another telling will be produced, in a sequential context in which a telling from B had just occurred. A indeed shifts his gaze first and then his posture accordingly and acts as a recipient. It might seem quite bold to claim that ".h e quel giorno" would be enough to project a telling, but the previous telling by B (shown below as example 2.6), produced few seconds earlier, started with a similar phrase: "Pensavo a Porto Marghera. Quella volta che mi sono svegliato e si sentiva…" (I was thinking about Porto Marghera. The time in which I woke up and you could hear…).

**(2.6) 2PCOMP_Porto 10:49**

```
01     (1.0)
02 B:  No pensavo  a     Porto Marghera
       No think 1S about Porto Marghera
       No I was thinking about Porto Marghera
03     (.)
04 B:  hm quella volta che mi son svegliato e   si sentiva iu:: iu::
       hm that   time that me  woke up      and cl. hear   iu   iu
       hm the time in which I woke up and you could hear iu:: iu::
05     ini[ziai ((S fa gesto tipo "avere paura" see Fig. 2.1))
       started 1s
       I sta[rted ((B makes emblematic gesture of "being scared" Fig. 2.3))
06 A:      [Ah si'?
            Oh yes
           [Oh really?
```



Figure 2.3. Emblematic gesture meaning "being scared"

Note the similarities between "quella volta che" and "e quel giorno che" (which is the expression he uses in example 2.5). A can be expected to use this resource to recognize this as a telling. A has just heard B using the former as a way of starting a telling. Moreover the turn initial "e" (and),[13] produced at a place in which the previous telling is over, suggests either an expansion of the previous telling or another similar situation of waking up scared: what the two stories have in common. Finally, tellings frequently contain a time reference in their beginnings and this often provides a claim about the "tellability" and "news" of what will follow (Sacks, 1986). By offering a time reference a speaker not only specifies when something happened, but also accentuates, on one hand, that the events narrated happened in a recognizable moment in time and the access the speaker had to the events narrated (imagine saying "last Monday" versus "thousands of years ago") and, on the other hand, it projects that what is going to be told is not something about the present, but it is located in some other time when the recipient was likely not with the speaker.

There are therefore potentially multiple features in the beginning of this TCU that help the recipient to project that an ETS is likely to be forthcoming. Moreover, there is a double display of orientation toward the beginning of the telling, first through gaze and then through a shift in posture. Both occur in the context of a competing activity that heavily relies on the participants' gaze: looking at pictures. This shows that even in the context of an ongoing engagement in a competing activity, the competing activity is abandoned and the recipient orients toward the extended telling as an activity having priority over the other activity. In my data set, participants are fast and effective in orienting their gaze toward the teller as soon as they can project a telling, even if this requires momentarily abandoning a competing activity. This suggests that a further analysis of the situations in which competing activities get prioritized over displaying recipiency to the telling is needed. It could emerge that sustaining an orientation toward a competing activity is a marked type of action and a way of displaying what kind of recipiency a participant is going to provide.

Let's see now what recipients' gaze behavior looks like in the alternative sequential environment: the initiating turn of an APBS. In the following example the participants are the same as the ones in example 2.5. They are looking at pictures, and B, the person producing the initiating action (a request for information), is the same individual who started the telling in example 2.5. Thus, everything seems to be the same, apart from the content and action

---

[13] For similar claims in terms of the function of turn-initial "and" see Van Dijk (1979) on the pragmatic aspects of the use of connectives and Heritage and Sorjonen (1994) on *and*-prefaced questions and their claim that "and" invokes not simply a relationship between facts, but rather a relationship between the current turn and its interactional setting.

performed through the first TCU of B's talk. Our focus is the recipient's gaze behavior at line 9.

**(2.7) 2PCOMP_pagato 9:33**
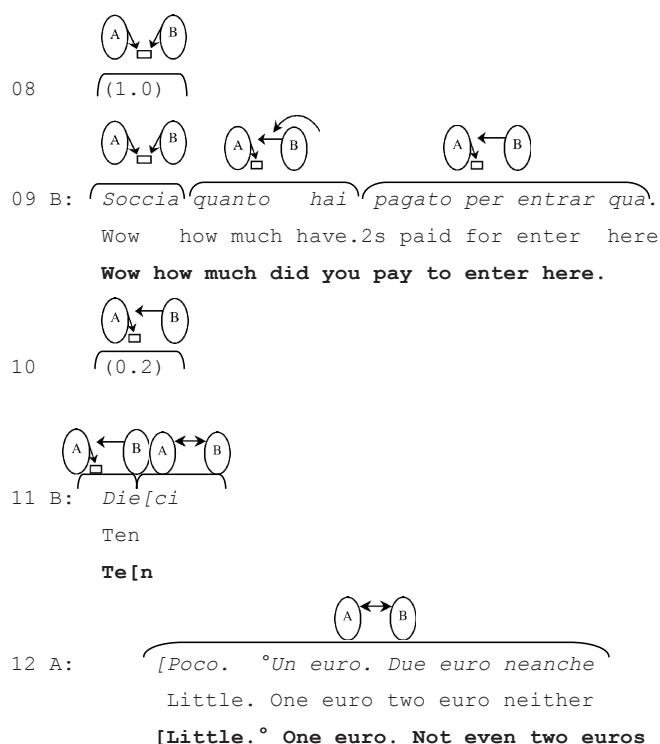
```
           A ↘ ↙ B
08      ⌐(1.0)⌐

           A ↘ ↙ B    A ←B     A ← B
09 B:   ⌐Soccia⌐quanto   hai⌐pagato per entrar qua⌐.
           Wow   how much have.2s paid for enter   here
        Wow how much did you pay to enter here.

           A ← B
10      ⌐(0.2)⌐

        A ← B A ↔ B
11 B:   Die[ci
           Ten
        Te[n

                      A ↔ B
12 A:       [Poco.   °Un euro. Due euro neanche
             Little. One euro two euro neither
        [Little.° One euro. Not even two euros
```

At the beginning of this extract both participants are oriented toward the pictures on the table (see line 8). At line 9 B first produces an assessment that sounds like an outloud (Goffman, 1981c) "soccia" (wow), most likely about what he is seeing in the picture (the dome of St. Peter's church in Rome). Then he asks a question about the cost of a visit to the dome. That request for information is our focus. The recipient does not look at the speaker at all during the turn's production. The speaker shifts his head orientation and gaze from the pictures to A while producing the beginning of line 9. This head movement is formally identical to the one described in example 2.5 during the beginning of the ETS. The only difference is that this time B looks at A when he ends the head movement while in the previous example he had a middle-distance gaze before looking at A. In example 2.5 the timing of B's head movement and A's shift in gaze orientation may have pushed for an understanding of the shift as a reaction to the head movement. This could mean that in example 2.5 A looked at B at that

76

moment not because of what was happening in the talk, but rather because of what was visible in B's shift in head orientation. Example 2.7 shows that B's head movement is not, per se, prompting A to look up. At line 9 B turns his head and looks at A while A keeps looking at the picture. Moreover, A continues looking down through the following silence and looks at B only when he starts to answer at line 12. It could be said that A continues looking at the picture during the gap because of the indexical "qua" (here) at the end of B's turn and because B's index finger is on the picture. On the other hand, the index finger has been there since the beginning of line 9. A has been looking at that picture since before the beginning of line 9 and the speaker is staring at A during that turn. Finally, the absence of speech disfluencies in B's turn notwithstanding (e.g., silences or phrasal breaks), B's gaze toward the recipient and the absence of recipient's gaze suggests that the recipient's gaze is not normatively required or that Goodwin's rule (1980, 1981) can be relaxed here because of the competing activity of looking at pictures.

A comparison of example 2.5 and example 2.7 shows a clear difference between the recipient orientation toward the speaker during the first TCU of a telling vs. the first TCU of a sequence initiating action. In these examples the participants are the same, the competing activities they are involved in are the same and so is the environment where this interaction occurs. Moreover, in both examples the participants are looking down before the focus TCU and in both cases the speaker turns toward the recipient in the beginning of the TCU. The different behavior observed in these two examples, however, could be accounted for in terms of different understandings and orientations toward what is happening in the talk and what activities those TCUs begin. Recipients treat the beginning of a projected extended telling as a position where their gaze should properly be directed at the speaker whereas they treat beginnings of APBSs as not making gaze at speaker relevant.

The two alternative hypotheses explaining recipients' gaze toward the speaker presented throughout this chapter can be alternatively summed up as follows: recipients look toward speakers at the beginning of an ETS because

a) someone is talking and the participant role of recipient requires looking toward the speaker most of the time
b) the speaker's TCU is in a particular sequential environment and is therefore initiating a particular activity (i.e. s/he will make particular social demands for the recipient, including producing appropriate responses at the right time)

Hypothesis a) has been proposed by prior literature on the topic, while hypothesis b) has been advanced during the course of this chapter (initially in section § 2.2.3) and then supported by qualitative evidence. The examples observed in this section point toward two additional issues that need to be addressed in the remainder of this chapter:

1. Is this behavior affected by whether the participants are involved in competing activities (e.g. cooking, eating, beading)?

2. How do we account for those ETSs in the data set in which the recipient does not look toward the speaker before completion of the first TCU?

In what follows I provide a negative answer to the first question and will deal with the second in section 2.3.4.

## 2.2.6 Competing Activities and Their Relation to the Distribution of Gaze

In the previous examples we saw that sometimes people are involved in competing activities while participating in a conversation. This work does not assume that the participants' gaze orientation toward other objects is a sufficient explanation for their not gazing at speakers while being recipients. If people only looked at the speaker during the beginning of a telling when they were not involved in any other competing activity, this would suggest that recipients only sometimes look up and keep looking, namely when they have nothing else to look at. This is clearly not the case. Participants are constantly involved in other activities throughout the data set analyzed here and yet they systematically react to the beginnings of extended tellings in the same ways. In examples 2.5 and 2.7 I have shown that if a participant is involved in the same competing activity but the sequential environment of the conversation changes, that participant shows a different orientation toward the talk and as such her/his gaze behavior is systematically different. To stress this point and to deal with the alternative account that uses orientation toward other objects or activities as a reasonable, equally valid, option for the participants during a conversation, I show two additional examples in which the gaze behavior changes while the competing activity and the participants remain the same.

In the following example two male friends are chatting facing each other. B has asked A to tell him how to prepare some recipes, as he will have to cook for 60 people for several days and he is worried about it. Our focus here is line 20, where B begins a telling.

**(2.8) 2PPLAN_il problema 06:28**

```
16      (0.2)
```

A ← B

```
17 B:  Cosa¿
       What
       What¿
```

A ← B

```
18     (0.6)
```

A ← B              A ↓ B

```
19 A:  F(h)ate    tr(h)anquillamente    voglio dire non e' che (tu)
       Make2p     relax                 want   say  not is  that you
       You can d(h)o it ea(h)sily       I mean it is not that (you)
```

A ↓ B       A ← B    A ↔ B    A ↔ B

```
20 B:  Ho capito. Pero' il problema e' che   (0.1)    (0.4)
       Have.1s understood but the problem is that
       I understood. But the problem is that (0.1)  (0.4)
```

A → B                   A ↔ B

```
21     uno   da  quel che  ho    capito che e' quello che   arriva
       one   from that which have.1s understood which is that which arrives
       one   from what I understood that is the one that arrives
```

A → B          A → B          A ↔ B          A → B

```
22     lunedi' (0.2) se ne intende abbast- somma  m'ha chiamato  lui e
       Monday         it know.3s enough  well   me have.3s called he and
       Monday (0.2) knows about it enou-well he called me and
```

After B's open class repair initiation (Drew, 1997) at line 17 and A's repair at line 19, B starts his turn at line 20 claiming he understood A's insistence on not worrying about the amount of cooking. Then he starts another TCU in which he counters B's minimization of the difficulties of the event, by explaining what the issues are (fundamentally there are only two people cooking for the first day, including himself, and he is not much of a cook). A, during lines 16-20, is looking down, first at a piece of paper with the menu and then at some keys he has picked up from the table. After the words "il problema e' che" (the problem is that) the recipient looks up toward the speaker and sustains his gaze through the beginning of the

telling, as seen also in example 2.5. What we see then is that A does not look at B while listening to the repair initiation (line 17) or doing the repair (line 19), nor during the first TCU of B's turn at line 20, when the latter stresses that he understood A's point. Rather, A looks up when B has projected that some explanation or details about why B is still worried will be delivered. In other words at the beginning of a projectable extended telling.

Similarly, see example 2.9, which is taken from the same interaction and where A is again the recipient. They are still talking about recipes and A has started going through the menu in front of him to see how he can help B. The menu he is looking at is from the year before and this is why he is using the future tense throughout this extract. The target in this extract is the turn at line 5-6 where B suggests what he is going to do.

**(2.9) 2PPLAN-prendo 08:45**

```
01  A:  Allora gli spaghetti alla carbonara sicuramente ci saranno
        Okay   the spaghetti      carbonara certainly   cl. will be
        Okay spaghetti carbonara definitely will be there

02      per[che' li   mettono] s[empre]
        because them put       always
        beca[use they put them in] eve[ry time]

03  B:      [Che schifo]      [Mangiano lo]ro
             How disgusting     Eat.3p   they
            [How disgusting]   [They eat it]

04      (0.6)

05  B:  .hhh Allora facciamo cosi' io prendo fuori della
             Okay   do.1p    so   I  take.1s out    some
        .hhh Okay let's do this I'll get some

06      carta [eh¿ intanto
        paper  eh  in the meanwhile
        paper [eh¿ in the meanwhile
```

80

```
                                          Ⓐ↘□ □↙Ⓑ
07  A:              [Il condimento al tonno e' sempre quello:::
                     the dressing  at tuna  is always that
                   [The tuna dressing is always the same:::
08          [quindi    lo faranno] sicuramente
             therefore  it do.3p   certainly
           [so they will make it] for sure
09  B:     [Prendo   la  penna]
            Take.1s the pen
           [I'll get the pen]
```

A is looking at the menu and speculating about the probability of having the same food this year as the year before at lines 1-2. B says how much he dislikes spaghetti carbonara and then suggests what the two of them could do starting from what he is going to do: pick up a piece of paper to annotate A's cooking advice. While producing lines 5-6, B reaches for the papers underneath the menu, therefore entering A's visual space. At line 9, B says that he is going to take the pen and the pen is again beside the menu, within A's visual space. But A neither reacts nor looks at B. His entrance into A's visual space and touching objects close to his focus of attention do not lead him to look up toward B.

   If we then look more closely at the action at lines 5-6 we see that it is a proposal and at the same time B says that he will be reaching for the paper (under the menu A is looking at) and therefore will be intruding into A's space. The use of the tag "eh" with slightly rising intonation displays that B is looking for a go ahead by A (Sacks, 1986), which he does not get, and yet B nonetheless proceeds in picking up the paper. The turn at lines 5-6 is also proposing, indirectly, that A interrupts what he is doing and that they start looking at each item in the menu in a more detailed way, so that B can make some notes. So it proposes a re-start of the advice about the cooking. Finally, by detailing what he is going to do after having started his turn with a plural "facciamo cosi'" (let's do this), B is inviting A to infer the part concerning what A should do if B gets some paper and, as such, to recognize what "I'll get some paper in the meanwhile" means for him. What matters for us is that the TCUs in lines 5-6 constitute sequence-initiating actions and A does not look toward B. Lines 7-8, similar to lines 1-2, are outlouds and A stops talking immediately after the end of this stretch of talk, letting B start the advice seeking activity with "allora partiamo" (Okay, let's start) just one second after line 9.

In the prior example (2.8), we see that while the recipient is looking at objects on the table and therefore is clearly engaged in competing activities, he nevertheless looks up during the first TCU of a projectable extended telling. By contrast, he continues looking down during the sequence initiating action in example 2.9. Therefore, it is not simply that the recipient does not gaze toward the addressee because he is looking at something else. Rather, depending on what is going on through the talk, the relevance and appropriateness of a visual orientation toward the speaker is more likely to make the participant disengage from the competing activity or not. We saw this in examples 2.5 and 2.7 as well. This evidence supports the claim that gaze behavior in interaction displays priorities and recognition of relevance to specific sequential types of talk.

Goodwin (1984) proposes that a way to understand why gazing away from the speaker during a telling is often not sanctioned or problematic is that there is a relaxation of the gaze rule. However, the notion of "relaxation of a rule" that would otherwise require the recipient to gaze at a looking speaker, does not sufficiently explain the interactional negotiation of local priorities and accountability of contingent choices, but rather takes for granted that competing activities at times get prioritized over visual engagement in a conversation. For recipients, looking at the speaker may not always be relevant, but sometimes may be. We should see an orientation by the participants toward the relevance for the recipient to look toward the speaker at the beginning of an ETS, no matter what other activities are going on. My claim that recipients tend to look at speakers during tellings does not mean that the recipient never looks away or glances at other objects in the environment. This actually happens quite often. For example, before grasping an object, or putting food on a spoon, or moving a cup toward their lips, participants normally look at the objects they are going to move. As said in chapter 1, this is related to visual control of actions (e.g. hand movement) and to a perceptual function of gaze. In studies that examine where people fixate when they perform routine activities on their own (see Hayhoe & Ballard, 2005 for a detailed review), it has been found that participants do not focus on the most visually salient object, but on what is more important for the spatio-temporal demands of the job they are doing. They do not look at irrelevant objects, but do look at the relevant objects for the ongoing activity (e.g., a glass if they are going to drink) before reaching for them. Thus, fixations are linked in time to the progression of the task s/he wants to accomplish. This simply means that in order to grab a cup from the table I look at it before reaching for it (and I can monitor my hand grip while I am going to grab it). This is not necessarily to display to the other participants that I want to grab the cup but because the accomplishment of the task requires

coordination of my muscles in relation to spatial and temporal information. These glances will be brief and indeed related to the accomplishment of the task. If I am making a cup of tea, I have to look at the cup, at the tea bag and at the water boiler at certain points throughout the process. However, if I am simultaneously a recipient of an extended telling, I will display through gaze and body orientation that I understand and recognize that I should be looking at the speaker, by alternating glances toward the relevant objects and glances toward the speaker. Just continuing to look at the tea bags or at the cup should be an accountable action and not simply a reasonable alternative.

Example 2.10 shows how a recipient of a telling (A) manages handling a competing activity (stirring the pasta in a pot) and attending to a report about B's vacations. The participants here are two male roommates in their 20s and they are preparing lunch together. Notice in the three pictures how A (person on the left) first looks toward the pot with the pasta in Figure 2.4, but then keeps his body and gaze orientation toward the teller while using his stretched right arm to stir the pasta in the pot (Figures 2.5 and 2.6). They are already in the midst of a telling when this extract begins.

**(2.10) 2PRON1-mare 16:50**

```
01      (2.0)
```



```
        Figure 2.4
```

```
02 B:  No era bello perche' dalla catena di sopra e' come(Fig.2.4) se fos-
       No was nice  because from  chain  of above is like        if was
       No it was nice because from the chain on top it is as (Fig.2.4) if-
```

```
      Figure 2.5
03    quasi  come (Fig.2.5) la Liguria vedi   direttamente il  mare_
      almost like            Liguria see.2s directly     the sea
      almost like (Fig. 2.5) Liguria you (can) see directly the sea_
04       (0.2)
05 A: mm hm
      mm hm
      mm hm
06       (0.4)
```



```
      Figure 2.6
07 B: (Fig.2.6) Quando sei    li'   sopra…
                When   are.2s there above
      (Fig.2.6) When you are up there…
```

The telling continues after line 7. This example shows how a participant might go out of his way to display sustained attention toward the telling even when a competing activity might require his attention as well. First he looks toward the object he has to reach for, to control distances and his reach, then, when found, he re-orients toward the teller and displays attentiveness with his whole body posture and gaze.

84

This section has shown that participants in interaction discriminate their involvement in competing activities in relation to what the sequential environment of the talk is.

## 2.3 Generalizing The Claim

It has been noted above that recipients' gaze behavior during sequences of talk can differ between a first turn and a second turn and that recipients who were not already looking at the speaker often do not look up during the TCU that initiates an adjacency pair sequence. Nonetheless, it has not been established yet how generalizable these observations are and to what extent the pattern described at the beginning of ETSs holds. For this reason, a quantitative comparison of the following domains was developed: recipients' gaze behavior during the first TCU of an ETS vs. recipients' gaze during the first TCU of an APBS.

Before proceeding with a quantitative comparison, I will show how my work takes into account a couple of general problems that emerge from prior literature on the topic.

### 2.3.1 How These Observations Fit (or not) with Prior Literature

In the preceding sections I have provided qualitative evidence that participants' gaze behavior is influenced by the sequential organization of the talk and that the pattern described by Kendon and many others since only applies to participants involved in a multi-unit turn (as far as recipient's behavior is concerned). This points to two general problems with prior work in this area outside of the conversation analytic tradition:

1. Insufficient analyzed data, both in terms of the amount of minutes and in terms of the number of different participants and their diversity
2. Insufficient control of the sequential environment of the talk and of the actions performed through talk.

Most research in this domain is based on the analyses of 3 to 5 minutes of video recordings from each interaction. In my corpus, a single telling can take that length of time. Therefore some behaviors may not have been observed because of the limited amount of data analyzed.

An additional complication is that we do not know the specific actions performed through the talk occurring in most of the previous research. In several studies the experimental task the participants were performing is described merely as "having a conversation" (e.g., Argyle & Graham, 1976; Kendon, 1967). In others the task for the participants is explicitly to tell (and listen to) a story (Bavelas et al., 2002) or to listen, wearing an eye-tracker, to somebody telling a story in order to be able to retell it to the experimenter (e.g., Gullberg & Holmqvist, 2006). In the latter two studies, the tasks and the pattern found (recipient looking at speaker's face most of the time) confirm the initial observations made in this chapter that when somebody produces a projectable multi-unit turn of talk (such as a storytelling) the recipient will start looking at the speaker very early and will sustain the gaze throughout it. For the other studies it is impossible to dismiss the possible presence of reasonably long tellings that would explain the misleading generalization that what matters in terms of participants' gaze behavior is only (or at least mainly) their participation status.

This chapter combines a qualitative and a quantitative approach to the analysis of gaze behavior. The analysis relies on data from 10 dyadic interactions (10 minutes from each) and 20 different participants (see Appendix A for further details about these interactions and their participants). These analyses demonstrate that the sequential environment of the talk is a critical organizational feature of gaze behavior. The following sections outline some of the findings about extended tellings in the conversation analytic literature.

## 2.3.2 Conversation Analytic Literature on Extended Tellings

There has been a significant amount of work in Conversation Analysis on the beginning of extended tellings and how opportunities for their delivery emerge throughout a conversation. Sacks (1974; 1992 [1964-72]: 222-228) showed that stories are sequential objects that articulate with the conversational context in which they occur. Stories can begin with story preface sequences, minimally composed of two turns. Prefaces usually project the launching of a telling and usually involve a characterization of it that will help the recipient recognize when the telling is over and how to react to it. Jefferson (1978) studied how stories can either be elicited or solicited, how they are locally occasioned by being either triggered in the course of turn-by-turn talk or can be methodically introduced into turn-by-turn talk and how these tellings can be sequentially implicative for the talk that follows their completion.

Marjorie Goodwin (1982) described how "instigating stories" begin (i.e. stories meant to create conflict by reporting what a non-present person said about the current recipient). Among their common features, the principal character of the story is a non-present person and the story reports actions performed by this party that can be perceived as offensive for the recipient. Charles Goodwin (1986a) showed how diversity in the audience to a story, in particular with respect to the different experiences and relationship that recipients might have with the events being talked about, can affect the way a story is told. Lerner (1992) described how the production of a telling can be "assisted" by other participants. He showed that there are three main ways of assisting the initiation of a story: by prompting it, by provoking it or by producing a reminiscence recognition sequence (solicit + recognition). Moreover who will deliver the story is usually interactionally arranged in the course of the preface.

C. Goodwin (1984) focused also on the relationship between gaze behavior and the production of a telling, though he did not analyze specifically what recipients of the telling do with their eyes at the very beginning of the telling. He argues that apart from displaying engagement and attention, recipient gaze fundamentally follows the rule described in some of his previous works (C. Goodwin, 1980, 1981) according to which if the speaker is looking at the recipient then the recipient should be looking back. He also suggests that in some sequential positions the rule can be 'relaxed' though not ignored so that the participant can attend to competing activities. In particular, he says that a non-gazing recipient may nod in different sequential positions compared to when the recipient is gazing, to display an orientation toward the possibility that looking away is a dispreferred activity. Goodwin also suggests that the body behavior of the storyteller, in particular assuming particular postures and looking toward the recipient, can be considered signals toward the recipient that a telling is about to start. However, apart from Sacks' work on the verbal beginning of tellings and Goodwin's single example of a "storytelling posture" by the incipient teller, none of these works explain in detail which features make clearly recognizable or projectable, from the very beginning of these extended tellings, that an extended telling will be delivered, and whether recipients display in any way their understanding of it before reaching the first TRP. Later in this chapter I will attempt to fill this gap.

In what follows I provide a quantitative comparison of the participants' gaze behavior in the two sequential environments above mentioned.

*2.3.3 Looking Up during First TCU*

Prior sections illustrated a difference in the organization of gaze in two sequential environments. In this section I show that recipients systematically look toward the speaker at the beginning of an ETS, whereas they do not systematically do so at the beginning of an APBS.

**2.3.3.1 Beginning of Tellings vs. Beginning of Adjacency-Pair-Based Sequences**

Two types of evidence show that recipients look toward speakers in response to beginnings of ETS.

First, recipients not only look up toward speakers, they also sustain their gaze through the transition relevance place and well into the following TCUs. In the 35 instances analyzed here it is not just that recipients look up quickly and then look away or toward something else before the completion of the TCU. Rather, they engage in the pattern previously described of sustained gaze toward the speaker. In this way they display that the looking up represents the beginning of that pattern, and therefore a display of recognition and understanding that they are/will be the recipient of a multi-unit turn.

Second, recipients behave differently in adjacency pairs from the way they behave in extended tellings. If recipients look up simply because they hear someone speaking then there should be no difference. If there is a systematic difference, then it likely lies in the sequential structure of the talk. This would also confirm the observation made in connection with examples 2.1 and 2.2 and what distinguished them from examples 2.3 and 2.4.

Let us look closer at the second type of evidence and focus on what recipients do with their gaze during the first recognizable TCU of an extended telling. The data analyzed here consist of 100 minutes of interaction involving a total of 20 different participants. I randomly selected 10 minutes from 10 different dyadic interactions and then coded for their content. As previously described, these interactions were not experimentally controlled and the participants were free to say and do anything they wanted, including producing extended tellings. I collected all the instances of extended tellings in the corpus without paying any attention to gaze behavior. These data yielded 68 instances of extended tellings. I then further reduced the collection to the extended tellings that started with the recipient not already looking at the speaker in order to assess whether something was happening to recipient's gaze direction during the first TCU of the telling. The instances were therefore reduced to

38.[14] Which TCU would count as the first one was determined *post hoc*, once the entire extended telling had been selected for the collection of 68 previously mentioned. The empirical question is whether there is a change in the recipient gaze direction during that TCU and whether the recipient starts looking toward the face of the speaker.

I then went back to the same corpus and collected all the TCUs that initiate an APBS (i.e. all the actions that could initiate a new sequence, and not ones that would constitute expansions of an ongoing one). This time the collection yielded 135 cases. These TCUs implement many different actions including requests for information, noticings, directives and announcements. Again, in order to observe differences in the gaze behavior of the recipient of that TCU, I reduced the collection to only instances in which the recipient was not already looking at the speaker before the beginning of the TCU. There were 103 remaining instances. The question then is: do recipients look up toward the speaker as they did during the first TCU of ETSs?

Table 2.1 shows the contrast between the two sequential environments and Figure 2.7 depicts this contrast graphically.

*Table 2.1* Recipient gaze behavior during the first TCU of ETS and first TCU of APBS

| | Extended-Telling Sequences (ETS) | Adjacency-Pair-Based Sequences (APBS) |
|---|---|---|
| **Recipient gazes at Speaker** | 35 (92%) | 43 (42%) |
| **Recipient does not gaze at Speaker** | 3 (8%) | 60 (58%) |
| **Totals** | 38 (100%) | 103 (100%) |

---

[14] The fact that in 30 of the 68 instances the recipient was already looking at the speaker does not constitute a problem for the claim here. Most of the tellings that began with the recipient already looking at the speaker were produced in second position, either as second stories or in response to questions by the other participant, or even as explanations of something they had just mentioned. This means that the other participant was likely oriented toward the speaker because the course of action was not possibly complete at that point, even if a telling was not necessarily projectable. This supports the findings described in chapter 4 about the withdrawal of gaze being related to an analysis of possible completion of the sequence of actions.
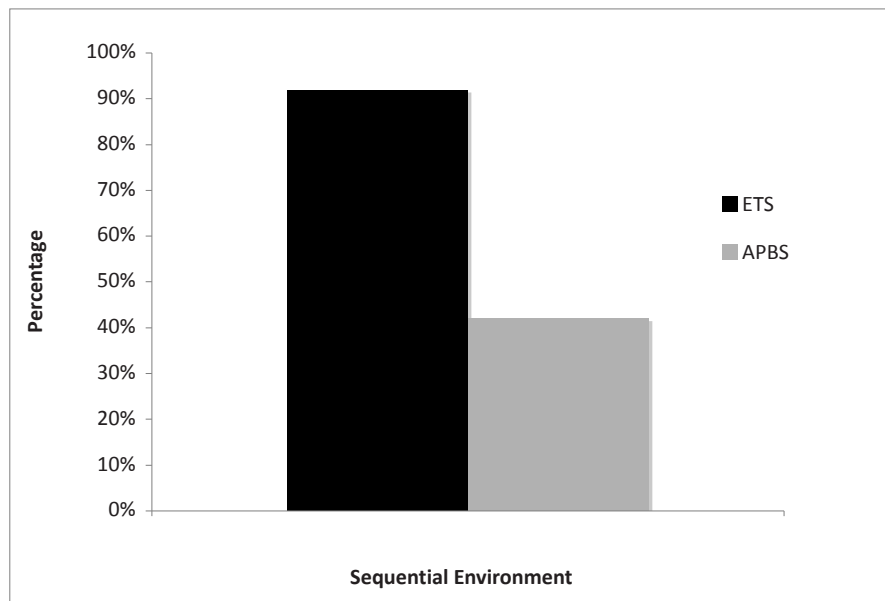
*Figure 2.7* Percentage of first TCUs by sequence in which the recipient turns gaze toward speaker.

In 92% of all cases recipients bring their gaze to the speakers once they hear that an ETS is projectably beginning. This distribution clearly contrasts with the one displayed in Table 2.1 for APBS. Here the recipient looks at the speaker in the minority of the cases (only in 42% of them), suggesting that participants can systematically differentiate their gaze between a TCU that constitutes the beginning of an ETS and a TCU that constitutes the beginning of an APBS before that TCU reaches completion.

In order to establish whether the observed difference was statistically significant, the relationship between gazing at speaker and the sequential environment of the TCU was tested using a logistic regression. This further allowed for a correction of the standard error for the clustering of sequences in the 10 interactions.

*Table 2.2* Results of Logistic Regression predicting Recipient Gaze at Sequential environment.[15]

| Variables | Odds Ratio | 95% Confidence Interval |
|---|---|---|
| Extended Telling | 16.28*** | 5.05, 52.45 |

The results of the logistic regression are shown in Table 2.2 as an odds ratio with the 95% confidence interval. They show that if the TCU is the first one of a projectable ETS then

---

[15] *** denotes p<.001

90

the odds that the recipient will gaze at the speaker during that TCU go up 16.28 times in contrast with TCUs representing the beginning of an APBS. The results were adjusted for interactions to monitor whether what happens in one particular interaction would mainly drive this result in the direction just presented, but this is not the case. Thus, the difference in gaze behavior is generalizable and systematic with respect to sequential environment. These results contrast with the general claim of what I called Kendon's gaze patterns and instead confirm the initial observation discussed with respect to examples 2.1 and 2.2 vs. examples 2.3 and 2.4: recipient gaze behavior in projectable ETS is different from recipient gaze during the first TCU of an APBS. These results also show that recipients usually engage in sustained gaze toward the speaker within the first TCU of the ETS. They do not have to wait until a transition relevance place.

There is, then, a systematic practice implemented by recipients of extended tellings, namely looking toward the speaker as soon as they detect the beginning of a telling and sustaining this orientation throughout the telling. This appears to be due to affordances and requirements of extended turns of talk and not simply to their being recipients of any turn of talk.

The distributional evidence provided here shows that it is a replicable pattern. It shows that recipients treat gaze to speakers as something that they should do in ETSs but not necessarily in APBSs. The next section addresses the issue of the normativity of recipient's gaze behavior in extended tellings by presenting a deviant case.

## 2.3.4 A Deviant Case

In documenting the interactional practice of recipients looking toward speakers' faces during the first TCU of a telling there were 3 cases (out of 38) where this did not occur. In what follows I discuss how we can account for one of these cases[16] and how these three cases add to our understanding of this practice. Example 2.11 is one of the three cases. As will be shown, the recipient does not look at the addressee during what I initially labeled as the first TCU of the extended telling. Upon closer analysis though that TCU is equivocal as to whether it is closing the prior telling or initiating a new telling. The fact that the TCU is

---

[16] Of the two remaining examples, in one case the extended telling occurs without being projected, simply by having the speaker adding TCUs that end up constituting a telling while the recipient neither responds nor looks at the speaker. In the last case the recipient is trying to assemble a plastic rose and looks intently toward the objects in front of her saying outloud "Okay wait I need," while the speaker starts the telling (i.e., she was not paying attention to the speaker of the ETS and considered the competing activity priority).

pivotal could explain the gaze behavior of the recipient. The participants in this example are the same as in examples 2.1 and 2.3. Here they are talking about a friend of theirs who is considered to be very insensitive. The focus is the recipient's gaze at lines 6 and 10.

**(2.11) 2GSTUDYING_trattore 37:47**

```
01 A:   Proprio u- un trattore
        Exactly a- a  tractor
        Really a- a tractor
02      (0.7)
03 A:   mm- mm- va   dritta   per la sua strada .hh
        mm- mm- goes straight for   her way
        mm- mm- She does not deviate from her way .hh
```

```
04      Poi  non ha  piu' parlato assolutamente
        Then not has more talked  at all
        Then she did not talk anymore at all
```
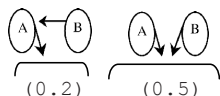
```
05      (0.2)    (0.5)
           (0.7)
```

```
06 A:   Idem  [quando la Lea] si e' lasciata    con  M(h)atteh
        Ditto  when     Lea   cl.had broken up with Matte
        Ditto [when Lea] broke up with M(h)atteh
```

```
07 B:        [(A cioe')]
              I mean
             [(I mean)]
```

```
08      (0.5)
```

```
09 A:   .hhh  che  se- c'ero  anch'io pero'   quella sera
        .hhh  that if- cl.was also I  but     that   evening
        .hhh  that if- I was there too however that evening
```

92

```
         ⓐ↓  ⓑ↓   ⓐ←ⓑ↰        ⓐ←ⓑ
10   siamo usciti tutti  insieme (0.8) e   la Lea
     are   exit   all    together      and the Lea
     we went out all together (0.8) and Lea
                              ⓐ←ⓑ
11   gliel'   ha detto mi sono lasciata con  Matte .h
     her that has said me am   left      with Matte
     told her I broke up with Matte .h
```
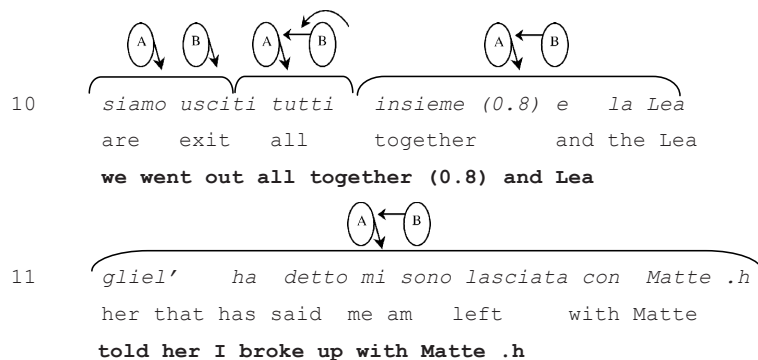
Line 4 is the last part of the previous telling, where A was talking about the day in which their friend C had behaved "like a tractor", being extremely cold and unfriendly when she heard that A had broken up with her boyfriend. The talk at line 4 refers to the silence following C's cold reply to A's bad news. Whereas the recipient keeps looking at the speaker into the following silence, the speaker has already lowered her gaze before the completion of the TCU. During the silence the recipient lowers her gaze as well and both participants seem oriented toward the possible completion of the telling. Then, at line 6, A produces the possible pivot turn. On the one hand, it could be heard as a story preface in which the speaker announces that the main point of the story will be describing how C did the same thing to Lea. On the other hand, the use of the term "idem" (ditto) suggests the possibility that the turn is a re-completion of the telling, stressing that what happened to A was not simply a unique case but rather something common, as it already happened with another friend of theirs. In this sense, line 6 could be further supporting evidence for the choice of the extreme characterization of C as a "tractor". B is drinking some iced tea from a glass and she is not looking at A while doing it. The additional evidence that B has not projected the production of an extended telling at that point comes from her behavior during line 9 and the beginning of 10, when she puts the glass down on the table on her right and she looks at it while A is talking. At this point, indeed, line 9 could be understood as another pivot TCU. The TCU to be taken as one unit (because of its prosodic contour) is "c'ero anch'io pero' quella sera" (however I was there too that evening). On the one hand it provides additional evidence concerning the reliability of the information at line 6 (A was there when this happened). On the other hand, it provides additional epistemic evidence of her knowledge of what she is reporting. So line 9 could be either looking backwards at what had just been said or looking forward to what has not been said yet. At line 10, after the first two syllables of the

word "usciti" (went out) B starts turning her head and her gaze toward A and from then on looks at A throughout the telling, including the 0.8 seconds of silence following "tutti insieme" (all together) at line 10.

Once A has said that something similar happened to another person they both know (and that A was there and witnessed it) the clause "we went out" becomes hearable as part of a telling. The fact that the event is located in the past, the choice of words such as "sera" (evening), and the fact that A was also present all create a semantic compound that makes plausible an understanding of "we went out" as an event that is related to what has been mentioned in the previous two TCUs. Moreover, the absence of any specification of the subject in Italian (there is no "we" but first person plural is specified in the conjugation of the verb and it is not clear who could be the "we", apart from A, C and Lea) adds to the relevance of linking the initiating utterance with what preceded. Even just focusing on the choice of the verb "to go out", we see that it projects a description of other things that happened during that evening, as "going out" represents only the first step in the possible schema of what normally happens when people meet in the evening. It is also a piece of information that contributes no crucial additional evidence to the negative characterization of C, so it cannot be perceived as a third piece of the re-completion of the telling. All these features allow B to recognize what is going on at that point and to look up and sustain her gaze toward the speaker, as happens in all the previous examples of multi-unit tellings.

A closer look at this deviant case has therefore shown that the recipient is not really deviating from the otherwise common gaze pattern, but rather she is analyzing the conversation on-line and interpreting a pivot turn as a possible end of the previous telling, until she has reason to interpret the next piece as actually a piece of a new telling. At that point the recipient gaze pattern observed in previous examples is immediately engaged, therefore confirming the systematicity of the practice.

This shows that deviant cases of the gaze practice described in this chapter can be reconciled with the normativity of the practice because participants' behavior is explainable in terms of interactional contingencies. In what follows I show how this orderly gaze behavior is practically implemented and I show that the findings that emerge can tell us something about human cognition and behavior parsing.

## 2.4 On Projectability

In the sections above I have shown that recipients look at speakers early in an extended telling. In this section I show how recipients can detect an extended telling in advance. In particular I show how recipients' gaze becomes a possible key to access their understanding of what is going on in the interaction. A list of features will be presented that should account for how participants can project which kind of action will be performed and whether something is going to be an extended telling or not.[17] Finally, some implications for research on decision-making, social cognition and heuristics will be discussed.

### 2.4.1 Projecting an ETS

The participants in these data look up toward the speakers of extended tellings systematically at the very beginning of the telling. How can a recipient recognize an extended telling right at the beginning, i.e. in the first TCU? I want to propose and test four[18] candidate features:

1. The speaker is gazing at recipient
2. The sentence type of the TCU
3. The action performed by the TCU
4. Lexical choice in the TCUs

These four features are not mutually exclusive and participants very likely rely on more than one of them, or even all of them at the same time in order to recognize ETSs. This analysis will suggest that some of these features seem to matter more than others, without

---

[17] Projection in social interaction is crucial not only for the fast unfolding of turn-taking, but also for the capacity of producing a sequentially appropriate next action while the first is still underway (i.e., in overlap). A few scholars have investigated this mechanism not from a psychological point of view but from an interactional one. As Mondada (2006: 118) summarized: "projections can be observed within prosodical, syntactic (Auer, 2005), turn constructional (Schegloff, 1996c; Selting, 2000; Ford, 2004), sequential (Drew, 1995), and gestural (Streeck, 1995) organizational practices."

[18] It has been suggested that a fifth feature should be the posture the speaker of an ETS assumes before launching into a telling, as suggested by Goodwin (1984). However, examples 2.5 and 2.7 show that often there is absolutely no difference between the posture a speaker assumes before launching a telling and the posture a speaker assumes before initiating an APBS. A further look at the instances of ETS and APBS this dataset show that speakers do not reliably assume one or more specific posture for ETS, clearly systematically distinguishable from the ones they assume for APBS. Moreover, often there is no shift in posture by the speaker at all before the beginning of an ETS.

completely rejecting some additional influence of the others. In presenting the analysis performed on each of the four features above mentioned, I will present fragments from examples previously shown in this chapter (or new ones, in case no example previously shown contained the dimension of interest). As such, the numeration of the examples refers to the number that the example had when it was first presented within this chapter.[19]

Let us begin by returning to our current knowledge of the database of ETSs and APBSs previously examined. Table 2.3 shows that there are 38 instances of beginnings of ETSs and 103 instances of first pair parts of APBSs that start with the recipient not already looking at the speaker. In 92% of the first TCUs of ETSs the recipient starts looking toward the speaker, while this happens in only 42% of the first TCUs of first pair parts of APBSs.

*Table 2.3* Current knowledge about the data set

| Sequential Environment | Number of instances | Recipient starts to look at Speaker |
|---|---|---|
| ETS | 38 | 92% |
| APBS | 103 | 42% |

The first feature –speaker gaze toward recipient during the TCU– is proposed as a candidate given Goodwin's claim (1980, 1981, 2000b, 2006) that speaker and recipient gaze could be considered as two parts of an adjacency pair where the presence of the first (speaker's gaze toward recipient) makes conditionally relevant the occurrence of the second (recipient's gaze toward speaker). It is possible that recipients look up more often during the beginning of extended tellings because the speaker tends to look at them in those situations. Example 2.8 could be supporting evidence for this claim.

**(2.8) 2PPLAN_il problema 06:28**



20 B: *Ho capito. Pero' il problema e' che (0.1) (0.4)*
      Understood but   the problem is that

      **I understood. But the problem is that (0.1)   (0.4)**

---
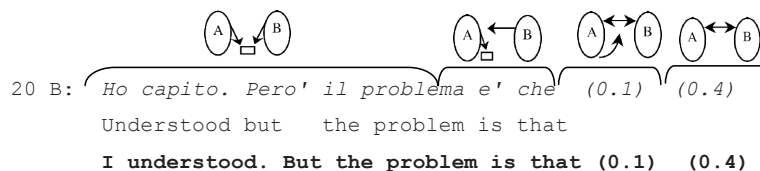
[19] Some of the extracts presented in chapter 2 will be presented again in the following chapters, to show that multiple practices can be deployed one after the other on the same conversational fragment. This will also show that potentially any extract from a dyadic conversation might show evidence of the various practices described in this work.

In this example the speaker looks toward the recipient before the latter looks up. However, there are many cases (e.g. 2.1, 2.2, 2.5, 2.11) in which this is not the case. Let's look for example at the beginning of the extended telling shown earlier in example 2.2.

**(2.2) 2GSTUDYING_mamma 29:36**



```
05 B:   Mi  fa   mia  mamma hai    conosciuto dei  bei      ragazzi? Si'
        Me does my   mum   have.2s known      some handsome boys     Yes
        My mum says to me have you met some handsome boys? Yes
```
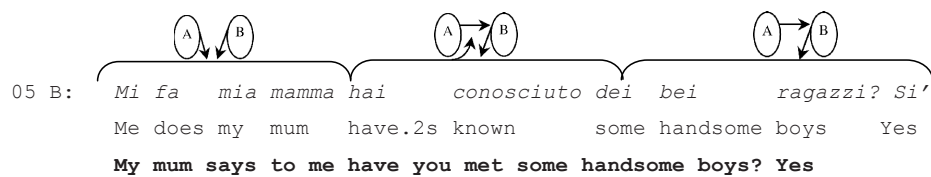
In this example the speaker never looks at the recipient while the recipient looks toward the speaker early in the TCU. The qualitative analysis shows that it is unlikely that speakers' gaze is systematically making relevant recipients gaze toward them, but a more quantitative approach is needed here. I therefore looked at participant's gaze behavior across the database. Tables 2.4 to 2.7 show that it is unlikely that speaker's gaze toward the recipient is the driving force of recipient's gaze behavior. Although speakers tend to look up toward the recipient a bit more during the first TCU of an extended telling (60.5%) than during the first TCU of an APBS (43.7%), the number of instances in which the speaker is not gazing and the recipient nonetheless looks up toward the speaker are too high to support such an analysis as a primary account (see Tables 2.5 and 2.7). Indeed, in 34.2% of the cases of extended tellings and in 14.5% of the cases of sequence initiating actions the recipient looks toward the speaker even if the speaker is not looking.[20] The percentage of TCUs in which there is no gaze by any participant in APBS is also remarkable. As shown in Table 2.7, 41.7% of sequence initiating turns (APBS) are produced without any gaze toward each other by any of the participants and even more remarkable is the percentage of TCUs in which there is no mutual gaze during the course of the entire first TCU (72.8% for APBS and 42.1% for ETS). These percentages counter the claim by Goodwin that mutual gaze is the default during face-to-face interactions and that not engaging in mutual gaze during a turn would be accountable or problematic.

---

[20] Additional evidence against a Goodwinian account is that in my data speakers who do not have recipient's gaze do not usually produce cut off, phrasal break or pause to secure the recipient's gaze before the completion of the turn. Specific evidence for this point is provided in Rossano et al. (2009) and in the following chapters of this dissertation.

*Table 2.4* Gaze behavior during first TCU of ETS.

| Participant | Gaze at other participant | Gaze away from other participant |
|---|---|---|
| Speaker | 23 (60.5%) | 15 (39.5%) |
| Recipient | 35 (92.1%) | 3 (7.9%) |

*Table 2.5* Instances of no mutual gaze in first TCU of ETS, divided by gaze configuration.
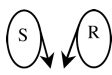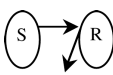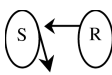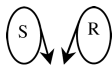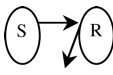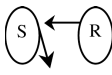
| (S→ R↓) | (S→R↓) | (S←R↓) | Total no Mutual Gaze |
|---|---|---|---|
| 2 (5.3%) | 1 (2.6%) | 13 (34.2%) | 16 (42.1%) |

*Table 2.6* Gaze behavior during first TCU of APBS.

| Participant | Gaze at other participant | Gaze away from other participant |
|---|---|---|
| Speaker | 45 (43.7%) | 58 (56.3%) |
| Recipient | 43 (41.7%) | 60 (58.3%) |

*Table 2.7* Instances of no mutual gaze in first TCU of APBS, divided by gaze configuration.

| (S→ R↓) | (S→R↓) | (S←R↓) | Total no Mutual Gaze |
|---|---|---|---|
| 43 (41.7%) | 17 (16.5%) | 15 (14.5%) | 75 (72.8%) |

The second feature I proposed was sentence type. Relying on sentence types for determining whether the recipient should be looking at the speaker means relying fundamentally on specific morpho-syntactic features of the TCU. I considered the distribution of three sentence types: interrogative, declarative, imperative. We might expect that, for instance, declaratives would be more likely to be associated with the recipient's gaze because tellings often start with turns in a declarative format, while interrogatives would be more likely associated with first pair parts in APBSs and therefore less likely to get the recipient's gaze.

Example 2.7 shows a TCU with interrogative morpho-syntax (it is a wh-question):

**(2.7) 2PCOMP_pagato 9:33**

```
09 B:  ⌐Soccia⌐quanto    hai⌐pagato per entrar qua.
        Wow    how much have paid  for enter   here
      Wow how much did you pay to enter here.
```
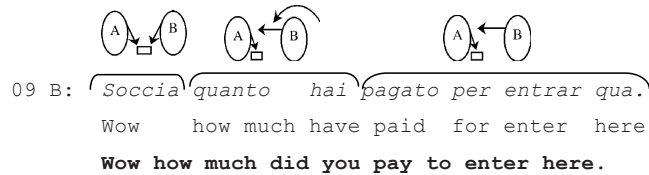
In this example the speaker produces a first pair part by asking how much the recipient paid to enter the dome of Saint Peter's church.

Example 2.4 is another example of sequence initiating turn but this time the TCU has declarative morpho-syntax.

**(2.4) 2GSTUDYING_testa  29:07**

```
09 A:  .h No io veramente son scarburatissima ho     un mal di testa
          No I  really     am  out of steam     have.1s a bad of head
      .h No I am really out of steam I have a headache
```
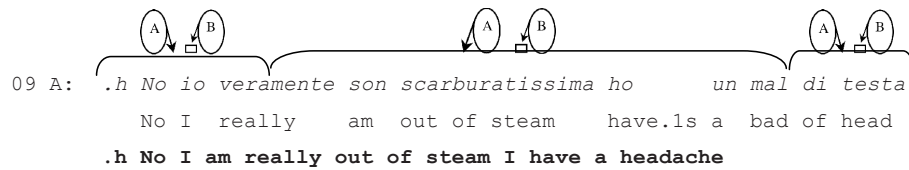
Our target here is the TCU "io veramente sono scarburatissima" (I am really out of steam) and it is in a declarative form. Example 2.12 is instead an example of imperative morpho-syntax. The participants are two girls sitting opposite each other eating a piece of cake.

**(2.12) 2GC-pezzo 28:00**

```
01 B:    [Dai    prendi  un  altro pezzo
         Come on take.2s one other piece
         [Come on Take another piece
02 A:  N:o io davvero
        No I   really
       N:o I really
```
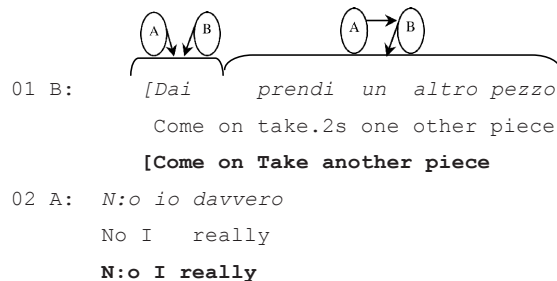
At line 1 B tells A to take another piece of the cake that is on the table in front of them, but A refuses. Line 1 is another example of a sequence initiating turn but this time with imperative morpho-syntax.

If a specific sentence type is more strongly associated with the beginning of extended tellings rather than with sequence initiating turns we could consider the sentence type the driving force of this specific gaze behavior. Table 2.8 shows the distribution of sentence types by sequential environment.

*Table 2.8* Distribution of instances of sequence-initial TCUs belonging to a specific sentence type

| Sentence Type | Extended Telling Sequence | Adjacency Pair Based Sequence |
|---|---|---|
| Interrogatives | 2 (5.3%) | 44 (42.7%) |
| Declaratives | 36 (94.7%) | 46 (44.7%) |
| Imperatives | 0 | 13 (12.6%) |
| **Total** | 38 (100%) | 103 (100%) |

There are two things to note about the distribution of sentence types by sequential environment: first, in this corpus imperatives are only used to initiate a sequence of talk and not to initiate tellings.[21] Second, in this corpus declaratives are the main way in which participants start a telling while they are used a bit less than half of the time to initiate a sequence.

Table 2.8 also shows that there is a striking similarity between the percentage of declarative sentences at the beginning of extended tellings (94.7%) and in first pair parts of adjacency-pair-based sequences (44.7%) and the percentage of instances in which a recipient looks up at the beginning of an ETS (92.1%) and at the beginning of an APBS (41.7%). Could sentence type be the feature that affects gaze behavior during a conversation? We can check this by examining whether recipients tend to look up only (or mainly) during declaratives when they are in APBS. This is shown in Table 2.9:

---

[21] In theory it is possible to project a story telling by producing an imperative: "listen to this!"

*Table 2.9* Instances of gaze toward speaker by recipient divided by sentence type

| Sentence type | Extended Telling Sequence | Adjacency Pair Based Sequence |
|---|---|---|
| Interrogatives | 2/2 (100%) | 20/44 (45.5%) |
| Declaratives | 33/36 (91.7%) | 20/46 (43.5%) |
| Imperatives | 0 | 3/13 (23%) |

Table 2.9 shows that the sentence type of the TCU is not (at least alone) driving the recipient looking up toward the addressee. Indeed, the three instances in which the recipient does not look up at the beginning of a telling all occur with a declarative TCU, but more importantly, of the 43 total instances in which a recipient looks up during a sequence initiating turn only 43.5% are in declarative forms. In APBSs the looking up by the recipient is quite evenly split between interrogatives and declaratives, though the percentages are only half of what we see for the same sentence types when deployed at the beginning of an extended telling. In other words, one might assume the rule to be "if declarative, suspect the occurrence of an extended telling", but if this were the case, then we should see that at least 90% of the time the TCU under examination is in a declarative format the recipient looks at the speaker. What we see, instead, is that when the declarative starts a telling, recipients look toward the speaker almost always, while when the declarative represents the beginning of an adjacency-pair-based sequence, then the recipient looks up toward the speaker only 43% of the time. This means that participants must use some other cue to distinguish among declaratives and what they actually project. Therefore, recipients do not seem to treat the sentence type of the first TCU as critical to whether they should look toward the speaker or not.

The third feature under examination is <u>action type</u>. The actions that will be considered here are not the actions or course of actions that the entire telling is performing, such as storytelling or doing a report, but rather what kind of action the first TCU of the telling is implementing, for example announcing or assessing. The following examples represent some of the action types that can be found in the data. Example 2.1 "una volta io ho aumentato" (once I have raised the price) illustrates an announcement.

**(2.1) 2GSOFA_aumentato 23:32**

```
02 A:  pero' in effetti non saprei   neanch'io co-<↑una volta
       but   in effect  not know.1s  neither  I  wh- one time
       but actually I would not even know wh-<↑once
```

```
03     io ho    aumentato
       I  have  raised
       I have raised ((the price))
```
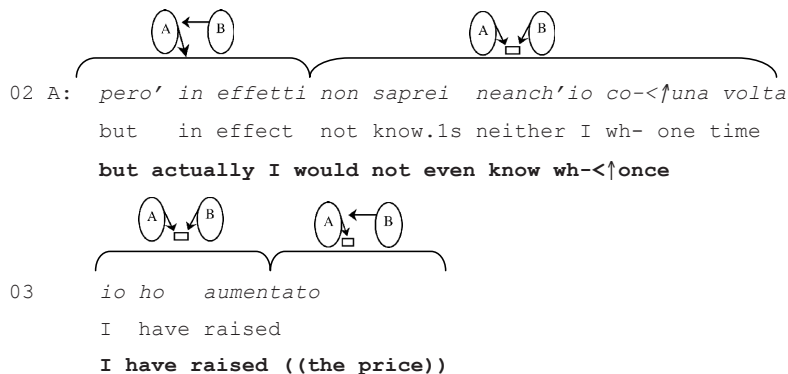
Before line 2, B had mentioned that she finds it difficult to ask for more money for her babysitter job and A had said that B has to ask for more anyway, although A does not know either how to ask for more money. Then, A cuts off her turn and says that she actually did it once. By reporting what she has done in first position (so not having been asked if she had ever done it) she frames this TCU as an announcement for B: here is something she has done that she has not told B about yet.

Example 2.3 is a request for information.

**(2.3) 2GSOFA_come   25:16**

```
07 B:  Come t'  e' venuto in mente
       How  you is come    in mind
       How come you thought about it
```
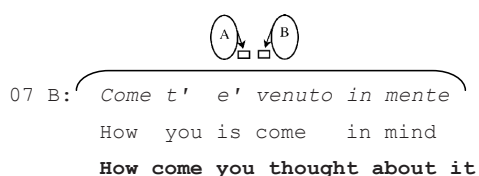
Example 2.8 is the beginning of an explanation.

**(2.8) 2PPLAN_il problema 06:28**

```
20 B:  Ho capito. Pero' il problema e' che  (0.1)  (0.4)
       Understood but    the problem is that
       I understood. But the problem is that (0.1)  (0.4)
```

```
21  uno    da quel che ho capito che e' quello che   arriva
    one    from that which understood which is that which arrives
    one    from what I understood that is the one that arrives
```
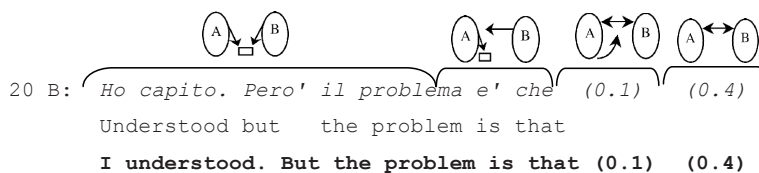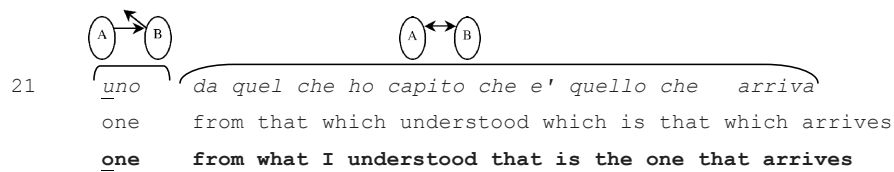
Example 2.13 is an assessment (a negative one) about a girl who is said to behave like a "bulldozer" when interacting with her friends.

**(2.11) 2GSTUDYING_trattore 37:47**



```
07      (1.5)
```



```
08  A:  HUUU A volte  e' un c a t e r p i l l a r
        huuu at times is a  bulldozer
        HUUU Sometimes (she) is a bulldozer
```
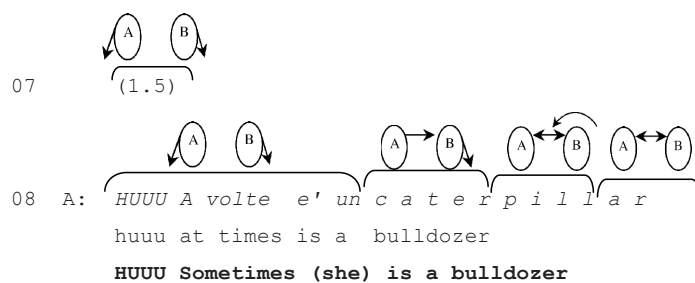
For recipients, to rely on the feature of action means that the recipient should look at the speaker once a certain action is recognizable. If we find that recipients looking up toward speakers strongly correlates with specific types of actions and that those specific types of actions are mainly occurring in one sequential environment rather than the other, we can account for the different gaze patterns with the actions a speaker is implementing through a TCU.

Table 2.10 and 2.11 show the distribution of action types[22] among the first TCUs of ETSs and APBSs, respectively, and show how often the recipient looks at the speaker divided by action type.

---

[22] I have consolidated the action types around general categories of action such as reporting, assessing, announcing, etc. I kept track of whether the recipient's gaze seems to behave similarly among the actions of one cluster. By this I mean that if the recipients look at speakers half of the time in one small group and the same happens in another one that I have labeled as two types of reports I clustered them as reports.

*Table 2.10* Action types and gaze behavior in first TCU of ETS

| Action Type | Number of Instances | Recipient Gaze toward Speaker |
|---|---|---|
| Announcement | 19 | 17 (89.5%) |
| Assessment | 2 | 2 (100%) |
| Report | 8 | 7 (87.5%) |
| Explanations | 7 | 7 (100%) |
| Other | 2 | 2 (100%) |

*Table 2.11* Action types and gaze behavior in first TCU of APBS

| Action Type | Number of Instances | Recipient Gaze toward Speaker |
|---|---|---|
| Announcement | 11 | 6 (54.5%) |
| Assessment | 21 | 11 (52.3%) |
| Report | 11 | 4 (36.4%) |
| Directive | 11 | 2 (18.2%) |
| Offer | 8 | 3 (37.5%) |
| Request | 36 | 14 (38.9%) |
| Explanation | 3 | 1 (33.3%) |
| Other | 2 | 1 (50%) |

Table 2.10 shows that the main way in which the extended tellings in my dataset get started is through 'announcements', which represent 50% of the entire instances. Table 2.11 shows that the most represented type for sequence initiating turns is the action type 'request'. On the other hand, while in extended tellings the recipient looks toward the speaker independently of the action performed, in the case of adjacency-pair-based sequences only announcements and assessments get the recipient to look at the speaker for slightly more than 50% of the cases. This is especially relevant for those actions such as announcements, assessments, reports and explanations that we also find in extended tellings and that in that environment are usually associated with the recipient looking toward the speaker. None of the action types found in the sequence initiating turns have a correlation with the occurrence of recipient gaze that goes above chance. It therefore seems reasonable to state that action types are not (by themselves) a driving force guiding the recipient gaze behavior.

104

Before proceeding to the next and last feature, I would like to sum up what we have seen in this section. The occurrence of the speaker's gaze, the sentence type and the action type of the TCU have been shown not sufficient alone to predict the recipient's gaze toward the speaker. This does not mean, though, that they are irrelevant. Indeed, from each analysis and numerical description we can pick one or two elements worth considering for further analyses in this domain. For instance, that the speaker seems to be looking at the recipient more at the beginning of an ETS than at the beginning of an APBS could be displaying a different orientation by speakers toward these sequential environments and the actions performed there. It is not the main cue, but it could be relevant and certainly worth exploring with a larger data set.

The fact that most of the ETSs start with declaratives and at the same time only half of the APBSs start with declaratives could be worth exploring further as well. Moreover, the fact that ETSs do not usually start with imperatives could be an additional hint for the recipient who is trying to project what the speaker is going to say and do.

Finally, the fact that certain action types tend not to occur in both sequential environments could be another important negative cue for recipients. If a recipient hears a request, it is unlikely that this is going to be the beginning of an extended telling and as such it could be less relevant for her/him to look at the speaker. Announcements, then, represent the kind of action that most frequently attracts recipients' gaze in the beginning of APBS and announcements are also the main action type through which ETSs tend to begin. Moreover, directives (and therefore imperative sentence types) less frequently attract a recipient's gaze. The specific action that the first TCU of a sequence is performing may not be the sole cue, but it may account for part of the observed pattern. There is, however, an issue that we need to consider to make sense of the next step in this analysis: how are turns of talk constructed and designed?

This is not the place to initiate an appropriate analysis of turn design[23] and action formation[24] but it seems clear that there are many features that are recurrent and maybe

---

[23] It is clear that a turn is designed in a certain way to implement specific actions, therefore the two are related.
[24] For a recent summary of what is known about action formation and recognition, see Levinson (in press). Though mainly limited to the investigations performed by few individuals, conversation analysts have produced extensive work on the formation and recognition of a few social actions, such as compliments (e.g., Golato, 2005; Pomerantz, 1978), complaints (e.g., Drew, 1998; Drew & Holt, 1988; Drew & Walker, 2009; Monzoni, 2008), assessments (Goodwin & Goodwin, 1987, 1992; Pomerantz, 1984a), requests (Curl & Drew, 2008; Drew & Walker, 2010), invitations (Davidson, 1984, 1990; Drew, 1984).

unique to specific types of actions or specific clusters of actions.[25] There is empirical evidence that suggests that not just the verb but other word types or practices (e.g., pitch resets) or semantic domains could help to project the occurrence of an ETS rather than an APBS: sometimes the recipient looks toward the speaker just 1 or 2 words into the TCU and sustains the gaze afterwards. How can recipients project the beginning of an extended telling with such a limited amount of information?

To answer this question, Gerd Gigerenzer and his collaborators' approach to the notion of heuristics is relevant. With the term heuristics they refer to mechanisms by which humans make decisions. Rather than considering these mechanisms to be imperfect versions of optimal statistical procedures, they treat them as adaptive strategies deployed by humans and animals to make inferences about the world under limited time and knowledge. In a recent article, Hutchinson and Gigerenzer (2005) propose a closer collaboration with biologists to develop better models of heuristic algorithms based on the descriptive findings of biologists. In particular they stress that two of their most famous heuristics, the Take the Best heuristic and the Recognition heuristic could be valuably used to understand animal and human behavior.[26] The *Take the Best* heuristic (Gigerenzer & Goldstein, 1996) proposes that having to make a decision given different cues, a participant takes the cue that discriminates better between two alternatives and has a positive cue value and should ignore the remaining cue(s). However, if the cue examined does not discriminate, then the algorithm goes back to the cue values and searches for another one that discriminates better between the options. The *Recognition* heuristic (Goldstein & Gigerenzer, 2002) proposes that given two or more options, if we cannot recognize a specific option, it is very likely to be a less important one (e.g., which is the biggest city: Chemnitz or Berlin?). People would normally opt for the one they can recognize (in this case, Berlin).

This focus on recognizability resonates with a famous quote by Harvey Sacks in which he defined culture as "an apparatus for generating *recognizable* actions" (Sacks 1992 [1964-72]: 226, Vol. I). Gigerenzer's attention toward simple mechanisms that can explain how people decide one way or another resonates with the concerns of this chapter and so does his stance toward heuristics as adaptive strategies, as answers to the problem of not having unlimited time and computational power. Do any of the above-mentioned heuristics

---

[25] See Austin's (e.g., 1962) or Searle's (e.g., 1970) work on speech acts and their suggestion to use the verb type to determine which action is an utterance performing.
[26] See Enfield (2009a) for a recent attempt to utilize heuristics to describe human behavior in social interaction.

work once applied to recipient's gaze behavior in ETSs? In what follows I address this question first by presenting the possible cues at play.

The following distributional evidence shows that certain turn design features are relevant for helping recipients project that an extended telling is in progress. The fact that participants often look toward the speaker of an extended telling very early in a TCU suggests they could use the presence of particular words, and their semantic domains, as possible cues to project what will come next and what sort of uptake will be relevant – recipiency or answering/acting etc. If this is the case, we should see different turn design features at the beginning of a telling than in a question or other sequence initial actions. This would imply that just hearing one or more specific words in the first TCU would be enough to determine whether it is relevant to direct my gaze toward the speaker. For instance, if a speaker uses a deictic and a point close to the place where her/his recipient's eyes are already oriented, the latter will typically not look up at the speaker's face but rather look in the direction indicated by the pointing finger.

In these data there are various turn design features occurring immediately before the recipient looks toward the speaker. I identified six features[27] of turn design that were recurrently present in first TCUs of ETSs and four features commonly present in the design of the first TCU of APBS.

---

[27] The features here listed are only six because only six features could be found in more than 10% of the first TCUs of ETSs in my data. It seems plausible that more than six could be found and described. For example, something such as "pitch resets" might attract recipient's gaze and might be used to initiate extended tellings, to distinguish them from previous turn-by-turn talk. A larger database and work developed in other languages might shed more light in this direction and provide further features to add to the list.
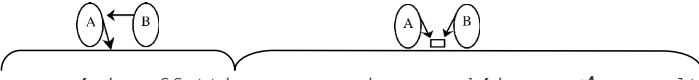
*Table 2.12* Possible heuristics in initial TCUs.

| ETS | APBS |
|---|---|
| 1) First person pronoun or first person inflected in the verb | a) Second person pronoun or second person inflected in the verb |
| 2) Time references (usually about past or future but not about the present) | b) Deictics (verbal and/or visible) |
| 3) Initial mentioning of third persons | c) "Ma" (resumption marker) |
| 4) Verbs that display epistemic access to a certain fact or event | d) Modals (want, should, could, etc.) |
| 5) Mentioning of actions' first step in a schema of a larger event or activity | |
| 6) Extreme characterizations of individuals, places or facts | |

Before discussing how recurrent these features are in the data and therefore their significance as cues, let us examine some examples of the items that usually appear at the beginning of extended tellings. The first design feature listed for extended tellings in Table 2.12 is the use of the first person pronoun "I" or "we". The beginnings of extended tellings often contain references to the first person singular pronoun "I", while first pair parts of APBSs are more likely to contain references to the second person singular pronoun "you". In a pro-drop language such as Italian the first and second person are often only inflected in the verb, but in the sequential environments we are comparing it often happens that the pronoun is explicitly produced. Apart from the presence of the pronoun, it is clear that in many beginnings of ETSs the speaker says something about herself/himself, while in many beginnings of APBSs the speaker says something about the addressee, or at least puts her/him as subject of the main verb. Most storytellings and reports are about the speaker, or something that the speaker has experienced, seen, heard, etc. On the other hand, when one asks someone something, it is likely that the second person pronoun will be relevant.

Example 2.1 shows an instance of a first TCU where the recipient looks up toward the speaker and it contains a first person pronoun

**(2.1) 2GSOFA_aumentato 23:32**
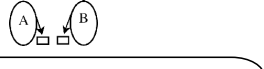
```
02 A:  pero' in effetti non saprei   neanch'io co-<↑una volta
       but   in effect  not know.1s  neither I  wh- one time
       but actually I would not even know wh-<↑once
```

```
03     io ho    aumentato
       I  have  raised
       I have raised ((the price))
```

In this example the recipient starts looking up after the production of the first pronoun "io" (I) and during the production of the verb "aumentare" (to raise).

Example 2.3 contains an example of second person pronoun in a TCU in which there is no gaze up by the recipient. In this example neither of the participants looks at the other.

**(2.3) 2GSOFA_come   25:16**

```
07 B:  Come t'  e' venuto in mente
       How  you is come    in mind
       How come you thought about it
```

Another recurrent design feature at the beginning of extended tellings is the use of temporal expressions that locate the events at a time different from the present, and as such potentially in a time when the recipient was not present. These can range from "the day" to "the first time she entered my house", for example. It should be noted that not just any time reference seems to occur at the beginnings of extended tellings but rather ones that refer to recurrent times (often, sometimes, etc) or to the past or to the future but not to the current situation at the time in which the conversation unfolds. People do not usually start a telling by saying "adesso" (now) though they can say 'today' in the sense that it refers either to a time prior to the moment of speaking or to a time that follows the completion of the current conversation.

Example 2.5 contains an initial time reference to which the recipient reacts by raising his gaze.

**(2.5) 2PCOMP_rosso 10:59**

```
04 B:    .h E    quel giorno   che: (0.3)
         And that day          that
         And the day in whi:ch (0.3)
     *= B raises head from pictures toward A but looks mid distance
        and not to A
```

Here the recipient looks up after the words "quel giorno" (that day), which refer to a specific day that will be identified only by what happened to the speaker. It is a non-recognitional time reference for the recipient.

A third turn design feature is the initial mentioning of third persons. This sensitivity to new characters resonates with an observation made a-propos of Chekhov's plays. In a Chekhov play, if the opening curtain reveals a gun hanging over the fireplace, you can be sure it will be shot before the final curtain falls. In the same way, if all of a sudden a new individual is introduced into the conversation, very likely something important will be said about this individual or about something this individual has experienced or done. Example 2.2 is a case in point.

**(2.2) 2GSTUDYING_mamma 29:36**

```
05 B:   Mi fa   mia mamma hai    conosciuto dei  bei    ragazzi? Si'
        Me does my  mum    have.2s known      some handsome boys    Yes
        My mum says to me have you met some handsome boys? Yes
```

In this example the recipient starts moving her eyes toward the speaker immediately after the production of the person reference "mia mamma" (my mum). Clearly related to it is a fourth feature of turn design: the use of verbs that provide information about the epistemic access the characters had to some event or news like 'seeing', 'hearing', 'knowing', 'believing', etc. In the following example the two participants are talking about preparation for an exam and participant A has asked B whether B has notes that he can borrow. B says that he does not

110

have notes and after 1.1 second of silence (line 1) he starts an ETS by saying "ho visto che" (I have seen that). He will then continue explaining how he knows about where other notes are and what happened to him when he went looking for them.

**(2.13) 2PEXAM-visto 47:28**

```
01    (1.1)

              Ⓐ  Ⓑ        Ⓐ→Ⓑ        Ⓐ→Ⓑ
02 B:  Ho    visto che  ce ne  sono in giro.  Al Dobbi anche
       Have.1s seen  that cl them are  in around at Dobbi also
       I have seen that there are ((notes)) around. At Dobbi also
```

In example 2.13 we can see that the recipient shifts his gaze toward the speaker after the latter has produced a verb of seeing that displays his epistemic access to what he is going to say next. The speaker produces line 2 while looking toward a computer screen and does not shift his head orientation during the first TCU.

Fifth, the choice of verbs such as "going out" projects a telling about what happened once the people were out because of its being a first item in a schema of what happens when people go out together, as in the following example. The two participants are talking about a mutual friend who seems to be particularly insensitive in terms of love matters. Example 2.11 shows the lines that exemplify the first action in a schema.

**(2.11) 2GSTUDYING_trattore 37:47**

```
           Ⓐ Ⓑ                Ⓐ  Ⓑ
09 A:  .hhh  che  se- c'ero  anch'io pero'   quella sera
       .hhh  that if- cl.was also I  but      that    evening
       .hhh  that if- I was there too however that evening

           Ⓐ Ⓑ  Ⓐ←Ⓑ        Ⓐ←Ⓑ
10     siamo usciti tutti  insieme (0.8) e  la Lea
       are  exit   all    together    and the Lea
       we went out all together (0.8) and Tea
```

111

The target here is line 10 and the TCU that starts with "siamo usciti" (we went out). On the last syllable of "usciti" (went out) the recipient start shifting her head orientation toward the speaker.

Finally, extreme characterizations or formulations (see Edwards, 2000; Pomerantz, 1986; Sidnell, 2004) can be used "to counter challenges to the legitimacy of complaints, accusations, justifications and defenses; to propose a phenomenon is 'in object' or objective rather than a product of the interaction or the circumstances; to propose that some behavior is not wrong, or is right, by virtue of its status as frequently occurring or commonly done" (Pomerantz, 1986: 219-220). They can also be used to project the upcoming explanation of why something or somebody is characterized in first position as "unbelievable", "terrible", "a mess", "a bulldozer". Example 2.13 shows an extreme characterization in which a girl is said to behave like a "bulldozer".

**(2.11) 2GSTUDYING_trattore 37:47**



```
07        (1.5)


08 A:    HUUU A volte  e' un  c a t e r p i l l a r
         huuu at times is a  bulldozer
         HUUU Sometimes (she) is a bulldozer
```
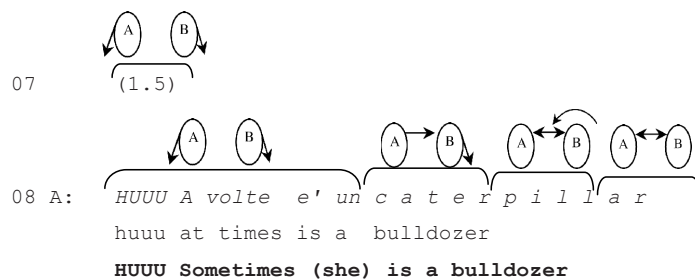
In this example we see that after the second syllable of the extreme characterization- the word "caterpillar" (bulldozer)- the recipient starts turning toward the speaker and they engage in mutual gaze.

I have shown that each of these turn design features can appear before the recipient starts looking toward the speaker, but to establish whether these features are actually associated with or attract recipient gaze requires an analysis of their distribution with respect to the sequential environments. I show this in Table 2.13.

*Table 2.13* Frequencies of turn design features in first TCU of an ETS or an APBS

| Features | ETS | APBS |
|---|---|---|
| 1) I, me | 19/38 (50%) | 21/103 (20.4%) |
| 2) Time reference | 16/38 (42.1%) | 9/103 (8.7%) |
| 3) Third person | 16/38 (42.1%) | 13/103 (12.6%) |
| 4) Epistemic access | 13/38 (34.2%) | 5/103 (4.9%) |
| 5) First action in schema | 7/38 (18.4%) | 1/103 (1%) |
| 6) Extreme characterizations | 4/38 (10.5%) | 10/103 (9.7%) |
|  |  |  |
| a) You | 3/38 (7.8%) | 47/103 (45.6%) |
| b) Deictics (verbal or visual) | 1/38 (2.6%) | 40/103 (38.8%) |
| c) But | 3/38 (7.8%) | 23/103 (22.3%) |
| d) Modals | 2/38 (5.3%) | 12/103 (11.6%) |

Table 2.13 reports the instances and percentages of the occurrence of an item in the first TCU of an ETS or an APBS. The first six occur at the beginning of ETSs at least 10% of the time (up to a maximum of 50%), while their occurrences in first pair parts of APBS is minimal and much less likely (from 20% for the use of I to 1% for actions that are first steps in a schema). On the other hand, if we look at the last four items we can see the opposite effect: they all occur much more often in sequence initiating turns (from a minimum of 11.6% of times up to 45.6% of times).

This shows that their distribution is systematically different and that the likelihood of having one of these items present at the beginning of a TCU could help recipients project which kind of turn the speaker is going to produce.

I have previously stressed that recipients look at speakers only 42% of the time during sequence initiating turns and I suggest here that, for example, if we hear a time reference we are more likely to look toward the speaker because we can project that a telling will be produced. If this claim is correct then we should see that in the first TCUs of APBSs in which the recipient looks up, a time reference should be more frequent than in the TCUs in which the recipient does not look toward the speaker. This is indeed the case. A time reference is present in 16.3% of the first TCUs of APBS in which the recipient looks at the speaker, while it is present in only 3.3% of the first TCUS of APBS in which the recipient does not look at the speaker. The recipient is therefore 5 times more likely to look up toward the speaker

when there is a time reference. This difference is statistically significant.[28] The same holds, for example, for the presence of third person references. In this case a third person is present in 23.3% of first TCUs in which the recipient looks at the speaker while it is present only in 5% of the first TCUs of APBS in which the recipient does not look at the speaker. Even in this case, the difference is statistically significant.[29]

This suggests that recipients are more likely to look at the speaker during the first TCU of an APBS that includes one of the features mentioned above. Participants therefore appear to orient to the occurrence of one or more of these items during the course of a TCU to project whether the turn is going to be the beginning of an ETS or an APBS.

The final piece of evidence comes again from what we may consider deviant cases. In a bottom up inferential probabilistic model it should be possible that participants hear a feature more often associated with an ETS, they look toward the speaker and they keep looking until they realize that in that particular instance the speaker was not starting a telling but just producing a single utterance. At that point we should see something happening in the gaze behavior of the recipients that displays that the heuristic procedure they followed failed in that specific case. Example 2.15 is a case in point.

Example 2.14 is taken from an interaction in which two friends are chatting sitting on a sofa (same participants as in Examples 2.1 and 2.3). A has just mentioned that she had been dreaming about a political figure of the 20th century from Saudi Arabia and that she was shocked about that dream because the exam about history of Arabic countries will be in one month.  Then they laugh about the dream and they are silent for a couple of seconds before the beginning of this extract. Our focus is what happens during lines 1 to 4.


**(2.14) 2GSOFA_faccio 25:47**



```
01 A:   Me    e' un casino Chia e' un casino tutta quella storia
        Shit  is a  mess   Chia is a  mess   all   that  story
        Shit it is a mess Chia it is a mess all that story

02      li'   non ce  la   faccio
        there not cl. that make.1s
        there I cannot make it
```

---

[28] Fisher Exact Test p = 0.03
[29] Fisher Exact Test p= 0.01

```
                    A ↔ B
03 B:    ⌒Hu hu hu [hu⌒
         Hu hu hu [hu

                         A ← B
04 A:           [Non ce la   faccio⌒
                 not cl. that make.1s
                [I cannot make it

          A ← B      A ↓↓ B
05        ⌒(0.4)⌒   ⌒(0.4)⌒
            (0.8)

          A ↓↓ B   A ←↑ B   A ← B
06 A:  ⌒E  Guros "Ma⌒ si'⌒ dai (0.2)⌒
       and Guros  But  yes  give
       And Guros "But yes   come on (0.2)

                    A ← B
07 A:  ⌒te  lo  leggi cosi' per cultura  generale.⌒
       you that read  so   for culture  general
       you read it just for general knowledge.
```
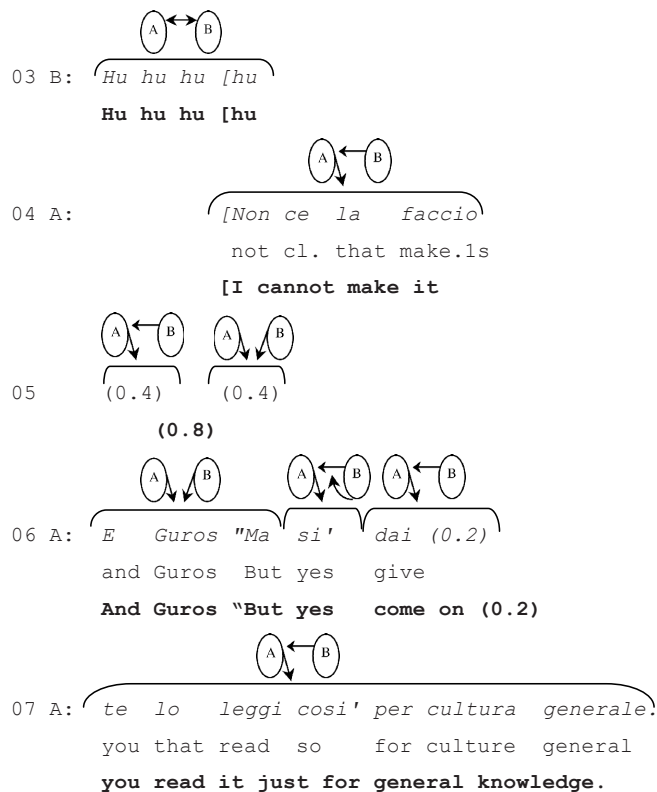
A's turn starts at line 1 with an outloud (shit) and then an extreme characterization of her situation (feature 6 of ETS) "e' un casino" (it is a mess) and a vocative (Chia, B's nickname). At this point B, who was looking at a biscuit she had in her hand, looks up toward A and sustains the gaze. While an address term can attract the gaze of the recipient, an extreme formulation in first position could be projecting a telling that accounts for the use of that characterization. But what follows in A's turn is a repetition of the assessment with the specification of the object of the predication of the assessment and finally A states her concern about her possible success in that exam (lines 2 and 4). B laughs and keeps looking at A, apparently oriented toward the possibility of the occurrence of an ETS, but the next piece by A does not further the progressivity of the talk (Schegloff, 1979, 2007b; Stivers & Robinson, 2006), and it is rather a repetition of the last TCU of line 2. Line 4 is produced with the speaker looking down, rather than toward the recipient as before. This could show that something else is going on, and that potentially there will be no telling. B keeps looking toward A for 0.4 seconds into the transition relevance place and then she re-orients toward the biscuit she was holding in her hand. At this point, given the orientation of both

participants, the re-engagement into eating a biscuit by B and the silence, it looks like participants are treating the complaint by A as possibly complete and no further talk on the topic seems to be on the way. B's sustained orientation toward A has indeed been abandoned and she is re-engaging into the activity she had momentarily interrupted (eating a biscuit). But at line 6 A starts a telling by using the conjunction "e" (and), the mentioning of a third person non present and not mentioned in the immediately preceding talk, and starts quoting what this person said about the preparation for the exam. After "e Guros ma" the recipient re-orients toward the speaker, who is looking mid distance in a sort of empty space, and again sustains gaze toward the speaker. The telling then follows.

This example shows that participants can orient toward features of the talk as potential projections of an ETS and sustain their gaze until it seems that the telling will not be produced. On the other hand, if the telling is then started, the recipient will orient toward the speaker and sustain the gaze as in all the previous examples. This confirms, then, that there are specific features of talk in specific sequential positions that can evoke the possibility of an upcoming telling. Participants display an orientation toward them exactly like in the cases in which the telling actually occurs immediately after them, as at the moment in which these features occur the actual occurrence of an ETS is only a possibility, not a certainty.

I have shown that recipients can start turning and looking toward speakers after each of the six design features I have listed for ETSs.[30] I have then shown that these features tend to occur more in an ETS than in an APBS and that at least the first four of them tend to occur more in the first TCU of APBSs in which the recipient looks toward the speaker than in the ones in which the recipient does not look at the speaker. Finally I have provided evidence that participants can use these features occasionally incorrectly to project an upcoming ETS that is not actually going to occur. At that point a participant may look away and abandon the orientation toward the other participant. All this provides evidence of the deployment of bottom-up heuristics by recipients to project the occurrence of a telling versus, for example, a request.

---

[30] During a discussion of these results, a commentator claimed that the 6 features that get the recipient to look toward the speaker simply indicate interactional displacement. In other words that the topic of the conversation is changing, and as such that what will be said is different from the previous activity. This is a plausible hypothesis and would certainly apply to the time reference feature, yet some features do not appear to have this function, but rather are specific of telling environments. For example, in what way does using the first person pronoun indicate displacement? And in what ways does an extreme characterization constitute a displacement? Moreover, if they are the first thing that gets said after a lapse (a long silence), as is often the case, then anything that gets said would constitute a displacement.

I have also stressed that often the recipient looks toward the speaker after very few words and suggested that it is impossible, at that point, to project the exact action that the TCU will implement. However, it is possible to infer to which 'ball park' the turn will belong, especially in terms of activity and sequential environment. This means that if the speaker starts a turn of talk with 'yesterday' I can guess there could be a telling or a report but, as we will see, for instance, in the following example, it is impossible to determine whether the action performed through the TCU will be an announcement or an assessment. On the other hand, a proper description of the patterns observed should say whether recipients always react after one of those design features or whether it happens more often that they need to listen to more than one of those features before they orient toward the speaker. Before showing the distribution of features that occur before the looking up of the recipient, I show two examples where it is clear that more than one feature is present. Example 2.15 shows a first TCU of an ETS in which there are two design features before the recipient turns toward the speaker. In this example the speaker starts a telling by assessing how her favorite soccer player played the day before. The previous topic of the conversation was an Arabic politician and the preparation for an exam on the history of the Middle East.

**(2.15) 2GSOFA-giocato 26:46**

```
01    (0.4)
```



```
02 A:  Niente  ieri    Sheva ha giocato  di  merda
       Nothing yesterday Sheva has played  of  shit
       Well yesterday Sheva played "like shit"
```
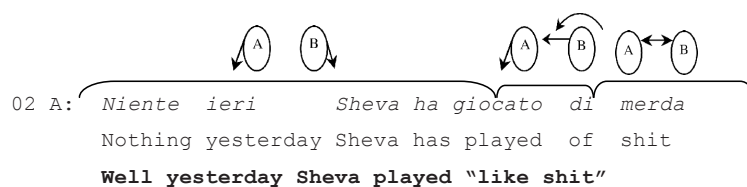
The recipient starts turning on the word "giocato" (played), after the speaker has produced a time reference (yesterday) and introduced a new character in the conversation (Sheva). The following example, Example 2.6, has been previously presented in this chapter without a transcript of gaze behavior. If we look at what happens during the first TCU of the ETS we can see that the recipient reacts after three design features have been produced.

**(2.6) 2PCOMP_Porto 10:49**

```
03     (.)
```

```
              Ⓐ  Ⓑ                              Ⓐ→Ⓑ Ⓐ→Ⓑ
04 B:  hm quella volta che mi son svegliato e    si sentiva iu:: iu::
       hm that   time that me  woke up      and cl. hear   iu   iu
       hm the time in which I woke up and you could hear iu:: iu::
```
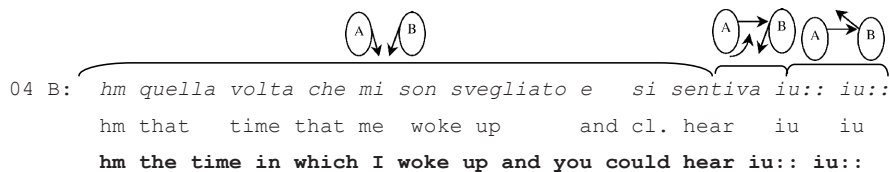
The recipient, A, starts turning on the word "sentiva" (hear). Before that, the speaker had used a time reference (the time), a first action in schema (to wake up) and the first person (I), although inflected in the verb.

It is possible, then, to have more than one feature occurring before the recipient reacts, moreover, as we can see, the speaker does not display any sign of trouble for not having the recipient gaze after the first design feature. The question is, rather, whether it is common for participants to react only after more than one feature. Table 2.14 shows that this is not the case and that only in a minority of cases the recipient looks up after multiple cues.

*Table 2.14* Distribution of features before recipient looks toward speaker in first TCU of ETS

| Number of design features before R looks at S | Instances |
|---|---|
| 1 | 25/35 (71.4%) |
| 2 | 5/35 (14.3%) |
| 3 | 5/35 (14.3%) |

If we consider only the 35 cases of ETSs in which the recipient looks toward the speaker, in 71.4% of the cases the recipient reacts after just one cue. There are, nonetheless, quite a few cases in which the recipients rely on more cues before looking toward the speakers. How can this be reconciled with the idea that participants rely on turn design features to project what the sequential environment of the talk will be?

Here is a simple way of conceptualizing it: imagine a semantic domain, such as time reference, and let's say that it corresponds to building block X (see Figure 2.8). When a recipient hears X s/he looks toward the speaker. On the other hand, given that a human being does not always react like Pavlov's dog, it is possible that the participant will consider X just a little cue and will wait for another cue to react, let's say Y. As we can see from table 2.13, a time reference is 5 times more likely to occur in the first TCU of an ETS than in the first TCU of a sequence initiating turn, but it still occurs 9% of the times in sequence initiating turns. So it is not the case that if the participant hears a time reference then s/he can be sure

118

that s/he is listening to the beginning of an ETS. This is why it is not unreasonable to wait for more evidence, especially if the same participant is potentially involved in some other competing activity. Sometimes Z will have to occur too, other times, Y and Z by themselves will be enough and A actually occurs after the recipient has already started looking at the speaker, simply because it is part of how a certain turn is designed. Every building block would have a different force in terms of cuing for a certain type of action or the occurrence of an ETS or an APBS. This force would come from the likelihood of being the discriminator between the two options, or at least from the likelihood of seeing it associated with a specific type of action. This force would not have to be constant, as there could be other contextual factors that could push the recipient to react to just one cue or to wait for more.

1) [ X ] R↑[31]→TCU continues

2) [ X | Y ] R↑ →TCU continues

3) [ Y | Z ] R↑ [ X ] →TCU continues
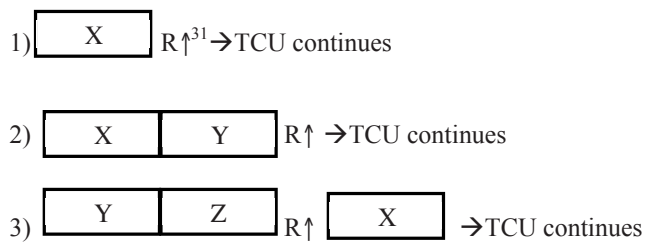
Figure 2.8. Scheme of features configurations that allow for probabilistic inference of sequential environment.

Neither the *Take the Best* nor the *Recognition Heuristic* correspond to the mechanism used by recipients to decide when to look toward the speaker and whether to sustain the gaze or not. Rather the occurrence of more cues one after the other, even if the second and third are less discriminating than the first one, increases the discriminating power of the whole. There is no neglecting or ignoring of the previous cues but rather a summative effect. If, for example, a time reference is 5 times more likely to occur at the beginning of an ETS than at the beginning of an APBS but the probability of being listening to an ETS is only 2.5 times higher if I hear a first person pronoun, once I hear "Yesterday evening I" the odds that this will be an ETS become higher. If both features per se would suggest it is more likely that the current TCU is the beginning of an ETS, once you see both of them one after the other the odds that this is not an ETS become lower and lower. It is a matter of joint probability of the occurrence of these features one after the other in the same utterance that explains why they are very powerful cues. It appears then that participants operate at a simpler computational

---

[31]R↑ means recipient looks up towards the speaker

119

level: they apply the elementary *probability rule for independent events*: P(A & B)= P(A) * P(B). So if P(A)= 0.5 and P(B)= 0.5 then P(A & B)= 0.25. It is highly unlikely they would occur one after the other by chance and so they become powerful identifiers of extended telling sequences.

We could say that if one knows that something is a telling, the latter is very likely to have some specific turn features in it. If one does not know what something is but s/he finds some of these features, it is plausible to infer that it is a telling. To recognize and appreciate the meaning of a part, we need to have an idea of what the whole might be.

Studying the timing of recipients looking up toward the speaker in ETSs opens a sort of window into their cognition. It provides us with clues concerning the online comprehension process of the recipient. Once we know that recipients regularly look up toward speakers and sustain their gaze during a projected telling, we can try to trace back not only which action performed through a TCU allowed them to recognize the possible beginning of a telling, but also which words within the same TCU were potentially enough to cue the possibility of an extended telling. And in the future we can possibly establish which feature seems to matter the most in projecting that a telling is going to begin, knowing that they are likely correlated.

In this section I discussed four features that participants may rely on and tested them as plausible hypotheses. I have provided the distributions of different features that could be used by recipients to determine whether looking toward the speaker is particularly relevant or appropriate. I have shown that while it is impossible to reject any influence on recipient's decisions, the first three (speaker's gaze, sentence type and action type of the TCU) do not account in a major way for the orderly recipient gaze behavior and in particular for its timing. I then proposed that participants are using the occurrence of words belonging to specific semantic domains to abductively infer whether an ETS or an APBS will be produced. I have listed some that clearly emerge from this data set. I have provided illustrations of them and I have provided their comparative frequencies. I have shown that most of the time just one of them is sufficient for the recipient to turn her/his gaze toward the speaker and I have explained how to account for the situations in which more than one cue is present before the recipient reacts. By doing this I have shown that turn design features are likely used as cues for determining whether a TCU is part of an ETS or of an APBS. Finally, I have proposed that participants in a conversation would apply the elementary probability rule for independent events to discriminate between sequential environments, rather than using more complex heuristics.

120

*2.5 Discussion*

This chapter shows that Kendon's distinction between speaker's and listener's gaze patterns is not sufficiently accurate to describe what happens during a dyadic conversation. A more fine-grained analysis demonstrates that those gaze patterns apply only to ETSs. In other words, Kendon's rule works if you replace the general words 'speaker' and 'listener' with 'teller of an ETS' and 'recipient of an ETS'. In a different sequential environment, such as in an APBS, the recipient's gaze behavior is much more variable. The evidence comes not only from a sampling of ETSs and APBSs such as question-answer sequences but also from a comparison of recipient's gaze behavior during the first TCU of an ETS and the first TCU of an APBS. The data show that when a recipient is not looking at the speaker before the onset of those TCUs, if something in the TCU projects (and retroactively the TCU actually constitutes) the beginning of an ETS, then the recipient almost always (92% of the time) looks up toward the speaker at that point. By contrast, if the TCU constitutes the first move in an APBS, only around half of the time do recipients look up toward speakers. This also demonstrates that participants in interaction do not look toward the speaker as soon as they hear some sounds produced by the speaker. Rather, recipients look at speakers once this orientation becomes socially relevant. The looking toward an addressee is therefore not merely an instinctive response to an auditory stimulus but it can be a constitutive component of a social action, with its own implications.

Recipient gaze and body orientation can be seen as a systematic way of displaying a fine grained online parsing of the ongoing conversation The recipient projects in which activity he is going to be involved next before the TCU is complete and his gaze orientation does or does not change accordingly. This displays that participants treat listening to the beginning of a story telling or of an explanation as a different type of interactional engagement when compared to listening to a question, a noticing or a directive.

In a paper on the invariants of human behavior, Herbert Simon (1990) suggested that even though "the fundamental goal of science is to find invariants, such as conservation of mass and energy and the speed of light in physics" (p. 1) "some of the most important invariants in science are not quantitative at all" (p. 2) but have rather a qualitative structure. Talking about psychology and distinguishing it from classical mechanics he claimed that

*its laws are, and will be, limited in range and generality and will be mainly qualitative. Its invariants are and will be of the kinds that are appropriate to adaptive systems. Its success must be measured not by how closely it resembles physics but by how well it describes and explains human behavior (p. 2)*

He added that "human rational behavior is shaped by a scissors whose two blades are the structure of task environments and the computational capabilities of the actor" (p.7). Finally, he proposed that "because of the limits on their computing speeds and power, intelligent systems must use approximate methods to handle most tasks. Their rationality is bounded" (p.6).

If what we are looking for are invariants of human behavior and we are aware that they must be qualitative and cannot be only quantitative, we should embrace an explanation of how people manage to project what the speaker will be saying that is interactionally and cognitively plausible. We need to consider the possibility that there are simple heuristics that participants use which do not require overly costly computations nor the calculation of too many levels of theory of mind or implicatures. I suggest that one way that participants project whether a TCU is the beginning of an ETS or not is by relying on a probabilistic association between turn design features and the sequential environments in which they tend to occur. To succeed, however, one must know what the whole (an ETS or an APBS) might look like. I propose a simple mechanism of abductive inference that could be carried out under the limited amount of time and knowledge that participants have. This does not deny that the turn design features mentioned above can be combined compositionally to form the actions that participants in interaction consistently orient to. Nonetheless, I suggest that for recipients to decide whether they should orient their gaze toward the speaker or whether they can continue looking down, a proper analysis of the sentence type or of the exact action type of a TCU is not necessary and is practically impossible anyway if all they have heard is "the other day" or " My mum says". The fact that participants look toward speakers at those points works as additional evidence of the plausibility of my claim.

Perhaps this gaze behavior is systematic because the participants know each other well. It is a general claim in social psychology that closeness, intimacy and level of acquaintance between participants affect the amount of gaze between participants (see e.g. Argyle & Dean, 1965; Bente, Donaghy, & Suwelack, 1998; Cordell & McGahan, 2004; Exline, 1963). Usually in experimental work this variable is taken care of by using

122

unacquainted subjects. It remains an empirical question whether friendship affects the principles here outlined.

The findings of this chapter display that participants treat listening to the beginning of a story telling or of an explanation as a different type of interactional engagement when compared to listening to a question, a noticing or a directive. This finding is particularly important when compared to Goffman's analysis (1981a) of the different types of hearers that might occur in a social situation. He distinguished between 'official' and 'unofficial' participants. Within official participants he distinguished between 'addressed' and 'unaddressed recipients' and among the unofficial participants between 'eavesdropper', 'overhearer', 'bystander' and 'audience'. What matters for us is that within addressed recipients, one must further distinguish between a recipient of an extended telling vs. a recipient of the first pair part of an adjacency pair sequence. In this way, we can further refine our understanding of participation in a social situation and recognize the specific behavior that an individual is expected to implement in order to act like an attentive and proper recipient.

The implications and relevance of the findings presented in this chapter can be seen from two different perspectives. On one hand, from a participant point of view, the looking up and then sustaining of the gaze toward the speaker could be used by the speaker as a recipient signal of her/his potential involvement as an extended telling recipient. This occurs much earlier than the traditional place in which the recipient signals recognition of her/his being a listener of an extended turn of talk: the first transition relevance place at TCU completion. The findings of this chapter suggest that even before producing a continuer, nodding or withholding talk in the first transition relevance place, the looking toward the speaker at specific places and sustaining of the gaze into the TRP can work as a cue for the speaker. The importance of this can be fully understood once the high vulnerability to derailment of the beginning of extended tellings is taken into account.

On the other hand, from an analytic point of view, this confirms that a participant in interaction not only auditorily monitors what is going on, but systematically analyzes and projects what kind of behavior is going to be relevant next from her/him. Looking toward the speaker while s/he is producing a multi-unit telling is evidently one of the behaviors that are relevant as soon as the telling is projectable, given the systematicity of the pattern observed. Indeed, the recipient's sustaining of the gaze toward the speaker throughout the telling seems

to indicate that the preferred gaze pattern for the recipient in a multi-unit telling requires a sustained look toward the speaker.

A final comment is on the possible reason for such a difference in gaze behavior between ETSs and APBSs. Two answers, not necessarily mutually exclusive, could be provided. The first answer to this question lies in the phenomenological notion of 'aboutness' and in the way in which gaze, also through its indexicality, can be used to display 'aboutness'. Talking about story telling, Harvey Sacks asked (1992 [1964-72]: 768, Vol. I): "what is a story about, by virtue of the fact that it's between those two? Stories are 'about' – have to do with – the people who are telling them and hearing them". If a multi-unit telling is about the participants, their relationship, etc. then in the same way in which we look at pictures to do "looking at pictures" or we orient toward the dishes to do "washing the dishes", then looking at the speaker during a telling becomes a simple acknowledgment of what a story is about: it tells something about the teller, his dreams, issues, what he finds funny or scary, annoying or entertaining. As such it tells something about the relationship with the recipient by doing basic socialization. In the same way in which I have to look at the pictures if I understand that what we are doing together is "about" the pictures and requires "looking at pictures" then looking at the teller of a story is not simply a matter of attention or lip reading but rather something fundamentally social. It becomes a way of displaying an understanding of what we are doing together and what a telling is (also) about: not just human beings, but the specific human being who is producing it and the one who is listening to it.

The second answer to the question of why recipients look at speakers during tellings deals with the notion of tellability and what it takes to produce an ETS. Among the things that matter, in order to produce an extended turn of talk, there is the existence of an audience to which the story can be told in so many utterances. As I said earlier in this chapter, once a participant starts talking s/he is entitled to produce one TCU (Sacks et al., 1974). To be able to produce more than that, a speaker will have to either project that a longer turn will be produced or make sure that possible transition relevance places are passed while the same participant keeps speaking (e.g. speeding up the talk, latching the end of a TCU to the beginning of another one). One important consequence of trying to produce an ETS consists of the necessity of having other participants to collaborate in this endeavor by withholding producing turns at talk. If the speaker needs to know whether s/he will have an audience to the telling that will align with the fact that this could require multiple TCUs and their collaboration, having a visible cue like sustained gaze toward speaker as soon as the telling is

recognizable as such becomes a useful resource. The idea is that by looking at the speaker very early in the first TCU of an ETS, the recipient displays that s/he recognizes that a possible telling is in progress and will act as a recipient. Additionally, by looking toward the speaker a recipient can heighten the input of information facilitating a proper calibration of appropriate uptake. It should be noted that in this case the issue is not just whether the recipient is paying attention to the speaker and therefore whether s/he is listening at all. The issue is the general accomplishment of a course of action that requires the collaboration of at least another individual. By sustaining gaze toward the speaker a recipient displays an understanding of what the course of action will be and the potential willingness to act as a recipient where it would otherwise be possible to take the floor and become the speaker.

While the first answer deals with a possible intrinsic understanding of what the things we do or say are "about", the second one takes a more "action" oriented perspective and sees recipient's gaze as an active cue for the speaker, rather than a mere display of attentiveness. It shows action recognition and shared intentionality: "I can see what you are trying to do and I am going to align with it to make it possible".

As mentioned in chapter 1, one line of thought about gaze behavior tends to conceptualize it as something fundamentally responsive to the environment and potentially redundant from a communicative point of view: I orient my eyes and my attention where it is more needed, usually in order to accomplish specific tasks. Another way of thinking about it considers gaze behavior as something that can shape the environment and the actions accomplishable in it. It can work as a way of projecting involvement and attention, it can work as a pledge of commitment, it can be actively directed toward the world to initiate engagement into an activity that had not been pressing and it can display 'aboutness'. The direction of my eyes and my head displays not just what I am currently attending but potentially what I am going to do next. The findings described in Hayhoe and Ballard (2005) (we look at the object that we are going to use or reach for before we actually reach for it) confirm the consistency of this view. The rest of this dissertation will focus on the latter way of thinking about gaze and will present different practices through which gaze can be used to do things in the world rather than being simply the gateway between the perceptual world and our brain.

# 3 Gaze as a Method of Pursuing Responses

**No endeavor is in vain;**
**Its reward is in the doing,**
**And the rapture of pursuing**
**Is the prize the vanquished gain.**
Henry Wadsworth Longfellow *The Wind Over the Chimney*

## *3.1 Introduction*

Chapter 2 showed how recipients of a turn of talk use different patterns of gaze behavior depending on which course of action (an extended telling or an adjacency-pair-based sequence) the first TCU projects. In that chapter the focus was on recipient gaze behavior during the first TCU of a turn. This chapter focuses on speaker gaze behavior and on a different interactional environment: the transition relevance place after possible turn completion. More specifically, this chapter examines the first turns in APBSs and shows what happens after the occurrence of a first pair part when the latter is not responded to promptly. It will show that speakers can use gaze to mobilize a response by the recipient when a response is not forthcoming.

The literature on gaze in social interaction has long emphasized its regulatory functions. They were first mentioned in relation to turn taking in conversation (Duncan, 1975; Duncan & Fiske, 1977; Duncan & Niederehe, 1974; Kendon, 1967), in particular in terms of signaling when somebody is going to start or stop producing a turn of talk and pass the floor to another participant. More recently, further work on gaze behavior in interaction has highlighted how gaze contributes to turn allocation and next speaker selection when there are more than two participants involved in a conversation (Lerner, 2003) and how children can distinguish a monitoring look from a sanctioning one by the caregiver and react accordingly (Kidwell, 2005). Another regulatory function that has been described is how gaze elicits a response from another participant (Kendon, 1967, Bavelas et al., 2002), which has been described as a symbolic function of speaker gaze during an interaction (Bavelas et al., 2002).

In general, the regulatory functions of an interactional practice can be more or less sensitive to both contextual features and participants' identities. Sacks, Schegloff and Jefferson (1974) distinguished interactional organizations (e.g., turn-taking) that can be

"context-sensitive" and "context-free". The organization of several related practices might be "context-free" if the specific context and participants do not affect the relevance or operation of an interactional organization (e.g., the turn-taking system). This means that every time some condition occurs (e.g., at the completion of a turn at talk), one of the practices that constitute that system can be implemented (e.g., turn allocation mechanism). On the other hand, a specific practice might be "context-sensitive" if the context might affect the actual implementation of that practice (as other variables might be at play and so the practice might have a differential effect).[1] "Context-free" and "context-sensitive" are not necessarily mutually exclusive, however, as some aspects of specific interactional practices might be "context-sensitive" while the system of practices might remain generally "context-free".

Within a larger "context-free" system of practices aimed at mobilizing responses (Stivers & Rossano, 2010), this chapter documents a practice that has a "context-free" dimension in that it can be deployed in any interactional situation, both institutional and ordinary, and the identities of the participants do not affect the implementation of this practice. However, this practice is also "context-sensitive".[2] In specific sequential environments, participants in conversation use gaze to both solicit a response and to display their understandings of where they are in a course of action. Specifically, by bringing their gaze to recipients during silence following a first action, speakers orient towards the relevance of a response by the recipient (i.e., they treat the responsive action as missing). And, recipients typically respond to this gaze practice by producing a relevant second pair part (SPP) to the initial first pair part (FPP). By looking toward the addressee in that sequential environment speakers employ gaze to pursue a response to a FPP, thereby treating the responsive action as missing.

So, rather than simply focusing on gaze as a regulatory turn-taking mechanism, I will focus on its indexicality. That is, gaze as a pursuit asks the recipient to consider what has been done before its occurrence and to perform an appropriate responsive action. By the time the speaker's gaze shifts towards the recipient, the recipient must "look backward" at what has just been said or done and, at the same time, "look forward" to produce an appropriate next action. In this sense, gaze provides a partial solution to the key interactional issue of

---

1 See, e.g., Heritage's investigation (2011) of the practice of prefacing a turn with "oh", in terms of "the limits of the context-free meaning of [the] practice and its context-sensitive uses" (p. 233). See also Lerner's work (2003) on the context sensitive operations through which forms of addressing can be used by participants to select next speakers in multi-party conversations.
2 "Context" here refers to a different level of granularity than the one outlined in "context-free", as this time it refers specifically to the conversational context, the sequential context of the talk, and not the social situation or social identity of the participants.

obtaining a response to a specific action. How can interactants get others to cooperate in an appropriate way in the development of a specific course of action? And how do recipients know when and how to do that if they fail or actively withhold reacting to the initiating action of another? By describing the organization of the practice of pursuing a response via gaze this chapter will offer a glimpse of some of the systematic resources people can rely on to deal with this interactional problem. In what follows, I first outline prior claims in the literature about the role of gaze in soliciting a response.

## *3.2 Speaker Looking Toward Addressee*

The sustaining of gaze by the speaker towards another participant has previously been described as a way of signaling the relevance of some response. Kendon (1967) describes how gaze can be used by interactants in dyadic conversations to regulate each other's behavior:

> *In looking up, which we have seen that [the speaker] does briefly at phrase endings, and for a longer time at the ends of his utterances, he can at once check on how his interlocutor is responding to what he is saying, and signal to him that he is looking for some response from him. And for his interlocutor, these intermittent glances serve as signals to him, as to when [the current speaker] wants a response from him. (Kendon, 1967: 42)*

Kendon also provides quantitative evidence to support the claim that the speaker looking up towards the recipient at the end of a long utterance[3] functions as a signal to the recipient that a response is expected. The long utterances that ended with extended looks got no response or a delayed one in only 29% of the cases, while the ones that ended without speaker's gaze got no response or a delayed one in 71% of the cases (Kendon, 1967: 36-37). In this way Kendon not only claims that speaker gaze displays the relevance of a response but also that it affects the timing of a response. In his seminal work, however, Kendon only deals with the general dynamics of turn taking. He does not make a distinction between sentence types

---

3 Kendon calls "long utterances" any utterance longer than 5 seconds. He is therefore likely referring to what was labeled as multi-unit turns in chapter 2; that is, turns that often constitute extended-telling sequences (e.g., storytellings or reports).

(such as interrogatives, imperatives, etc) or between action types performed through those turns at talk. Instead, he only talks in terms of long or short utterances.[4]

Duncan and colleagues (Duncan, 1975; Duncan & Fiske, 1977; Duncan & Niederehe, 1974) shares with Kendon his observations about the function of signalling the handing over the turn  for the speaker practice of looking up towards the recipient as the former approaches the end of her/his turn. Indeed, Duncan et al. list this as one of the basic cues deployed by participants in interaction to regulate turn taking. In trying to verify previous claims about the occurrence of looking towards a recipient approaching the end of the turn, Torres et al. (1997) find that of all turn endings in their data, only 16% include a look towards the recipient by the speaker. Moreover, these "look-towards" represent only 15% of all the speaker "look-towards" in their data.  Relying on information structure, and, in particular on Halliday's (1967) notions of *theme* and *rheme*, they find that "co-temporaneous rheme and end of turn always elicits a look-toward" (Torres et al. 1997: 8) but also that "45% of all the look-towards [other participant] were not associated with either the end of the turn or the beginning of rhematic material". This leads them to assert: "we still cannot account for the majority of gaze behavior (look-towards in particular) with the association of information structure and turn-taking" (Torres et al. 1997: 8). De Ruiter (2005), looking at task based dialogues, confirms the lack of a systematic relationship between gaze and turn-taking in general, therefore re-affirming the need for a better description of gaze functions in face-to-face interaction.

Starting from Kendon's claim and looking at a storytelling environment, Goodwin (1981: 103) observes that "nods are not only seen by the speaker, but seem to be organized precisely so as to be seen. They begin just after the speaker, who has turned her head away, returns her head to the recipient." Similarly, in storytelling environments produced in an experimental setting, Bavelas et al. (2002) find that "the listener tended to respond when the speaker looked at him or her, and the speaker tended to look away soon after the listener responded. Together, speakers and listeners created and used the gaze window [mutual gaze] to coordinate their actions" (2002: 576-577). Their notion of *gaze window* describes a mutual gaze situation in which it is not just that the speaker's gaze elicits a response but rather the listener's response seems to terminate the speaker's gaze. By "listener responses" they refer

---

4 Kendon actually talks about "short questions" as a type of utterance in which the speaker will look steadily at the recipient because the speaker "expects" an answer. Then he shows that this is the case in 75% of the questions in his corpus and accounts for the remaining questions in terms of utterances in which a lot of hesitant speech occurs and there is a high level of emotionality by the speaker. This, in his opinion, explains the looking away.

mainly to "mm hm", "uh huh" and "nods", which have been labeled in the literature as back channels (Yngve, 1970) or continuers (C. Goodwin, 1986b; Schegloff, 1982). They assert that one of the main features that distinguishes the gaze window from the pattern previously described in turn exchange is the fact that there is no exchange of roles between speaker and listener.

One of the main conclusions that Bavelas et al. draw is that gaze has a symbolic function, meaning that the speaker is not looking at the listener to monitor her/him for action but rather to solicit a response. A second conclusion is that "the listeners understood the meaning of the act because they responded immediately and appropriately" (Bavelas et al., 2002). Nonetheless, it remains unclear in which way the occurrence of a response before the withholding of gaze becomes evidence that listeners are responding specifically because of the gaze. Indeed, because of the coding system adopted by Bavelas et al., somebody could have been looking for 10 seconds and if s/he looks away after a response then the response would be considered related to the sustaining of the gaze, while the actual talk is not taken into account. Other signals to solicit a response are simply not taken into account, and, as such, the specific actions performed through talk are ignored. This suggests that the claim of a relationship between speaker gaze and listener's response needs further specification. What is particularly needed is stronger evidence that listeners are actually responding to speaker gaze.

In their work on the social organization of word searches, Goodwin & Goodwin (1986) refer to the occurrence of speaker gaze towards recipient in two examples of word searches and they describe its function as a way of soliciting aid from the other participant. No systematic evidence for this claim is presented and in both examples this apparent solicitation through gaze is unsuccessful. However, later in the chapter it will be shown in which way their observation resonates with a wider and more systematic understanding of the occurrence and function of speaker gaze in specific sequential environments. The analysis presented here reinforces the evidence of a connection between the organization of gaze and different levels of organization of human action and talk in interaction. Moreover, it reinforces previous observations by providing empirical evidence for the necessity of including at least one other level of analysis: the sequential level. Indeed, one of the limits of Kendon's characterization of gaze patterns in dyadic conversation, is that if focuses mainly on the notion of utterance (turn) and the idea that participants have to take the floor from each other and keep talking, rather than focusing on the social actions the participants are implementing with their talk. In a similar fashion, Bavelas et al. do not explain why the

131

speaker is soliciting uptake at those specific moments and in which ways the timing of this practice is organized. Specifically, they do not analyze what the participants are doing each moment with their talk, which limits the understanding of the orderliness of the practice they describe. This chapter shows why not considering the social actions that participants are implementing through their talk might be a problem.

Recent cross-linguistic work that takes the sequential environment into account, in particular whether a turn constitutes a first pair part of an adjacency pair or a first action of a possible sequence, provides additional evidence for the existence of a relationship between the occurrence of speaker gaze and the expectancy and timing of a response. In particular, in a study that compares gaze behavior during questions in three unrelated languages and cultures, Rossano et al. (2009) show that, on average, speakers look at recipients during questions in 73% of cases.[5] They also show that speaker gaze behavior during questions is similar across the three cultures while recipient gaze behavior differs. Questions are overwhelmingly responded to and the ones that do not get responded to are produced mainly with the recipient not looking at the speaker (there is a significant correlation between lack of mutual gaze and lack of responses to questions in two of the three languages). Moreover, they find that when the speaker or the recipient look towards the other participant during a question, they tend to do so from the beginning to the end of the question. Delayed looking towards the addressee tends to occur at particular sequential boundaries (usually at the beginning of courses of action or following their possible completion).

Similarly, in a paper focusing on the resources that participants can deploy to mobilize a response by a recipient, Stivers and Rossano (2010) list speaker gaze as an important resource to deploy in actions such as assessments (and potentially other first-pair-part actions as well) to pressure the recipient for a response.[6] They show that in face-to-face interactions, when the speaker is gazing at the recipient during an assessment, the recipient usually responds to the initial assessment, while the assessments that do not get responded to are usually produced without speaker gaze. Therefore the claim is that speaker gaze plays a role in mobilizing recipient response. Additional evidence for this claim comes from my own work on question-answer sequences in Italian dyadic conversations (Rossano, 2010). In this work I focus on resources that speakers deploy to signal to addressees that they are producing a question and therefore that they are pressuring for a response. I show that when speakers

---

5 Notice that this percentage is perfectly compatible with the one presented by Kendon (1967: 45, 47) for "short questions" produced with speaker's gaze in dyadic interactions in English (75%) and by Beattie (1978: 13) for questions in general produced with speaker's gaze in dyadic interactions in English (76.9%).
6 This finding in fact first emerged from the dissertation research here reported.

132

produce polar questions that do not have a prototypical interrogative intonational contour (e.g. if rather than ending with rising intonation, they end with a flat one, or they even have a declarative contour) they are significantly more likely to look towards their addressee than when they are producing questions with prototypical interrogative contours. Given that polar questions in Italian can only be marked intonationally and not morpho-syntactically, the occurrence of speaker gaze towards the recipient during a question appears to work as an additional cue that her/his turn is something that should be responded to.

Finally, in a work focusing on the timing of responses to yes-no questions in multiple languages, Stivers and colleagues (including the present author) (2009) show that the occurrence of speaker gaze accelerates the speed of responses to participants' questions. However, in this study, recipient gaze is not taken into account, the number of participants involved in the interactions varies (i.e. they are not all dyadic or tryadic interactions) and the domain of investigation is specifically limited to yes-no questions. Still, the finding is supported in 9 out of 10 of the diverse languages investigated.

These four recent works have in common the focus on a very specific domain of investigation (questions or assessments) and the focus on the occurrence of gaze (or not) towards the other participants during the action examined. The first part of this chapter will provide three further developments to those claims. The first one consists of generalizing the original observations made on questions or assessments to all FPPs. This means that the current domain of investigation is expanded to include not only questions and assessments but also announcements, sanctionings, complaints, accusations etc. These actions have been suggested to make relevant a response once they are produced (e.g., in Schegloff, 2007b). Thus, these actions potentially provide us with a broader claim about the role of gaze in pressuring addressees for a response. The second development consists of analyzing not only speaker gaze but also recipient gaze behavior during FPPs to see whether the picture changes once recipient gaze is included, and thus when mutual gaze analyses become possible. In particular, it is possible that speaker and recipient gaze affect the occurrence of a response and its timing in different ways, which will be illustrated below. In addition to analyzing gaze behavior during a turn, this chapter also analyzes gaze behavior during the transition relevance place that follows possible turn completion. At turn completion the other participant can start talking. If the first turn was a FPP, the other participant actually has an obligation to produce a response. What happens if the response does not occur in a timely fashion? How long do participants wait for a response before pursuing it? And if gaze has a role in pressing for a response, can a speaker use gaze to actually pursue a missing response?

These questions will be addressed in the second part of this chapter. The final part of this chapter will show how the findings provided in §3.3 and §3.4 can provide a framework for understanding similar speaker gaze practices (i.e., pursuing a response, or at least pressing for one) occurring in interactional environments in which the relevance of a response or uptake by the recipient is less straightforward.

### 3.3 Gaze and Sequence Organization

The first thing to establish before proceeding with a more detailed description of when in a sequence of actions gaze can be used to push for a response is that just engaging mutual gaze during talk does not mean that the speaker is pressing for some sort of response by the other participant. The following example shows this clear problem for the notion of mutual gaze as a "gaze window" (i.e., when the speaker engages mutual gaze with the addressee, the addressee responds) (Bavelas et al., 2002). In example 3.1 two young women are chatting sitting at a table face to face. B is telling A about her grandmother and how weird she has become as she has grown old. At the beginning of this extract they are talking about the grandmother's age (she is eighty-four years old) and the beginning of line 4 is a display of astonishment at this information. The focus is the participant's gaze during the turn at lines 4-5.

**(3.1) 2GC-dottoressa  24:00**

```
01 A:   [Ottantacinque_
         Eighty-five
        [Eighty-five_

02 B:   Ottantaquattro
        Eighty-four
        Eighty-four

03      (3.0)
```

134

```
          A↔B        A←B              A↔B          A←B
04 A:  Me::. Io ho avuto un nonno (.) di mio papa'  (0.2)
       Wow   I had.1s   a granfather of my   dad
       Wow:: I had a grandfather (.) from my dad's side (0.2)

          A↔B          A←B           A←B         A↔B
05     ho      il singhiozzo (.) che e' morto a ottantasette
       have.1s the hiccup        that is dead  at eighty-seven
       I    have  hiccups   (.) who died at eighty-seven years old

          A↔B
06 B:  .hhh
       .hhh

        A↔B      A→B       A↔B
07     (0.4)    (0.9)     (0.2)       ((B turns her head away from A))
       (1.5)

               A↔B                    A←B
08 B:  hhh hhhu Si' hh Speriamo di no   e(h)ccoh
               Yes     hope.1p  of not I mean
       hhh hhhu Yes hh Let's hope not I m(h)eanh
```

At line 4 A starts producing a turn that is interrupted by a side description (I have hiccups) that accounts for the intra-turn silences, which she then goes on to complete (line 5). If we look at the participants' gaze behavior we see that B, the recipient, is looking all the time, while A looks at her 4 times during the course of that turn and still B reacts only at line 6, with a sigh and turning her head away during the silence. B makes explicit what she meant with that reaction by producing line 8. We can also observe that A keeps looking at B until the end of line 8, when she withdraws her gaze and looks at the table. This example shows that mutual gaze during a turn at talk does not invariably entail that the recipient will produce some responsive action or feedback. That is, the engagement of mutual gaze, by itself, does not make a response relevant; however, sometimes a response occurs after the engagement of mutual gaze. This suggests the need for a more fine-grained analysis that can assess when the engagement of mutual gaze might lead to a responsive action to what has just been done by the speaker.

Any account of speaker looking towards recipient and of mutual gaze needs to take into account what happens in examples like this and it should be able to explain when speakers who are looking at addressees secure responses and when they do not. This might provide us with some information concerning whether speakers looking at addressees exert pressure for a response from the co-participant in the interaction. In what follows, I will show that it is necessary to include the sequential organization of action in an analytic framework of gaze and how the visible modality of gaze direction maps onto this level of order in interaction.

### 3.3.1 Sequential Patterns in Conversation

The first step towards understanding the relationship between speaker gaze and sequence organization consists in seeing how the first maps onto the second. In a dyadic interaction, when a participant A produces a sequence initiating action, B's responsive action can occur in the following 5 possible combinations:

*Table 3.1* Possible sequential configurations following the production of a First Pair Part

| Configuration | Features |
| --- | --- |
| 1 | A first pair part (FPP) action is followed, without delay, by a second pair part (SPP) action produced by the other participant (e.g., a request for information is followed by an answer that provides that information) |
| 2 | A SPP occurs, but it is produced partly in overlap with the FPP |
| 3 | A SPP occurs, but it is delayed |
| 4 | A FPP does not get responded to promptly and after some silence the same speaker pursues a response |
| 5 | A FPP does not get responded to but the speaker of the FPP does not pursue it and what follows is a sequence of actions unrelated to the initial FPP |

136

Here is a graphic representation of the configurations in Table 3.1:

1.
A: FPP
Time
B: SPP

2.
A: FPP
Time
B: SPP

3.
A: FPP | Silence
Time
B: SPP

4.
A: FPPb | Silence | A: FPPp
Time
B: SPP

5.
A: FPP1 | Silence | A: SPP2
Time
B: FPP2
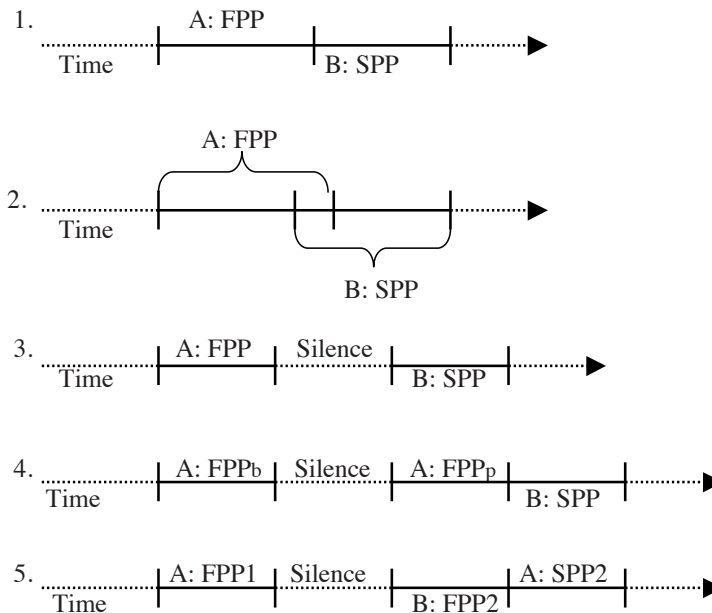
Configuration 4 can be further refined by distinguishing 4 possible types of pursuits:

a) the pursuit is done by repeating the FPP;
b) the pursuit is done by adding a delayed tag (e.g. *it's a beautiful day* (1.5) *isn't it?*);
c) the pursuit is done by further specifying the FPP (e.g. *how much does it cost?* (1.0) *How much does a helmet cost?*);
d) the pursuing turn modifies the polarity of the preferred response.

In pursuits *a*, *b* and *c* the polarity of the preferred response is not modified. Pursuit type *a* assumes that the lack of response might be due to a problem in hearing the FPP. Pursuit type *b* assumes that the recipient has heard the FPP and that some pressure to respond should be sufficient to elicit a response. Pursuit type *c* assumes that the lack of response is due to a possible problem in understanding. In pursuit type *d*, however, the speaker of the FPP, prefiguring a dispreferred answer because of the silence, facilitates the likelihood of obtaining a preferred response by backing down or upgrading the FPP or providing an alternative preferred answer (e.g. *did you take these notes?* (1.0) *did somebody give them to*

*you?*). The five configurations presented in Table 3.1 represent who speaks next and whether the FPP gets responded to or not and when. In this sense, configuration 4 can represent all fours situations (*a-d*) even though the way in which a response is pursued, and, as such, the response that will be provided, might vary substantially, as it will be illustrated below.

How are these patterns distributed in dyadic interactions? Answering this question required returning to the same dataset used in chapter 2[7] and characterizing all of the FPPs in terms of the five configurations outlined above. This first pass at the data did not consider the gaze behavior of the participants but instead only considered verbal pursuits. As in patterns 3-5, a SPP was considered delayed when more than 0.2 seconds[8] of silence occurs between the end of the FPP and the beginning of the SPP. Results are indicated in Table 3.2.

*Table 3.2* Distribution of verbal responsive patterns to First Pair Parts

| Possible Patterns | Number of instances |
|---|---|
| 1) FPP → SPP | 125 (29.9%) |
| 2) FPP → Overlapping SPP | 53 (12.7%) |
| 3) FPP → (Silence) → SPP | 174 (41.6%) |
| 4) FPP → (Silence) → Pursuit | 44 (10.5%) |
| 5) FPP → (Silence) → ∅ | 22 (5.3%) |
| **Total** | 418 (100%) |

Table 3.2 shows that almost 95% of the FPPs in my dataset get a responsive SPP, although 10.5% do so only after a verbal pursuit. However, the information in this table also suggests that more often than not the responses to first actions are produced with some delay rather than immediately after the FPP in dyadic face-to-face interactions.

We now turn to how gaze behavior maps onto these patterns. In the first sequential pattern, where the FPP is immediately followed by a SPP, the FPP is frequently produced with speaker gaze towards the recipient (in 77% of the cases). On the other hand, it is also the

---

7 Ten minutes of dyadic face-to-face interaction from 10 different interactions for a total of 100 minutes and 20 different participants. For further details see Appendix A.

8 Based on the data collected for the Stivers et al. (2009) comparison of question-answer sequences in 10 languages, it emerged that the average "on time" response for Italian conversation occurs 92 ms after the completion of the question. As such, 200 ms (0.2 s) corresponds to nearly 2 standard deviations above this average value. Moreover, of the 107 responses that occur after 200ms or more, only 6% are judged "on time" by a native speaker of Italian. All the others are judged to be late.

case that the recipient is often looking back at the speaker (in 74% of the cases). Participants engage in mutual gaze during those FPPs in 63% of the cases. Examples 3.2 and 3.3 illustrate these gaze configurations. In example 3.2 two friends are sitting on two couches talking about exams and how they usually study. A has just complained about how difficult it is for her to memorize what she reads and B has asked whether she underlines parts of the text while she reads. The focus is the speaker gaze behavior during the questions at lines 3 and 7 and the timing of the responses.

**(3.2) 2GSOFA-colori 34:27**

```
                                           A ←→ B
01 A:   e    mentre leggo   sottolineo   le  cose   che  secondo me
        and while  read.1s  underline.1s the things that according to me
        and while I read I underline th things that in my opinion

                              A ←→ B
02      sono piu' importanti [(che   s-)
        are  more important    which a-
        are more important [(which a-)

                                            A ←→ B
03 B:                           [Con   colori diversi?
                                 With colors different
                                 [With different colors?

                                    A ←→ B
04 A:   Con   colori diversi. ((+ nod)) Il risultato pero'   e'
        With colors different          The result    however is
        With different colors. ((+nod)) The result however is

                              A ←→ B
05      che  alla   fine sottolineo   tutto.
        that at the end  underline.1s everything
        that in the end I underline everything.
```

```
           ⟨A⟩↔⟨B⟩
          ⌒‾‾‾‾‾‾⌒
06         (0.2)


                ⟨A⟩↔⟨B⟩
            ⌒‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾⌒
07 B:      Cosa sottolinei   te=[che   colori] usi?
           What underline.2s you which colors  use.2s
           What do you underline=[which colors] do you use?


                    ⟨A⟩↔⟨B⟩
                ⌒‾‾‾‾‾‾‾‾‾⌒
08 A:          [T u t t o]
                Everything
               [Everything]


                      ⟨A⟩↔⟨B⟩
          ⌒‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾⌒
09 B:     Il  rosso e   il  blu  e   il  verde per i   t(h)itoli hhh .hh
          The red   and the blue and the green for the titles
          Red and blue and green for the t(h)itles hhh .hh
```
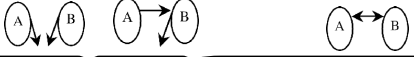
Here B asks A to further specify how she underlines the text while reading and A responds promptly to the three questions at lines 3 and 7. Although the answers are preferred in type, it should be noted that the one at line 4 ('with different colors') is not a type conforming answer (Raymond, 2003), it does confirming an allusion (Schegloff, 1996a), and the first one in line 8 ('everything') is a partial repetition of what A had just said at line 5, which was treated as not really informative by B. The main noticing here is that the questions are produced in mutual gaze and the answers are provided without delay.

Example 3.3 shows a similar pattern, only this time participants are not constantly looking at each other and the recipient tends to look away during the question and before the production of her timely answers (in 2 out of 3 questions). In this example two friends are sitting at a table (at a 90 degree angle) and the conversation has just started. We focus on the gaze behavior during lines 1, 3, 7-8.

140

**(3.3) 2GGOSS-stasera 00:15**

01 B:  .hh A(h)l l o r a stas(h)era cos' e' che fate
       S(h)o           tonight    what is that do.2s
       **.hh S(h)o t(h)onight what are you doing**

02 A:  Eh    andiamo a  Villa Chiara=
       Eh    go.1p   to Villa Chiara
       **Eh    we go to Villa Chiara=**

03 B:  =Ma  a  che   ora  vi  incontrate
        But at which hour you meet.2p
       **=But at what time do you meet**

04 A:  Vado  alle nove e  mezza dalla Gloria.
       Go.1s at   nine and half  to    Gloria
       **I go to Gloria's (house) at nine thirty.**

05     ((B annuisce))
       **((B nods))**

06     ((some turns with side sequence omitted))

07 B:          [Nove e   mezza ma]  andate subito    a  Villa
               Nine and half  but go.2p immediately to Villa
       **[Nine thirty but] do you go immediately to Villa**

08     Chiara alle nove e   mezza?=
       Chiara at   nine and half
       **Chiara at nine thirty?=**

141

```
         A ↗ B              A ↔ B
        ⌒⌒⌒⌒⌒⌒⌒⌒⌒⌒⌒⌒⌒⌒⌒⌒⌒⌒⌒⌒
09 A:   =Con- le   dieci con  gli altri.
         With the ten   with the others
        =With- at ten with the others.
         A ↔ B
        ⌒⌒⌒⌒⌒⌒
10      (0.4)
```

At line 1 the recipient starts looking at the speaker before the speaker turns towards her. By the time the question at line 3 starts, the participants are looking at each other and only the recipient withdraws gaze before the completion of the question. The same happens during the question at lines 7-8. Each of these questions are requests for information about A's plans for the night, probably meant to get an invitation by A. Participants engage in mutual gaze during each question and even though the recipient looks away before the completion of the last two questions, the answers are delivered without delay.[9]

Sometimes the speaker of a FPP looks at the addressee while the addressee does not look back. In my corpus, where the participants sit less than one meter away from each other, if the speaker has her/his head oriented towards the recipient, the recipient can be assumed to perceive this.[10] In situations in which the speaker looks at the addressee during the FPP but the latter is not looking back, often the responses are delayed. Example 3.4 illustrates this pattern. In this example, two friends are sitting side-by-side at a table looking at A's holidays pictures. At the beginning of this extract they are looking at a picture of the inside of a building in Rome. Our focus is the participants' gaze behavior during the FPP at line 2.

9 Pragmatically, the gist of the question is already clear when the recipient looks away during the questions at line 3 and 7-8, even though the speaker continues. It is possible that the recipient is looking away to think about the answer (see Kendon 1967 and Beattie 1979 on speakers looking away before speaking to help planning their own speech), nonetheless the answers to the requests for information are not delayed.
10 As mentioned in chapter 1, the human eye receives information from a field of about 200 degrees, although proper visual acuity provided by foveal vision is only about 15-30 degrees. In other words, our peripheral vision is very powerful even though the acuity of what we can perceive is quite poor.

**(3.4) 2PCOMP-pagato 9:33**

01 (1.0)

02 B:  *Soccia quanto hai pagato per entrar qua.*
       Wow    how much have paid  for enter   here
       **Wow how much did you pay to enter here.**

03 (0.2)

04 B:  *Die[ci*
       Ten
       **Te[n**

05 A:      *[Poco.  °Un euro. Due euro neanche*
           Little. One euro two euro neither
           **[Little.° One euro. Not even two euros**

06 (.)

07 B:  *Ah    beh*
       Oh    well
       **Oh    well**

08 (1.9)
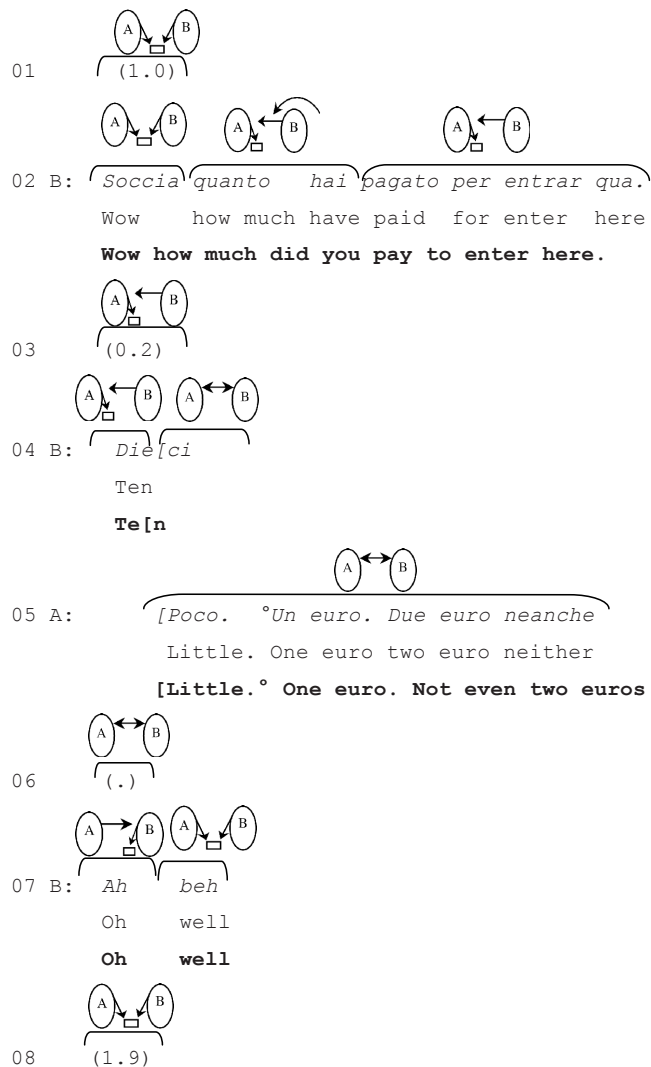
At line 2, B asks about the cost of entering the building represented in the picture. The speaker (B) turns towards the recipient (A) at the beginning of the question while the recipient keeps looking at the picture. After a delay in responding, B offers a candidate answer, but following this A looks towards B and begins a delayed answer.

On the rare occasions in which a FPP is not responded to, the participants are not usually looking at each other during the production of the FPP. Examples 3.5 and 3.6 exemplify this pattern. In example 3.5 two friends are sitting in front of each other talking

about exams and sometimes checking for information on the Internet. At the beginning of this example B is looking at his computer screen and continues looking at it, while A asks B to start asking him possible questions for the exam he has to take the next day. Our focus is the participants' gaze behavior during the request at lines 3-4.

**(3.5) 2PEXAM-chiedi 48:42**

```
01 B:  E   qui  ti: ti apre ((guardando lo schermo del computer))
       And here you you open.3s
       And here it opens ((for you)) ((looking at the pc monitor))

02     (4.4)  ((B sneezes))

03 A:  .hhh Scolta    mi chiedi un paio  di cose   della
            Listen.2s me ask.2s a  couple of things of the
       .hhh Listen ((can you))ask me a couple of things about the

04     rabbia e   della  brucellosis poi:[:::
       rabies and of the brucellosis then
       rabies and the brucellosis then:[:::

05 B:                            [Ma  qui  c'e'    un sacco
                                  But here there is a  bag
                                 [But here there is a lot

06     di roba  eh anche_
       of stuff eh also
       of stuff eh also

07     (0.9)

08 A:  Eh lo so
       Eh it know.1s
       Eh I know
```
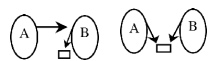
Here we see that A's request is completely ignored by B, and, rather than responding, B instead produces a noticing at line 5 while looking at the computer screen. At line 8 A responds to the announcement/noticing. In this example both speaker and recipient are not looking at each other, but while the recipient (B) is clearly involved in another activity, the speaker (A) is simply looking away from the other participant, only to turn towards him and the screen during the production of line 5. The FPP does not get a response and both participants were not looking at the one another during the FPP.

Example 3.6 shows a similar case, only this time the FPP is produced in overlap and is not responded to by the other participant, who instead keeps talking and completes his own turn failing to deal with the FPP entirely. In this example two friends are sitting in front of each other and B has just asked A for some advice in terms of recipes he could prepare for a boy-scout camp. During lines 46-47 B is looking down towards the table in front of them, mainly at a piece of paper and at a TV remote control that he moves slightly to the left. Our focus is the gaze behavior during the question at line 47.

**(3.6) 2PPLAN-scrivere 7:20**

```
41 A:  Come [sgua]ttero ce l'hai l' e(h)sp(h)erienza [hu hu hu]
       As scullery boy  cl. have the experience
       As  [scul]lery-boy you have e(h)xp(h)erience [hu hu hu]


42 B:       [(so)]                              [Si' ins(h)omma]
             know.1s                             Yes I mean
            [(I know)]                          [Yes I m(h)ean]


43     (0.6)


44 B:  M'aiuta.    Da solo no- non::
       Me help.3s  Alone no- not
       It helps me. On my own (I) don- I do::n't


45     (0.8)
```



```
46 B:  L[ui  fa   ah  vabbe'  vabbe'  va]bbe'  poi  non  ho  capi- =
       He do.3s ah    alright alright alright then not have.1s und-
       H[e goes "ah  alright alright al]right then I did not und- =
```

```
47 A:  [Cosa ti devo    scrivere scusa]
        What you must.1s write   sorry
       [What should I write for you excuse me]

48 B:  =di fatto   uno che   ci aiuti non so   se ci sia e (0.3)
        practically one which cl. help not know if cl. Is and
       =practically one who help us I do not know if there is and (0.3)
```
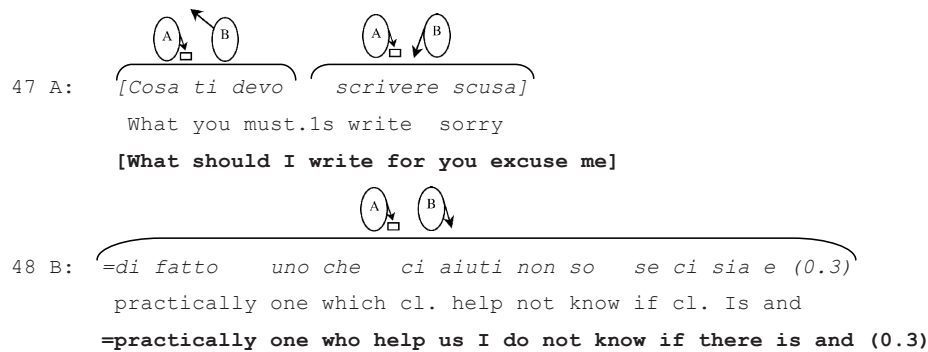
B does not respond to the request for clarification at line 47, and instead continues his turn-at-talk (beyond line 48; not shown). However, there is evidence that B has heard the question. In particular, we can see that immediately after A produces the question at line 47 (produced in overlap with the beginning of line 46), B looks briefly at A while saying "then not". This look towards A suggests that B heard A speaking and monitors what A is doing. However, B sees A looking down towards a piece of paper rather than looking back at him and so B immediately lowers his gaze and continues his turn.

Examples 3.5 and 3.6 show two situations in which a FPP is not responded to nor pursued:

1. when the other participant is involved in a competing activity and produces another unrelated FPP immediately after the first;
2. when the FPP is produced in overlap and the recipient continues producing his turn.

Concerning gaze behavior, we can see that in both cases the participants are not looking at each other during the production of the FPP. This is particularly interesting if compared to what we observed in examples 3.2 and 3.3, when both participants are looking at each other and the responses are produced on time. It is also important to notice the difference between example 3.4, where the speaker looks towards the addressee and holds his gaze towards the recipient, and examples 3.5 and 3.6. In example 3.4, the response is produced, although it is delayed, while in the latter two examples no response is produced at all. The speakers in examples 3.2-3.4 look towards the recipient whereas the speakers in examples 3.5 and 3.6 do not. A first qualitative observation then suggests that given the relationship between gaze and

direction of attention, if a participant, in particular a recipient, does not look towards the speaker during a FPP, this might lead to a delayed response to the FPP or no response at all. And yet, the occurrence of speaker gaze towards the addressee appears to play a role as well. To summarize, these extracts exemplify the following three patterns:

1. When participants engage in mutual gaze during FPP, the SPP is produced without delay.
2. When the speaker of the FPP looks at the addressee, who is in close proximity, the SPP is produced but it is delayed.
3. When both participants do not look towards the other during the FPP, the FPP is not responded to.

These three patterns deserve further investigation, particularly with respect to their robustness in a larger corpus and in terms of their relationship to the occurrence, and timing, of responses. The following section shows how speaker and recipient gaze differently influence the occurrence of a response and its timing.

## 3.3.2 Gaze During FPP, Response and Its Timing

In this section I intend to address two aspects of the relationship between gaze behavior and the occurrence of responses to FPPs:

1) whether gaze affects the provision of a response
2) whether gaze affects the timing of responses

To address these two questions I returned to the database of 100 minutes of naturally occurring dyadic interactions quantitatively analyzed in chapter 2. This database had already been coded for sequential patterns as shown in Table 3.2; however, I also coded for the following features:

1. speaker gaze during the FPP

2. recipient gaze during the FPP

3. mutual gaze during the FPP

4. face-threatening FPP

5. whether the FPP is a question

6. whether the FPP is an other-initiation of repair

7. preference compatibility of the SPP

8. type conformity of the SPP (if a yes/no question)

In what follows I account for my interest in each of the 8 factors. The first three factors concern the possible effect of gaze on the provision of a response to a FPP. In the previous section we observed a possible relation between the occurrence of mutual gaze and the occurrence of a response to the FPP. Looking at examples 3.5 and 3.6, in both cases neither party is looking at the other during the FPPs and the FPPs are not responded to, while in examples 3.2 and 3.3 both participants are looking at each other during the FPPs and the FPPs are systematically responded to. This immediately suggests that in a dyadic face-to-face conversation the occurrence of mutual gaze during a FPP could predict that a response will be produced, while lack of mutual gaze could be associated with lack of response.

Among the additional factors that might affect whether a recipient responds to a FPP is whether the FPP is what Brown and Levinson (1987) call a "face-threatening act". This term is used to refer to the kind of actions that could be socially perceived as challenging or negative because they attack the recipient's "face" (see also Goffman, 1955, 1959, 1967), such as challenges, accusations, criticisms, sanctionings etc. So the face-threatening status of a FPP should be included in a model that tries to understand systematic predictors of the occurrence of response or not. The fifth factor to take into account is whether the FPP has a question format, given that an interrogative formatting might add pressure for a response (see, e.g., Stivers and Rossano, 2010). The sixth factor that could affect the likelihood for a FPP to obtain a response is whether the FPP initiates repair (see, e.g., Schegloff, Jefferson, & Sacks, 1977). The hypothesis is that if a FPP is initiating repair, it is more likely that the recipient responds to it in order to repair her/his own talk.[11] Regarding the seventh factor, extensive research on the preference of SPPs (e.g., whether an invitation is accepted rather than rejected) has shown that

11 Another reason for responding might be that the repair initiation, in some cases, could be understood as prefiguring a disalignment or doing challenging (see, e.g., Schegloff, 2000b). However, a bivariate analysis of the effect of the FPP doing other initiation of repair suggests that there is no specific effect of this feature on the occurrence of a response.

preferred responses are produced faster and with no delay, while dispreferred ones are usually delayed (see, e.g., Pomerantz, 1984a; Sacks, 1987; Stivers et al., 2009). Finally, recent work by Raymond (2003) has shown that with respect to polar questions, type conformity of the SPP (i.e., whether the recipient responds with a 'yes' or 'no' rather than with some other form) affects delay in the production of SPPs. In particular, type conforming responses are produced faster than non-type-conforming ones.

To address the question of whether gaze affects the provision of a response, a logistic regression analysis was conducted on the predictors of response after a FPP, where a FPP is considered responded to if it receives a response without a pursuit by the speaker of the FPP, regardless of its exact timing.

*Table 3.3* Predictors of response including mutual gaze. Results of logistic regression with data clustered by interaction. [12]

| Variables | Odds Ratio | 95% Confidence Interval |
|---|---|---|
| Mutual Gaze | 3.95*** | 2.25, 6.92 |
| Face threatening FPP | 1.51 | 0.66, 3.42 |
| Repair | 1.26 | 0.42, 3.80 |
| Question | 0.93 | 0.59, 1.45 |

Table 3.3 shows that the status of a FPP as face-threatening, a repair initiation or a question does not predict whether the FPP will be responded to or not in any significant way, while the occurrence of mutual gaze makes it almost 4 times more likely that the first action will be responded to than if it did not occur. This shows that mutual gaze by the participants during the FPP is a good predictor of the occurrence of a response after its completion.

The second question to be addressed in this section concerns the effect of gaze on the timing of the response. The relevant timing here is not the exact length of silence after which the response occurs, but rather the timing of a response conceptualized according to turn-taking rules and the details of sequence organization (see, e.g., Schegloff, 2007b). In this sense, if we exclude the turns that do not get responded to (and those which must be pursued), only three scenarios are possible:

---

12 *** denotes p<.001

1. The FPP gets responded to promptly, without an audible delay (i.e., within 0.2 s)

2. The FPP gets responded to but in overlap

3. The FPP gets responded to but after some silence

Given the complexity of the accounts for an early response (i.e., in overlap), I decided to exclude cases of overlap (#2) from this quantitative analysis, together with the FPPs that did not receive a response. The reason these were excluded is that often an overlapping SPP shows affiliation and optimal mutual understanding, strong disagreement or to prevent further talk on a specific topic. Moreover, there is no reason to believe that if participants engage in mutual gaze than the response would actually occur before the completion of the FPP, which is an interactionally marked action anyway. For all of these reasons, the investigation was reduced to turns that get a timely response and those that are responded to after a delay. As mentioned above, a response was coded as delayed if it occurred more than 0.2 s after the completion of the FPP.

We saw in examples 3.2 and 3.3 that FPPs during which participants engage in mutual gaze get a timely response. This suggests that the occurrence of *mutual gaze* during the FPP is probably a good predictor of the occurrence of a *timely* response. Among other factors that can independently influence the timing of a response there is its *preference*, intended as an interactional preference[13] (e.g., an invitation should be accepted, a request should be granted) and not a subjective or psychological notion of preference.

Another factor that could play a role in the timing of a response is its *type conformity*, as shown in the case of yes-no interrogatives by Raymond (2003). The results of the logistic regression analysis presented in Table 3.4 are obtained by considering as predictors of a timely response (versus a delayed response) only the occurrence of mutual gaze, whether the response is preferred or not and whether it is type conforming.

13 See, e.g., Pomerantz (1984a) or Sacks (1987) for a definition of preference in social interaction.

150

Table 3.4 Predictors of timing of response, including mutual gaze. Results of logistic regression with data clustered by interaction.[14]

| Variables | Odds Ratio | 95% Confidence Interval |
|---|---|---|
| Mutual Gaze | 2.16** | 1.27, 3.66 |
| Preference | 2.53*** | 1.71, 3.76 |
| Type Conformity | 0.89*** | 0.85, 0.93 |

Table 3.4 shows, as predicted, that preferred responses to a FPP, type conformity of the response and mutual gaze are all good predictors of whether the response will occur without a noticeable delay. In particular, the likelihood of getting a timely response goes up 2.16 times if participants engage in mutual gaze during the FPP and it increases to 2.53 times if the SPP is preferred.

The results presented in Tables 3.3 and 3.4 provide quantitative evidence that confirms the qualitative observations outlined with examples 3.2, 3.3, 3.5 and 3.6, namely, that the occurrence of mutual gaze is a good predictor of whether the FPP will be responded to or not and whether this response will be timely. It is not the *only* predictor but it is certainly a relevant one even when evaluated in combination with other variables. However, just knowing that mutual gaze is a significant predictor does not tell us much about the different possible roles that speaker and recipient gaze can play in predicting a response and its timely delivery. In example 3.4, for example, we saw that sometimes only one participant looks toward the addressee (in that case the speaker looks at the addressee) and we saw that the response was delayed. It could be that mutual gaze is a significant predictor because either the occurrence of speaker gaze or of recipient gaze is mainly driving the results and that each participant's gaze might be contributing something very different to the interaction. If this were the case, mutual gaze would be a byproduct of behavior independently enacted by each participant, rather than something that participants strive to achieve during conversation. To address this issue, two further logistic regression analyses were run, this time substituting mutual gaze with *speaker gaze* and *recipient gaze* during the FPP. Table 3.5 corresponds to the same dataset used for the analysis presented in Table 3.3 and shows the role of speaker and recipient gaze as predictors of the occurrence of a response.

---

14 ** denotes p<.01 *** denotes p<.001.

*Table 3.5* Predictors of response including speaker and recipient gaze. Results of logistic regression with data clustered by interaction.[15]

| Variables | Odds Ratio | 95% Confidence Interval |
|---|---|---|
| Speaker Gaze | 3.05*** | 1.85, 5.05 |
| Recipient Gaze | 1.40 | 0.98, 2.02 |
| Face threatening FPP | 1.55 | 0.64, 3.73 |
| Repair | 1.24 | 0.47, 3.28 |
| Question | 0.92 | 0.60, 1.41 |

Table 3.5 shows that speaker gaze is the factor that best predicts whether a response will occur or not after the FPP, while recipient gaze does not quite reach significance. This suggests that while recipient gaze plays a crucial role for the occurrence of mutual gaze (shown to be a significant predictor of the occurrence of a response in Table 3.3), the most important predictor of response is *speaker gaze* during the FPP. If the speaker looks towards the recipient during a FPP it is 3 times more likely that the FPP will be responded to. If we combine these results with the ones in Table 3.3, we see that once the recipient looks back at the speaker and they engage in mutual gaze, the likelihood of getting a response increases and becomes 3.9 times higher than if mutual gaze was not obtained. This shows that speaker gaze matters more than recipient gaze for the occurrence of a response to a FPP, and the importance of recipient gaze is conditional on speaker's gaze, i.e., on entering a state of mutual gaze.

Turning to the factors affecting the speed of a response rather than the presence of a response, Table 3.6 shows the results of a logistic regression analysis of the predictors of a timely response to the FPP, considering the same dataset used for the analysis run for the predictors in Table 3.4. Speaker gaze and recipient gaze substitute for mutual gaze as variables here.

---

15 *** denotes p<.001

*Table 3.6* Predictors of timing of response, including speaker and recipient gaze. Results of logistic regression with data clustered by interaction.[16]

| Variables | Odds Ratio | 95% Confidence Interval |
|---|---|---|
| Speaker Gaze | 1.43 | 0.81, 2.55 |
| Recipient Gaze | 1.98*** | 1.65, 2.38 |
| Preference | 2.50*** | 1.67, 3.77 |
| Type Conformity | 0.88*** | 0.84, 0.92 |

In contrast to the findings in Table 3.5, Table 3.6 shows that speaker gaze is not a good predictor of a timely response to the FPP, while recipient gaze is (and so are type conformity and the preference of the response). In particular, if the recipient gazes at the speaker during the FPP it is 2 times more likely that the recipient will respond promptly than if s/he is not looking at the speaker. If we compare these results with the ones presented in Table 3.4 we can see that the likelihood of getting a timely response becomes slightly higher when the participants engage in mutual gaze (2.16, rather than 1.98), and therefore speaker gaze matters once it is engaged in mutual gaze with the recipient.

### 3.3.3 Discussion

Tables 3.3 and 3.4 showed that the occurrence of mutual gaze during a FPP is an important predictor both of the occurrence of a response and of its occurring promptly, although other variables can play an important role as well. Tables 3.5 and 3.6 have further refined the finding by showing that for the occurrence of any response at all, what matters the most is actually the occurrence of speaker gaze, while when a response occurs, its timing depends mostly on recipient gaze. In other words, if the speaker is looking at the recipient, the recipient is more likely to respond to a FPP, but if the recipient is looking at the speaker, the recipient is more likely to respond promptly, without delays. The latter finding seems perfectly reasonable given that a gaze orientation towards the speaker during the FPP is not a default (see chapter 2) and that because of the biological limits of our foveal vision, looking at the speaker entails not paying attention to competing activities that could become a distractor. If the recipient is treating the speaker's talk as the main activity to attend to, then it

---

16 *** denotes p<.001

is plausible that this would make it easier for the recipient to process the information, detect the action(s) delivered in the FPP and react promptly.

The role of speaker gaze in the occurrence of a response to a FPP is consistent with and expands the claims made by Stivers and Rossano (2010) about the use of speaker gaze to pressure for a response to assessments and confirms the claim that this could be a mechanism at play in designing the action that can work as a first action in a sequence. Showing that this mechanism is significant across different types of FPPs is important because it demonstrates the significance of the sequential position of an action. Indeed, as in example 3.1, recipients do not react every time speakers look towards them, and yet, there are specific places (the completion of a turn) and specific types of actions (the ones that can be used as FPPs) after which a response can actually occur. The claim is therefore that speaker gaze can pressure for responses not at any point in time but rather when it occurs in specific sequential environments, in this case during a sequence initiating action. Later in the chapter I will show that this function of pressuring for responses works also during the transition relevance place following a FPP.

The importance of the sequential environment for the interpretation of speaker's gaze is important for the "gaze window" hypothesis by Bavelas et al. (2002). It is true, indeed, that the engagement of mutual gaze affects the likelihood of a response and its timely delivery, but this does not mean that a response could occur at *any* time after the occurrence of mutual gaze. The claim of this chapter is that there are, indeed, interactional moments, based on the organization of talk and action in interaction, on top of which gaze can place an additional pressure for response, without it being, by itself, a symbol of anything. It should be specified here that the sequential environment in which mutual gaze has been investigated in this chapter is different from the one studied by Bavelas and colleagues. Here, gaze behavior has been investigated during sequences of talk, in particular during FPPs, and the effect gaze has on recipient production of a response has been examined. Bavelas et al. (2002) studied what happens during storytellings, here labeled as extended-telling sequences in chapter 2. That chapter showed that different sequential environments (i.e., extended telling sequences and adjacency pair based sequences) require different gaze behavior by the participants.

Two more points need to be emphasized before proceeding to the next part of the chapter. First, consider the finding that the engagement of mutual gaze is an important predictor of the occurrence and timing of a response following a FPP. Mutual gaze is clearly an emergent property of natural conversation that occurs in relation to the specific gaze behavior deployed by each participant in a conversation. Yet, if one thinks about it simply in

terms of attention, it is clear that if recipients are fully attending speakers' turns at talk, recipients are more likely to respond without delay than not (as they might be distracted). Yet what is the speaker doing by looking at the addressee in a dyadic interaction? If it were simply a matter of monitoring the recipient, then it is not clear why speaker gaze would better predict the occurrence of responses than recipient gaze itself. The speaker must have a different reason for looking toward the recipient and I propose that this has to do with the capacity of adding pressure towards the occurrence of a response.

If speaker gaze matters more for the occurrence of responses and recipient gaze more for their timely delivery, the two participants in a dyadic interaction could be looking at each other for very different reasons. And yet, the occurrence of both participants' gaze could still help to secure the occurrence of a response and its prompt delivery. This matters because it reminds us that we cannot study gaze behavior without taking into account all of the participants in the interaction. Focusing only on speaker gaze in analyzing the occurrence of a response or focusing only on recipient gaze in analyzing its timely delivery could lead to forgetting that these two factors are related to each other and that both participants contribute to the final outcome.

Secondly, consider that each participant plays a different role in the development of a sequence of talk. Sacks (1992 [1964-72]: 55, Vol. I) claimed that the questioner has control over the conversation because he has the right to talk again after the answer by the other participant. He later modifies this claim, noticing that this is true only in part, because the recipient might be avoiding answering the terms of the question and therefore might be uncooperative. This observation becomes particularly relevant here if we extend it to adjacency pairs, in general, rather than limiting it to question-answer sequences. The individuals who produce first actions are the ones who initiate the sequence and the ones who know, from the beginning, whether they are designing their action(s) to secure a response. I here subscribe to the model outlined in Stivers and Rossano's work (2010), where canonical and non-canonical actions are distinguished and response relevance is not conceived as normatively on or off, but rather as something scalar, that can be publicly mobilized and therefore increased by stacking particular cues together while delivering a sequence initiating action. It is possible that initial actions are produced and thrown into the public arena to see if they get picked up and dealt with. This is sometimes the case, for example, for noticings, announcements, topic proferrings, or even assessments (see Stivers & Rossano, 2010 on non-canonical actions). A response or reaction to the turn might be appreciated, but if it does not occur, the individual who launched the turn does not treat it as missing. Speakers then know

that by looking at the recipient they can exert additional pressure and display that a response would be expected by the end of the turn. This also means that if the speaker is not looking at the recipient while producing a FPP, the recipient could potentially understand the utterance as not necessarily requiring a response. This is particularly relevant for all those actions such as noticings or assessments, which may end up not being responded to (Stivers & Rossano, 2010). On the other hand, the recipient is the person who produces the response, and therefore the one who decides whether a response occurs and when. This means that the recipient has the most power in affecting what happens after the sequence initiating action. Even in the case in which one would rather subscribe to a conception of question-answer sequences as a normative system, where response is normatively required, the actual compliance with the possible norm (i.e., responding) is usually costly—it takes time, it usually requires thinking about the response and, more generally, it can provide a benefit for the person who asks but does not necessarily provide any immediate benefit to the answer. And the cost of not complying during ordinary conversation is not necessarily high. In other words, the fact that most times complete strangers would respond to each other's questions without securing any immediate advantage remains a remarkable display of humans' natural cooperativeness (see, e.g., Tomasello, 2009). Although responses are often produced, the possibility of not receiving one exists. This explains why looking towards the speaker, displaying full engagement in the conversation and a sort of preliminary commitment to it via gaze can be an important cue to what the recipient of the initial action will do during the transition relevance place: namely respond, usually promptly. Yet it remains a crucial fact that this asymmetry in the development of sequences of action (where the speaker of a sequence initating action can only pursue but the recipient ultimately decides whether or not to respond) is always in place and requires both participants to rely on the norms of conversation but also on cooperation by the other participant. Gaze, by itself, does not determine response. However, speaker gaze is a resource that increases pressure for a response in specific sequential environments (as outlined in the remainder of this chapter). Similarly, recipient gaze is a resource for displaying upcoming "cooperation" towards the completion of the course of action and priority given to the conversation over other activities. Therefore the occurrence of mutual gaze during a sequence initiating action can project the occurrence of a response and its timing, but it does not make a response obligatory. The question is, then, what can speakers do if their first action gets no response, and when in the developing silence is it tantamount to "no response"? The remainder of this chapter addresses these issues and describes the practices that a speaker can implement to pursue a response.

156

## 3.4 Pursuing

In § 3.3.1 I outlined 5 different configurations that describe what happens after the occurrence of a FPP. Scheme 4 is of particular interest here. It shows that speakers can do something to favor the actual production of a SPP when it appears to be missing: namely produce another utterance that works as a pursuit of the missing response.

4.  ＋ A: FPP$_b$ ＋ Silence ＋ A: FPP$_p$ ＋ ........ ▶
    Time                              B: SPP

From the literature, we know that participants can deploy different verbal practices in order to pursue a response. One way participants can deal with silence after FPPs such as invitations, offers, requests and proposals (which can signal upcoming potential rejection) is by producing a subsequent version that addresses a possible inadequacy or trouble in the first version in order to make acceptance the preferred outcome (Davidson, 1984). Different forms of pursuit include subsequent versions of the FPP, following actual rejections, or following weak agreements, or after the absence of response at possible completion points, or, finally, after "other monitor spaces" (Davidson, 1984: 119). Participants monitor what follows the first action and produce subsequent revised versions of it in order to pursue acceptance, which is seen as preferred. In addition, there is evidence that there is a preference for agreement and contiguity in response to questions, and interactants can revise the structure of their turns to pursue that agreement (Sacks, 1987).

Pomerantz (1984b) has described three practices for pursuing a response after an assertion depending on the type of problem with the assertion:

1. If the problem is understood to be an unclear reference or term, then the speaker offers a less opaque reference to replace the problematic one.
2. If the problem is understood to be that some previously assumed shared knowledge is not actually shared, the speaker often tries to explicate the facts and information.
3. If the problem is understood to be that the recipient does not support or agree with the speaker's assertion then the speaker commonly reviews the assertion and possibly modifies it if overstated or inaccurate.

The basic idea is that speakers look for support or agreement to the actions they initiate and if they do not get it then they will try to figure out what went wrong and remedy the problem. The production of some assertion (Pomerantz does not use terms such as FPPs in that work) normally makes relevant uptake and if there isn't any, there may be a problem that the speaker can solve by revising his previous turn and adding something. Most of the observations made in the research outlined above are based on the analysis of phone conversations or audio recordings. As such, the role of visible behavior has not yet been studied. The next step then is to look at what role gaze plays in pursuing responses in face-to-face interactions.

### 3.4.1 Gaze in the Transition Relevance Place

In § 3.3, I showed that the presence or absence of speaker gaze has an influence on whether a FPP gets responded to or not (and partly on its timely occurrence if participants engage in mutual gaze). In particular, I showed that when speaker gaze occurs during the production of the FPP, the recipient is more likely to respond. The question then is what happens if the FPP is produced without speaker gaze and the FPP does not get an immediate response. If speaker gaze adds pressure on the recipient to produce a response, three scenarios should be recurrently observable in face-to-face interactions:

1. A speaker can gaze towards the recipient as an 'increment' (Ford, Fox, & Thompson, 2002; Lerner, 2004; Schegloff, 1996c), with less than 0.2 s between the completion of the FPP and the beginning of the movement of the eyes.
2. If a speaker who does not look towards the addressee during the FPP looks towards the recipient during the transition relevance place (TRP), the recipient might be likely to respond promptly to the FPP.
3. If the recipient does not respond promptly, the speaker may allow for some delay treating the delay as projecting a dispreferred response and if no response is provided, a response may either be further pursued verbally (while holding the gaze) or (more rarely) the speaker may abandon the gaze pursuit.

In what follows I provide evidence of the recurrent occurrence of these scenarios in natural conversation.
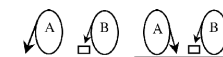
## 3.4.2 Post-Positioned Gaze
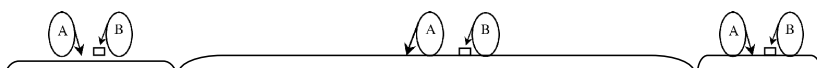
The first piece of evidence of the importance, for speakers, to look at the addressee as a way of obtaining a response comes from an apparent 'misfire' of the looking behavior, more precisely described as the post-positioned deployment of gaze (henceforth, PPG). This can happen in situations in which the FPP is not taken as making a response clearly conditionally relevant (but see example 3.8). Examples 3.7, 3.8 and 3.9 show this pattern. The first two examples show that the speaker starts moving her/his eyes towards the recipient before any silence occurs and the recipient responds promptly. What we are using as a measure is the fact that the speaker's shift in eye direction (and often head direction) occurs before 0.1 second of silence has elapsed. The time it takes for the eyes to reach a focus on the other participant's face depends on the speed at which the participants move their eyes and head. We are interested in when a shift in eye direction is projectable.

In example 3.7 two friends have just decided to stop studying for the day. A says she is really tired and has a headache (she adds "as well" because B has mentioned having a headache earlier on). Our focus is A's looking up towards B at line 11.

**(3.7) 2GSTUDYING-testa 29:07**

```
07      (0.7)


08 B:   Okay quindi    (0.3)
        Okay therefore
        Okay therefore (0.3)


09 A:   .h No io veramente son scarburatissima ho    un mal di testa
           No I   really   am  out of steam  have.1s a  bad of head
        .h No I am really out of steam I have a headache


10      anch''io
        also I
        as well
```

11
(0.1)  (0.1)

(0.2)

12 B: *Anch'io c' ho      mal di testa    ((annuendo))*
Also I  cl have.1s bad of head
**I have a headache too ((nodding))**

13
(1.1)    (0.3

(1.4)

14 A: *Poi  ho      detto sara'  il raffreddore sara': il-*
Then have.1s said  will be the cold      will   the
**Then I said it will be the cold it will be: th-**

15
*i   postumi dell'influenza, (0.5)*
the effects of   flu
**the side effects of the flu, (0.5)**
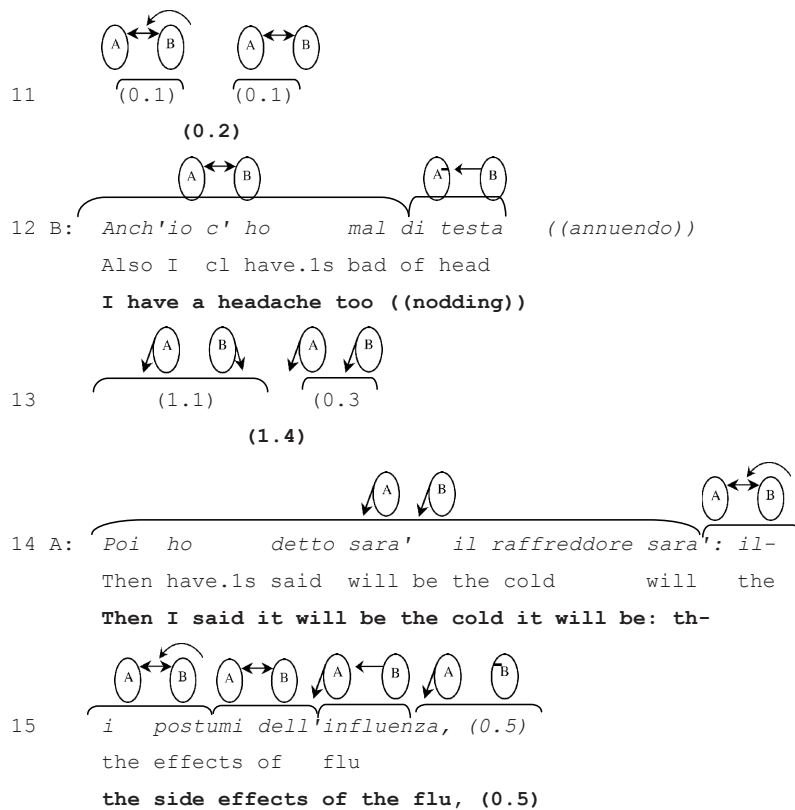
In this example, after 0.1 seconds of mutual gaze, B responds to A's "headache" statement by saying that she has a headache as well (she had already mentioned this few minutes earlier in the conversation). A starts moving her eyes towards B immediately after the completion of her turn at line 10, spending 0.1 seconds to reach the face of the interlocutor. In one sense, the TCU "ho un mal di testa anch'io" (I have a headache as well) could be taken as an announcement about her current physical condition.  In another, it can also be interpreted as an account for her previous statement "io sono veramente scarburatissima" (I am really out of steam). Both are produced with no speaker gaze to the recipient and yet we see A's gaze move towards B immediately after the completion of line 10 (I have a headache as well). The recipient then promptly responds to the announcement about the headache. The announcement about the headache at lines 9-10 already shows that A knows that B has a headache as well (the words "anch'io", me as well, indicate it), yet B repeats it and nods. The occurrence of nodding makes sense only in a situation in which B knows that A is looking at her (and indeed B can perceive A's gaze because they engage in mutual gaze during line 11). Independently of whether we believe that the verbal response at line 12 was occasioned by

160

the speaker gaze towards the recipient at line 10, the occurrence of the nodding can be perceived as a response because the speaker is looking towards the recipient during the TRP at line 11. In this example we therefore see that A starts shifting her gaze towards B while uttering the last sound of her turn, and B responds timely.

Example 3.8 shows a similar pattern. In this example, two friends are looking at pictures sitting side by side. A is showing B some pictures he took in Stockholm during a vacation and at line 1 he points towards a boat in one of the pictures. Our focus is B's looking up towards A during the silence at line 4 and A's immediate response at line 5.

**(3.8) 2PCOMP-barca 12:15**

```
01 A:  hh Questo e' l'ostello  in cui   ero ((indicando una foto))
           This   is the hostel in which be.1s
       hh This is the hostel where I was ((pointing on a picture))

02     (2.4)
```



```
03 B:  Cosa vuol   dire=sulla  barca
       What want.3s say  on the boat
       What does it mean=on the boat
```
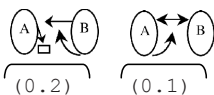


```
04     (0.2)      (0.1)
           (0.3)
```



```
05 A:  Si'
       Yes
       Yes

06     (1.0)

07 A:  E' un ostello meta' sulla  barca meta' sulla  terraferma
       Is a  hostel  half  on the boat  half  on the land
       It is a hostel half on a boat and half on the land
```

When B produces the FPP at line 3 in which he initiates repair by asking A for a clarification, both B and A are both looking at the picture on the desk. At the completion of B's turn at line 3, B turns and looks towards A, who looks up towards B and then responds (line 5).

In the last two examples we have seen that a participant can start moving her/his eyes towards the recipient before any noticeable silence occurred after the completion of the FPP (notice that at line 4 of example 3.8, B shifts his gaze towards A immediately and 0.2 s is the time it takes him to get his eyes from looking at pictures to looking towards A, not the time he waits until shifting his gaze). When the speaker looks towards the recipient, s/he meets the eyes of the recipient and they engage in mutual gaze. At that point the recipient promptly responds to the FPP. Often these slightly delayed responses are actually preferred responses to the FPP (e.g., acceptance of an offer). This suggests that once a first action is produced without speaker gaze towards the recipient, the speaker can start looking towards the recipient immediately at completion of the turn, so that the looking towards the addressee may be perceived as a post-positioned deployment of gaze or as a kind of increment (Ford et al., 2002; Lerner, 2004; Schegloff, 2001). According to Lerner (2004), increment initiators such as "to", "for", "because", etc. might occur in situations in which the turn is syntactically and prosodically complete, yet pragmatically there seems to be some bit of information missing. This missing information is then provided by the increment. Moreover, a piece of talk is *incremental* if it displays its connection to the previous turn at talk (e.g., "we went straight to visit Ted (1.0) *from the office*").

The interesting feature of this incremental post-positioned gaze (PPG) is that the element that is added is not talk, but rather the specification that the last TCU was designed to be responded to. It is crucial that the speaker's gaze to the recipient occurs before any audible silence because it means that its occurrence is not due to a perceived lack of response (i.e., there is no gap at the moment the looking up starts), but rather to pre-empt other possible problems (e.g., not getting a response to). In this respect, a PPG may appear to do an interactional job similar to increments (i.e., providing further pragmatic specification in terms of the action the turn is implementing). However, the impossibility of displaying a connection to the previous turn through gaze (apart maybe from its timely occurrence on the last sound of the turn) invites a more careful categorization of this practice and therefore the current labeling will be post-positioned gaze.

With this established, although speakers may look towards recipients as soon as they enter the transition relevance place, this does not necessarily secure an immediate response. Example 3.9 is a case in point. In this example, two friends are having a conversation facing

each other sitting on two couches. B has just asked A to help him with some recipes for preparing food for a boy-scout camp and at line 1 A asks about the number of people B will have to cook for. The focus is on the participants' gaze behavior during the silence at line 9.

**(3.9) 2PPLAN-piu 05:56**

```
01 A:   Quante    sono le  persone
        How many are   the people
        How many people are there


02      (0.4)


03 B:   Trenta
        Thirty
        Thirty


04      (0.1)


05 B:   °Cinni°
         Kids
        °Kids°


06      (0.5)
```
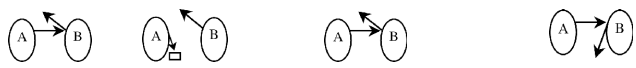


```
07 B:   Cinni, Ragazzi    ((A sta guardando un foglio con il menu'))
        Kids    adolescents
        Kids, Adolescents ((A is looking at piece of paper with menu'))
```
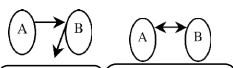


```
08 A:   Piu',
        Plus
        Plus,
```



```
09      (0.2)   (0.6)
            (0.8)
```

163
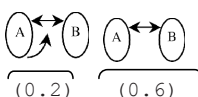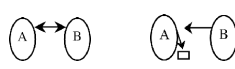
```
         A→B      A  B      A→B      A→B
                                             
10 B:  Piu'::::: tre   noi: e  loro i  capi  saranno: (0.4)
       Plus       three us   and them the bosses be.3p
       Plus::::: three of us: and them the bosses will be: (0.4)


         A→B    A↔B
11     quattro o cinque
       four     or five
       four or five


         A↔B
12     (0.3)

         A↔B      A←B
13 B:  Quaranta persone
       Forty     people
       Forty people
```

As in example 3.7, the speaker's gaze provides additional evidence for the recipient that this is a turn that should be responded to. In this example the speaker (A) of the FPP at line 8 looks up towards the recipient, who was already looking at him, immediately after the completion of the FPP. While producing line 8, A looks at the menu B has to prepare for the boy-scouts. Line 8 can be considered a FPP because, with it, A asks B to specify how many more people will be present in addition to the boy-scouts, although it does so in a very elliptical manner. The speed at which the speaker looks up towards the recipient makes it look like a natural bit of behavior belonging to the previous TCU (i.e. at line 8). Nonetheless, this time the response is delayed. The reason for this delay can be found in the delivery of the response: B does not know exactly how many more people will be present at the boy-scout camp. So even though they are looking at each other, the recipient starts responding only after 0.6 seconds of mutual gaze and while producing the answer he looks away, in the middle distance, displaying " thinking" (C. Goodwin, 1987; M. H. Goodwin & Goodwin, 1986), and he does so also in his talk with the sound stretches on the "u" of "più" (plus) and later on the words "noi" (we) and "saranno" (will be). A looks away briefly while B conveys that he is "thinking" (line 10) and resumes looking at him soon after B starts producing the

response. Finally, notice also that it is not self-evident that line 8 is a complete turn when it is delivered.

To summarize, in example 3.9 we have seen that even though speaker gaze has been post-positionally added to the end of the turn and mutual gaze is engaged, the response is nonetheless delayed because the recipient does not immediately know the answer. At the same time, this does not contradict the claim that speaker gaze pressures recipients for a response, as the delay can be accounted for contingently. The next section shows a case that could be taken as a variant of the practice outlined in this section or a member of the second practice that will be presented in § 3.4.4.

## *3.4.3 A Continuum of Pursuits?*

Example 3.10 shows something in between a post-positioned deployment of gaze and a full-blown gaze pursuit. In theory, a gaze pursuit could be described as another occurrence of PPG, as the gaze is post-positioned with respect to the FPP. However, the occurrence of speaker gaze before the occurrence of silence is interactionally different from the occurrence of speaker gaze after some silence has occurred. In the first case, the turn might be "pragmatically" incomplete, but no delay in responding has occurred yet, while in the second case the occurrence of speaker gaze might be dealing with the lack of response, as the absence of a response after a bit of silence is a noticeable phenomenon. It will be shown that the main feature that distinguishes these two practices is the timing of speaker's gaze towards the recipient.

In the following extract, two friends are talking about the exams for which they need to prepare. At lines 1-2 A asks B whether she has talked to someone about her exam and what to study. Our focus is the gaze behavior during the silence at line 6.

**(3.10) 2GSOFA-chiama 31:08**

```
01 A:  Ma   hai     sentito qualcuno? Quella tipa la'
       But have.2s heard   someone   that   type there
       But did you hear from someone? That woman/girl there
```

```
02      l'hai      ch- l' [hai chiamata?]
        her have.2s ca- her have.2s called
        did  you  ca-      [did you call her?]

03 B:                      [No non l'ho]      chiamata
                            No not her have.1s called
                           [No I did not] call her
```



```
04      (1.1)
```
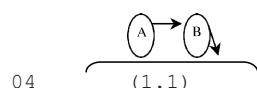


```
05 B:   La chiamo?
        Her call.1s
        ((Should)) I call her?
```



```
06      (0.1)  (0.2)  (0.2)
               (0.5)
```
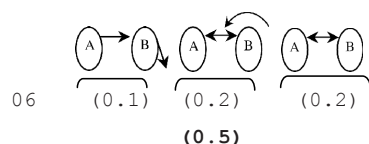


```
07 A:   [C h i a m a l a ]
         Call her
        [C a l l  h e r]
```



```
08 B:   [Magari domani   la] chiamo cosi' mi faccio   raccontare
         Maybe  tomorrow her call  so    me make.1s tell
        [Maybe tomorrow] I call her so she tells me ((about the exam))
```

Immediately following the first FPP at line 1, A refines the reference to the person B should have talked to by referring to "la tipa" (that woman/girl) in her second FPP. Producing a second more specific question immediately after a more general one without any audible silence in between is a pattern that appears to be the verbal equivalent of the gaze practice just described in § 3.4.1: that is, a way of facilitating the occurrence of a response to the FPP. Interestingly, it would not be appropriate to talk of a verbal pursuit in this case because there is no indication that the first FPP will not be responded when the second more specific

166

version is offered. Rather, it is clearly an attempt to clarify what the question is targeting and therefore an attempt to facilitate an appropriate answer. The question receives a dispreferred answer in overlap at line 3, and after some silence, B asks whether she should actually call that person. The question here seems rhetorical given that A has already implicitly suggested B should talk to this woman. B gazes towards A soon after the end of the FPP, but this time a brief silence occurs before the beginning of the movement (0.1 s). The occurrence of some silence before the speaker of the FPP shifts her gaze towards her addressee's face is what makes 3.10 a different case than the ones presented in § 3.4.1.

The question is: is B's gaze movement during the silence at line 6 being prompted by this gap or is it a misfire of the PPG practice presented in § 3.4.1? Although difficult to assess with some certainty, given that the main criteria for distinguishing the two practices is the noticeable absence or not of a response before the occurrence of the looking up towards the recipient, there are other important elements that show the link between the incremental use of speaker gaze and the pursuing use of speaker gaze. Specifically, the participants engage in mutual gaze during the silence following the FPP and the FPP is then promptly responded to. The relationship between the timing of speaker gaze, and the timing and occurrence of a response is the focus of the following section.

### 3.4.4 Pursuing Responses With Gaze

The second piece of evidence that speaker gaze is instrumental in obtaining a response from a recipient comes from the fact that gaze can be used to pursue a response when one is missing. The timing of speaker's gaze towards the recipient is different from the cases of PPG shown in examples 3.7-3.9. If the gaze is used to pursue a response, the looking up occurs only after some silence has occurred, and therefore after a noticeable gap in the sequential development of the course of action. However, when a speaker does gaze towards a recipient, a response is generally promptly provided. Thus, the timing of the response is an important piece of evidence because the response was absent before the speaker's looking up. Examples 3.11 and 3.12 illustrate this pattern.

In the following example, there are two university students sitting side by side in front of a desk, looking at pictures together and commenting on them. When this extract starts, the guest (B) has just recognized a famous actor in one of A's pictures. A had previously introduced the picture as "the evening of the concert" (in which A played the piano). Our focus is the gaze behavior during the silence at line 11.

**(3.11) 2PCOMP-mangiato 16:32**

```
01      (4.0)


02 A:   Tipo alla   mano poi  sai
        Guy  at the hand then know.2s
        Guy very easy going then you know


03 B:   Si' si'
        Yes yes
        Yes yes
04       (2.0)


05 B:   Ma si vede anche adesso perche' e': [c'era in una
        But cl. see also now    because is  cl. was in a
        But you can see it even now because he is: [he was in a
                                                    [B starts turning picture


06      trasmissione l'  altro giorno [.hhh
        tv show      the other day
        tv show the other day          [.hhh


07 A:                                   [hhh
                                        [hhh
```
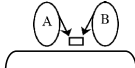


```
08 B:  Ma   ha     mangiato li'   con  voi?
        But has.3s eaten     there with you
        But did he eat there with you?
```



```
09           (0.4)


10 A:   hhh h ((looking at pictures))
```

11      (0.5)  (0.1)  (0.1)

        (0.7)


12 A:   E:hm n[o dopo sono andati a  far  la  cena  dopo li   ho
        Ehm  no  then were gone   to make the dinner after them have.1s
        **E:hm n[o later they went for dinner and afterwards**
                [B turns pictures back to the picture with the actor


13      raggiunti anch'io
        reached   also I
        **I joined them too**
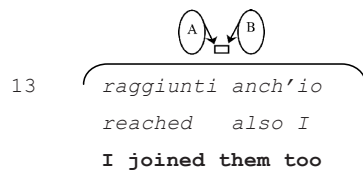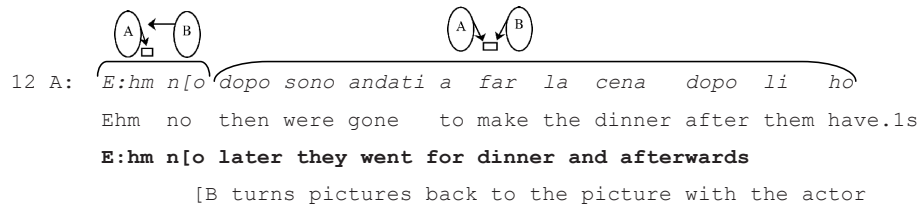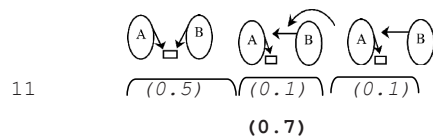

In this excerpt, the speaker of the FPP pursues a response by looking towards the recipient during the silence and this gaze pursuit obtains a prompt response. Let's take a closer look at how the gaze pursuit occurs. B's question at line 8 is responded to after 1.5 seconds of silence (lines 12-13). After having acknowledged and conveyed independent knowledge[17] of the "easy going" nature of the famous Italian actor (see lines 5-6), B asks whether the actor also had dinner with A and his friends (line 8). During the production of the turn at line 5 the speaker (B) continues to turn the pictures then stops to look at two new ones. He continues talking about the previous picture but at the same time takes a closer look at the new ones. At line 7, A produces a laugh token while looking at the new picture (that is later described by A as representing a dinner with his relatives and is not related to the one with the actor). When B asks the question at line 8 both of them are oriented towards this new picture. During the following gap (lines 9-11) A produces another laugh token (line 10) while continuing to look toward the new picture. This laugh token is not interpretable as a response to the question. A does not produce any further talk, and, as a result, after another 0.5 seconds of silence, B begins to turn towards A. At this point, A then provides the relevant next turn: an answer to

---

17 For a more detailed account of how and why participants would claim independent epistemic rights in assessing someone's personality or behavior, see Heritage & Raymond (2005) and Stivers (2005).

the question.[18] A starts talking before B turns the pictures back to the previous one. B's head turning and his looking towards A become an efficient way of pursuing a response that was missing.

Example 3.12 shows another pursuit of a response achieved by looking up towards the other participant. Here two young women have just finished studying together and are beginning to chat. A begins by noticing that B tends to scribble a lot. Our focus is A's gaze during the silence at line 3.

**(3.12) 2GSTUDYING-colori 30:20**

```
01      (1.6)
```
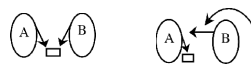
```
02 A:   Tu  colori   sempre
        You color.2s always
        You color all the time
```

```
03     (0.8) ((enacting coloring)) (0.1)  (0.1)  (0.1)
                   (1.1)
```

```
04 B:   Io e'    vero. L' ho      notat[o
        I  be.3s true  It have.1s noticed
        Me it is true.        I notic[ed it

05 A:                            [Sempre^h
                                 Always
                                 [Always^h
```

A's observation about B is what Labov & Fanshel (1977) call a 'B-event statement' and makes confirmation relevant (Heritage & Roth, 1995). However, this observation does not

---

receive a response until A looks up towards the recipient, after having mimicked B's scribbling on a piece of paper during the silence. B is not directly looking at A's face but is nevertheless less than 1 meter away, [19] and so arguably, B can easily perceive A's shift in head and gaze direction. B's response at line 4, though not type-conforming (Raymond, 2003), confirms A's observation and therefore appears to be preferred (see, e.g., Atkinson & Heritage, 1984; Pomerantz, 1984a; Sacks, 1987), yet it is produced in a dispreferred manner being delayed and occurring only once pursued.

In these two examples we have seen that by orienting towards the other participant during the gap following a FPP the speaker obtains a relevant SPP promptly. More specifically, in a situation in which the participants are not directly looking at each other, the speaker of the FPP turns towards the recipient who then answers the question. The only thing that changes in terms of bodily orientation and posture is the speaker's head and eye orientations. Immediately following these shifts in orientation, the recipient produces a sequentially relevant next (the SPP) in order to advance the sequence (and the preference of the SPP often cannot account for the previous delay in answering). The claim is that gazing towards the recipient in this sequential environment is a way of pursuing uptake. Participants in interaction orient to gaze in this position as a recognizable practice for pursuing response by immediately producing the appropriate SPP. Next, I examine deviant cases in which a participant implements the practice of pursuing a response through gaze and yet s/he fails to obtain a timely response or any response at all.

### 3.4.5 Gaze Pursuits That Fail or Get Very Delayed Responses

The fact that a participant can pursue a response by looking towards the recipient during the transition relevance place following a FPP does not mean that this attempt is *always* successful or that the response will *always* be produced promptly. In what follows, I show a series of examples in which the speakers of a FPP perform a gaze pursuit but are ineffective in achieving a response.

In 3.13, B pursues A's response with gaze, however, A does not respond. That said, in this case we nonetheless see an orientation towards the occurrence of a response soon after a

---

19 The distance between the faces and the bodies of the participants is always less than 1 meter. If the recipient is oriented towards the speaker, although looking down, s/he can perceive the change in gaze direction by the prior speaker because of peripheral vision. § 3.4.5 will provide examples in which the recipient cannot perceive the speaker's gaze direction for various reasons (for example because turned towards a computer screen), and in which a gaze pursuit cannot be responsible for the timing of the response.

gaze pursuit as the default and a lack of a prompt response as projecting a dispreferred response to come or at least a problematic one. When a response is not immediately forthcoming, the speaker of the FPP can treat the silence as prefiguring disagreement and back down. In this example A is teaching B how to make a recipe, and at lines 2-3, B asks whether adding some ham would improve it. Our focus is the gaze behavior during the silence at line 4.

**(3.13) 2PPLAN-prosciutto 09:17**
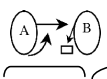
```
01 A:  Pero' non so      la  pan[na io la metterei
       But   not know.1s the cream  I  it put.1s
       But I do not know [I would put the cream
```
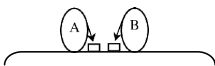
```
02 B:                              [Ma non ci   sta       bene
                                    But not cl. stay.3s well
                                    [But doesn't it fit well
```

```
03     anche il  prosciutto
       also   the ham
       also the ham ((in it))
```

```
04        (0.9)    (0.2)   (0.3)
                  (1.4)
```

```
05 B:  Di-  Vabbe' io [devo    fare quell che   mi dicono quindi =
       Di-  Alright I  must.1s do   that  which me say.3p  so
       Di-  It's alright [I must do what they tell me ((to do)) so=
```

172

```
                    A  ⟶  B              A  ⟶  B
                       □                     □
06 A:        [si'      si'      si'      si'      ci =
             Yes       yes      yes      yes      cl.
             [Yes      yes      yes      yes
```

```
                  A  B
                  □  □
07 B:    =si' tranquillo]
          yes relax
         =yes don't worry]
```

```
                A  B
                □  □
08 A:    sta      be]ne anche il  prosciutto
         stay.3s well  also  the ham
         the ham fits we]ll as well
```

When A produces the FPP at lines 2-3, he is looking down at the entire menu B has to prepare, while B is taking some notes about the recipe. After 0.9 seconds, B looks up towards A and sustains gaze towards him for 0.3 seconds before starting to back down (line 5). When B pursues a response with his gaze, A continues looking down. The delay in responding, together with the lack of recipient gaze towards him after a total of 1.4 seconds of silence, is interpreted by B as prefiguring A's disagreement. In anticipation of that disagreement, B backs down from his suggestion of improving the recipe by adding ham and says that he has to do what the organizers tell him to do, implicitly indicating that his suggestion was perhaps not appropriate. However, A actually produces a response to B's initial suggestion regarding the ham and confirms that adding ham would be a good idea. The way the answer at lines 6-8 is produced shows that A has been thinking about the suggestion by B and is willing to acknowledge that this could be a good variation. Given that A is the expert and B has repeatedly manifested his incompetence in terms of food recipes, it is clear that this stretch of interaction is highly affected by the epistemic rights in play between these two individuals. In particular, B seeks confirmation of the appropriateness of a variation to the recipe, but this turns out to be something that A has never considered. By delaying his acceptance of B's suggestion, A re-asserts his expertise by enacting a thoughtful consideration of the apprentice's suggestion. This example shows that when a speaker pursues a response through gaze, s/he expects the recipient to respond promptly, and if this does not happen, the speaker

displays an orientation towards this delay as projecting further problems in dealing with the original FPP.

A second deviant case is shown in example 3.14. Here, although a response is pursued via gaze, the response occurs only after a noticeable delay. However, the delay is accounted for by the dispreferred nature of the response. In this excerpt, a couple is sitting at a table eating and chatting. Before the beginning of this excerpt, A accused her boyfriend B of having almost entirely eaten a chocolate rabbit she had bought for dinner guests that evening. After some denying, he eventually admits having eaten the chocolate rabbit. She accepts his confession and closes the sequence with a sequence closing third "alright" (Beach, 1995; Schegloff, 2007b). Then they re-engage eating during the silence at line 1. Here our focus is the participants' gaze during the silence at line 3.

**(3.14) 2PLUNCH1-sorpresa 9:12**

```
01      (2.5)
```



```
02 A:  Io ci son rimasta male che non c' era  la sorpresa.
       I  cl. am left    bad that not cl. was the surprise
    I was disappointed because there was not the surprise.
   ((Inside the chocolate rabbit))
```



```
03        (1.0)    (0.2)    (1.3)
                  (2.5)

04 ((B makes gesture + facial expression like "what a pity", Fig. 3.1))
```

174

Figure 3.1. Frame representing line 4
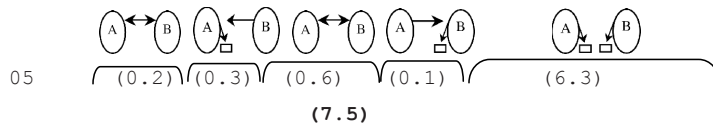
```
                     A↔B  A←B   A↔B   A→B        A B
                            □               □      □□
05      ⌒(0.2)⌒⌒(0.3)⌒⌒(0.6)⌒⌒(0.1)⌒⌒      (6.3)    ⌒
                            (7.5)
```

06 A:   *Beh insomma Angela ha vinto l'  Erasmus ad Heidelberg (0.2)*
        Well in sum Angela has won   the Erasmus at Heidelberg
        **Well Angela won the Erasmus for Heidelberg (0.2)**

A's turn at line 2 is both a complaint (not about B's actions) and an announcement (i.e. there was no surprise in the chocolate rabbit). A produces the talk at line 2 without looking at B and looks at him only 1 full second into the transition relevance place. They engage in mutual gaze and sustain it for 1.3 seconds before B produces a facial expression and a mildly empathetic gesture that can be glossed as "what a pity", which seems quite ironic. They engage in a back and forth of looking towards each other, probably due to the ironic nature of B's response, until both of them look down and then re-engage eating for some seconds.

Contrary to other examples shown in this chapter, B's reaction to the gaze pursuit is not immediate. It is actually quite delayed, as if he was resisting answering. If we consider the sequential environment in which it occurs we can find an appropriate account for this. B has just been accused of having eaten the entire chocolate rabbit without thinking about A and her guests and he has admitted it after having initially denied it. After 2.5 seconds, A seeks B's affiliation with her complaint for the absence of a surprise inside the chocolate rabbit that had just been the cause of an argument. The delay, after the sustained mutual gaze, and the teasing component readable in his facial expression display his unwillingness to easily affiliate with such a complaint. He appears to be delaying conveying his stance towards the other participant. In this respect A's looking up and down again during the following silence is a way of monitoring what B meant through his enactment: in particular whether it is something she should accept or take up in a specific way. The fact that A looks

back and forth at him after the visible response, rather than simply withdrawing gaze and closing the sequence, provides further evidence for this claim.[20] Example 3.14 thus provides some evidence that a participant orients towards a delay in the occurrence of a response after a gaze pursuit as a marked action and that pursuits do make relevant a timely response.

Although most gaze pursuits are responded to, occasionally this is not the case. Example 3.15 shows a gaze pursuit that is abandoned after no response is provided in the 1.3 seconds that follows the pursuit. This shows both the affordances of this visible behavior (i.e., gaze towards a recipient can be sustained and abandoned) and an orientation towards a lack of response within a certain amount of time as a probable display that no answer will be provided to the FPP. In this excerpt, B has just offered A the use of her helmet for the evening. At line 2 B warns A that she will get very upset if he loses her helmet. Our focus is their gaze behavior during the silence at line 3.

**(3.15) 2PLUNCH1-perdi 11:32**
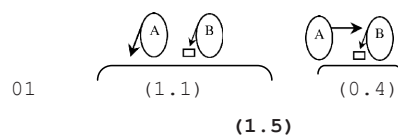
```
01        (1.1)        (0.4)
               (1.5)


02 B:  Se mi perdi   il  casco   vedi
       If me lose.2s the helmet see.2s
       If you lose my helmet you will see ((the consequences))


03   (0.2)  (0.2)  (0.2)     (1.3)       (0.7)
                  (2.6)


04 A:  Te  lo ripaghero' il  casco
       You it repay.1s   the helmet
       I will repay it the helmet
```
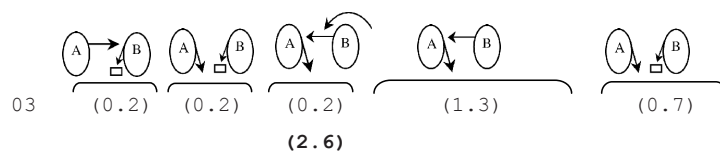
_____

20 See chapter 4 for evidence that A's behavior after B's response displays that A perceives the sequence as incomplete at that point.

```
05    (1.2)

06 B: Giallo ((il colore del casco))
      Yellow
      Yellow ((the color of the helmet))
```

Here, the recipient of the FPP (A) looks at B during its production and briefly into the following silence, yet he does not react to the warning/threat that B produces at line 2. After some silence, B looks up towards A to pursue a response and keeps looking at him for 1.3 seconds but he keeps looking down at his dish (line 3). After this sustained look and without accomplishing mutual gaze, B lowers her gaze for 0.7 seconds. At this point, however, A responds to the FPP by saying he will buy her a new helmet if he loses it. In this case, there is no reason to believe that A actually responds to the gaze pursuit, given that 2 seconds passed after B looked up towards him. A responds with such a delay because he perceives B's comment at line 2 as a threat and also a possible accusation that he normally loses things. If the FPP is perceived as being face threatening, a delay in responding and the seriousness of the response can be a display of his stance towards that action.

In this final example, the speaker does not abandon the gaze pursuit as in 3.15 above, but the recipient nevertheless does not look back but instead maintains focus on a computer screen. The dispreferred response that is eventually produced is delayed but is also occasioned by the gaze pursuit. In example 3.16 two friends are preparing an exam and B is talking about what A can find in a book they have to read for the exam. At line 6 A asks B whether B has read this book entirely. Our focus is their gaze behavior in the silence that follows.

**(3.16) 2PEXAM-letto 45:49**

```
01 B: pero' il  Treetti va  bene per radiografia.
      but   the Treetti goes well for radiography
      but Treetti is good for radiography.

02    Per ecografia  non dice niente.
      For ultrasound not says nothing
      Concerning ultrasounds it does not say anything.
```

03  (0.3)



04 A: ⌐ *mm*

   mm

   **mm**



05  (0.2)



06 A: ⌐ *Ma  l'hai  letto tutto il  Treetti¿*

   But it have.2s read  all   the Treetti

   **But have you read it all the Treetti¿**



07  (0.2)  (0.2)  (0.3)
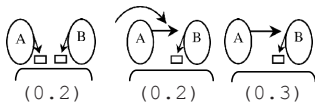
      **(0.7)**



08 B: ⌐ *Macche'* ((scuotendo la testa))

   *Not al all*

   **Not at all ((with a headshake))**



09  (1.6)



10 B:  *No del Treetti devi fare le prime: Cioe' le prime pagine*

   No of Treetti must do the first I mean the first pages

   **No  of Treetti you must do the firs:t I mean the first pages**

11     *che   poi   sono  (1.0)   (0.6)    (0.2)  quelle  che:  son*
       that  then  are                         those    that  are
       **that then are**         **(1.8)**         **those tha:t are**

12     *schematizzate   sugli  appunti  di  Marini.*
       schematized      on  the  notes  of  Marini
       **schematized on Marini's notes.**

13     (7.0)
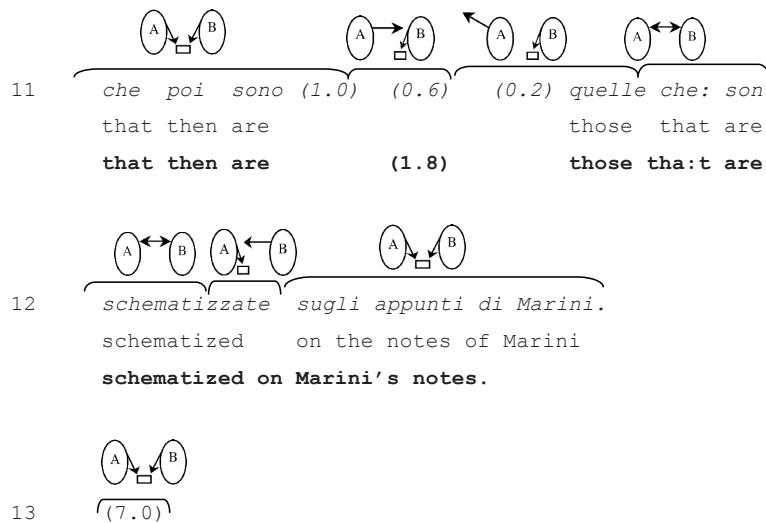
At line 6, A produces a FPP that does not receive an immediate response by B. After a bit of silence, A turns and looks towards him, though he is actually looking at a computer screen, surfing the Web. A sustains his gaze for 0.3 seconds until B produces a dispreferred response (line 8) and then explains what A has to study in the book (lines 10-12). In responding in this way, B first deals with the format of the question (line 8) and then deals with the action implemented by the turn at line 6: that is, the question of what A has to study from that book. The timing of B's response, however, does not seem to be affected in any way by the gaze pursuit, which B has very likely not even perceived. Rather, the timing is more likely affected by the fact that the response is dispreferred and that B is involved in a competing activity. Notice that B turns and looks towards A during his response at line 11-12, yet he produces the first part of his more detailed answer while continuing looking at the computer screen. The main point displayed through this example is that the timing of the recipient's response is not in fact due to the occurrence of a gaze pursuit, even though it might appear so. This example shows again that the speaker of a FPP can implement the practice of gaze pursuit but the contingencies of the interaction (the recipient looking at a computer screen rather than towards the speaker) can make it ineffective.

      In this section we have seen multiple examples of gaze pursuits that are not immediately successful. This invites further investigation about the actual timing of delayed responses, the timing of pursuits and how quickly participants tend to respond to them when they occur. The following section provides quantitative answers to these questions.

*3.4.6 Types of Pursuits*

　　The database was re-coded considering the occurrence of gaze pursuits before a response occurs, in order to examine how often non-answered responses are pursued, how long speakers wait before pursuing a response and whether pursuing a response normally occasions a prompt response. This re-coding produced the results presented in Table 3.7 (an updated version of Table 3.2).

*Table 3.7* Distribution of response patterns to first pair part turns including gaze pursuits.

| Possible Patterns | Number of Instances |
|---|---|
| 1) FPP → SPP | 125 (29.9%) |
| 2) FPP → Overlapping SPP | 53 (12.7%) |
| 3) FPP → (silence) → SPP | 157 (37.6%) |
| 4) FPP → PPG or verbal increment | 15 (3.6%) |
| 5) FPP → (silence) → Pursuit | 48 (11.5%) |
| 6) FPP → (silence) → Ø | 20 (4.8%) |
| TOTAL | 418 (100%) |

The first noticeable difference from Table 3.2 is the addition of a pattern that I called the occurrence of a verbal increment or of a PPG after the possible completion of the FPP that facilitates the occurrence of a response. This corresponds to the gaze patterns displayed in § 3.4.2 and to any similar modification done verbally, as shown in example 3.8. From now on both these practices will be called "incremental modifications". This pattern is observable in 3.6% of the cases. The number of unresponded turns without pursuits slightly decreased compared to Table 3.2 because in two of them the speaker had actually produced a gaze pursuit that did not get a response. Moreover, there is a decrease in the number of turns that get a delayed response without any pursuit. 15% of all FPPs (63, given by the sum of the ones with an incremental modification and the ones that got pursued) get modified either through gaze, verbally or with the occurrence of both to facilitate a response and only around

5% of FPPs are not pursued and are not responded to. Table 3.8 shows how the 63 FPPs are modified.

*Table 3.8* Distribution of modalities through which a FPP gets modified.

| Modality | Incrementally Modified | Pursued | Total |
|---|---|---|---|
| Verbal Only | 4/15 (26.7%) | 12/48 (25%) | 16/63 (25.4%) |
| Gaze Only | 9/15 (60%) | 16/48 (33.3%) | 25/63 (39.7%) |
| Gaze + Verbal | 2/15 (13.3%) | 20/48 (41.7%) | 22/63 (34.9%) |

Table 3.8 shows that nearly 2/3 of FPPs that are incrementally modified are modified solely through the looking up by the speaker towards the recipient, while a response is pursued with gaze alone in only 1/3 of the FPPs. In general, combining the two possible ways of modifying a FPP, a FPP gets modified around 25% of the time using verbal resources only, 40% of the time using gaze only (see total in Table 3.8) and around 35% of time by producing another TCU while holding the gaze towards the recipient. FPPs with speaker gaze towards the recipient constitute approximately 75% of all cases (47/63). Thus, most of the time when a speaker tries to pressure a recipient for a response, s/he will do so by looking towards the recipient.

If we now focus on the effectiveness of the modifications in terms of obtaining a response, we can see in Table 3.9 that 80% of the incrementally modified FPPs get responded to and 72.9% of the ones pursued get responded.

*Table 3.9* Effectiveness of modification on facilitating a response.

| Type | Responded | Not responded |
|---|---|---|
| Incrementally Modified | 12/15 (80%) | 3/15 (20%) |
| Pursued | 35/48 (72.9%) | 13/48 (27.1%) |

Of the 16 FPPs that were not responded to after the first modification, only 2 are not responded to at all (in both cases response had been pursued only with speaker gaze). In 14 instances a response is pursued further. Table 3.10 shows how responses are pursued the first and the second time.

*Table 3.10* Verbal vs. gaze modification of multiply pursued FPPs.

| Modality | First Pursuit | Second Pursuit |
|---|---|---|
| Verbal Only | 4 (28.6%) | 1 (7.1%) |
| Gaze Only | 7 (50%) | 4 (28.6%) |
| Gaze + Verbal | 3 (21.4%) | 9 (64.3%) |
| TOTAL | 14 (100%) | 14 (100%) |

Table 3.10 shows that of the FPPs that got modified/pursued more than once to facilitate a response, half of them had been initially pursued only through gaze and only a small number had been initially pursued both visually and verbally. If we then look at how a turn gets pursued a second time, we see that most of the time a speaker produces a verbal pursuit while looking at the recipient (64.3%), and, in general, it appears to be extremely rare (n=1/14) that a response is pursued a second time without speaker gaze towards the recipient. These results underscore the importance of gaze in pressuring for a response, although gaze alone is not necessarily the most effective modality to obtain a response (as shown in § 3.4.5).

If we consider all of the possible combinations of modalities through which the pursuit gets done, a remarkable picture emerges. Table 3.11 shows that only 4 out of 9 possible patterns actually occur in the corpus.

*Table 3.11* Patterns of pursuits in terms of modality when a response is pursued twice.

| 1st PURSUIT → 2nd PURSUIT | NUMBER OF INSTANCES |
|---|---|
| 1. Verbal Only → Verbal Only | 0 |
| 2. Gaze Only → Gaze Only | 0 |
| 3. Verbal Only → Gaze Only | 4 (28.6%) |
| 4. Gaze Only → Verbal Only | 1 (7.1%) |
| 5. Verbal Only → Verbal + Gaze | 0 |
| 6. Gaze Only → Verbal + Gaze | 6 (42.9%) |
| 7. Verbal + Gaze → Gaze Only | 0 |
| 8. Verbal + Gaze → Verbal Only | 0 |
| 9. Verbal + Gaze → Verbal + Gaze | 3 (21.4%) |
| TOTAL | 14 (100%) |

182

The results in Table 3.11 show that if a response had been first pursued using only one modality (e.g., only verbally or via gaze), the second time it will be either pursued using the other modality (e.g., pattern 3 or 4), or by using a combination of them (e.g., pattern 6 or 9). It is never the case that a failed pursuit using one modality is simply repeated using that same modality. Moreover, it is never the case that if the initial pursuit had been delivered using both modalities (i.e., verbally and with gaze), then the second pursuit is launched only relying on a single modality. This is particularly important for the claim put forward in this chapter, and in the work of Stivers and Rossano (2010), concerning how participants mobilize response and concerning a compositional view of action in interaction and pressuring for response. The results in Table 3.11 also show that the failure of a pursuit in one modality leads a participant to either attempt using the other modality (which leaves open the option of a potential third pursuit, as will be shown later) or using a combination of the two modalities in order to increase pressure on the recipient for a response. Indeed, it appears that pursuing a response with both modalities seems to be the most effective way of pursuing a response. And gaze is almost always deployed. When gaze is deployed, it is usually not abandoned in the second pursuit, while if gaze has not been deployed the first time, the second pursuit nearly always (13/14 instances) involves the speaker gaze to the recipient. These results also highlight the cumulative power of these modalities, because if both are initially deployed, they will both be deployed the second time, while if only gaze has been initially deployed, most of the time both will be deployed in the second pursuit.

Further, these patterns suggest that if speakers pressure for a response with only one modality, they are less effective than if they deploy both. However, if they pressure with both modalities initially and this does not prove effective, they have no other additional resources to deploy beyond repetition of the action. Tables 3.8 and 3.10 suggest speakers typically pursue a response with one modality, with gaze being the most likely to be deployed of the two. This suggests that if a response can be obtained by exerting the least noticeable pressure, such a pursuit is preferable (an idea that will be further discussed later on in this chapter). An additional piece of evidence about the tendency to cumulate modalities to maximize effectiveness comes from the fact that of the 14 FPPs that are pursued twice, 2 are actually pursued three times[21] and of these one is finally responded to and one is not. In both cases in which a response is pursued three times, the final time responses are pursued with both modalities. Here are the patterns of pursuit for these two cases:

---

21 No response is pursued more than 3 times in my corpus.

1) 1st Pursuit Only Verbal → 2nd Pursuit Only Gaze → 3rd Pursuit Both → No response

2) 1st Pursuit Only Gaze   → 2nd Pursuit Both      → 3rd Pursuit Both → Response


## 3.4.7 Timing of Responses and Pursuits

Let's now consider the timing of delayed responses, the timing of pursuits and the timing of responses after pursuits. The average delay time across the 157 responses that occur without any pursuit (and yet are delayed by more than 0.2 seconds) is 0.68 seconds.[22] Approximately 90% of the responses occur within 1 second of real time, confirming the "standard maximum silence" of 1 second described by Jefferson (1989) for conversations in English (see also De Ruiter, Mitterer, & Enfield, 2006; Sacks et al., 1974; Stivers et al., 2009). If we then look at the time it takes before a speaker starts to pursue a response (considering the 48 cases of actual pursuits and not the ones incrementally modified before any silence occurred) we find that the average time is slightly longer at 0.76 seconds.[23] Although this difference is not statistically significant, it still confirms something very important: on average, a speaker waits to pursue a response until the delay has exceeded the average delay time for a delayed response. If by that time the recipient has not yet responded, then the speaker may begin pursuing a response.[24] In my corpus, the approximate time when this tends to occur is about 0.6-0.7 seconds from the completion of the FPP.

With this information in mind, if we then consider how long a participant waits to pursue a response if s/he pursues it repeatedly, we find that there is no significant difference between the waiting time before the first pursuit and the second pursuit.[25] This suggests that each newly deployed pursuit provides a new opportunity for the recipient to respond and the silence tolerance does not change if one pursues a response twice or only one time. In other words, a speaker does not lose patience nor does s/he become more patient when faced with the absence of response.

---

22 The median is 0.60 seconds.
23 The median is 0.70 seconds.
24 It is possible, and in some cases it seems likely (e.g., in the interaction 2PLUNCH1 and in 2PEXAM), that the time a speaker waits before pursuing a response is calibrated on the typical speaking speed of the recipient and on the time that that particular individual takes to produce a dispreferred response. However, this corpus is too limited to allow a systematic empirical investigation of this possibility.
25 Considering responses to FPPs that are pursued at least twice, in my sample the average wait time before the first pursuit is 0.67, while the average wait time before the second pursuit is 0.81; however, the difference is not statistically significant (p=.345).

This finding is confirmed if we consider how often and how quickly a speaker who has modified or pursued a FPP obtains a response. In the database, of the 63 FPPs that are modified, 47 are responded to without further pursuits (75%). This result is not very different from the 80% of FPPs responded to (with the remaining 20% being the ones not responded to at all and the ones that are pursued) that we found in the general dataset for FPPs that are not themselves pursued. Eleven responses occur in partial overlap with the pursuit, 16 responses occur promptly and the remaining 20 have an average delay of 0.62 seconds (the exact same range as the 0.68 seconds found for the delay of responses to initial FPPs). These numbers suggest that pursuing does not necessarily make it more likely that a speaker will secure a response nor that s/he will secure one more quickly than following the original FPP. However, if a FPP does not get responded to within ± 1 second, it will likely remain that way. Pursuing responses nevertheless gives the speaker another chance at obtain a response in order to complete the sequence the speaker initiated with the first FPP.

### 3.4.8 Examples of Gaze + Verbal Pursuits

Examples of the pursuing patterns discussed in the previous section will be presented in this section. Moreover, further evidence of speaker sensitivity to a 0.6-0.7 seconds silence for the conversations in my dataset will be provided. The first two examples will show pursuits performed verbally only, without speaker gaze. The following two examples will show pursuits done both with gaze and verbally. Having already showed multiple examples of pursuits performed only through gaze, this section is meant to illustrate the patterns listed in Table 3.8 and show how pursuits are produced in attempting to obtain an absent response.

The first example shows how both participants orient towards 0.6 seconds of silence as the maximum length of silence before a response should be provided or will be pursued. In example 3.17 two friends are looking at pictures together and B, the guest, asks A about the content of the pictures. Our focus is the silence at line 3 and what happens next.

**(3.17) 2PCOMP-carta 10:17**

```
01      (1.0)
```



```
02 B:   Cos'e' quella roba  la'  ((indicando))
        What is that  thing there
```
**What is that thing there  ((pointing to the picture))**



```
03      (0.6)
```



```
04 B:   [La carta,]
        The paper
```
**[The paper,]**



```
05 A:   [E' il coso   per la c]arta
        Is the thing for the paper
```
**[It is the thing for the p]aper**



```
06      (0.3)
```
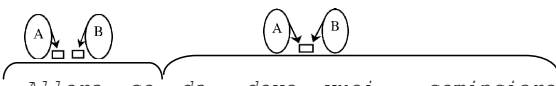
B points towards an object in the picture while producing the question at line 2 but A does not respond promptly. After a silence of 0.6 seconds, the recipient starts to respond, but simultaneously the speaker of the initial FPP starts pursuing a response by offering a candidate answer (which turns out to be the correct response). The point here is that both the speaker and the recipient orient to 0.6 seconds as a maximum silence. Notice also that both participants are looking down at the picture and not at each other and therefore the pursuit is done only verbally. Indeed, to adequately respond to this FPP requires both participants to focus on the object under examination.

Example 3.18 illustrates another example of a verbal pursuit in which the speaker does not look at the recipient. In this fragment, B is teaching A how to make flowers with

beads. They are both looking at the materials they are going to use for this task. Our focus is the gaze behavior of the participants during the pursuit at line 3-4.
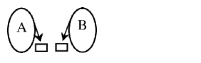
**(3.18) 2PERLINE-cominciare 04:43**
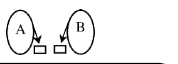
```
        A   B              A   B
         ⌐┐                 ⌐
01 B:  Allora co- da   dove vuoi   cominciare,
       Alright st- from where want.2s start
       Alright le- How do you want to start,


        A   B
         ⌐
02     (0.6)


                      A   B
                       ⌐
03 B:   .hh Vuoi   incominciare dall'   infilare perle,
            want.2s start        from the thread   pearls
        .hh Do you want to start from threading beads,


                     A   B
                      ⌐
04     ((con pollice sinistro alzato))(.) per fare:
                                        to  make:
       ((holding her left thumb up))(.) to make:


              A   B
               ⌐
05 A:  No. Allora (0.2) Du-
       No  So         so-
       No. So (0.2) le-


             A   B
              ⌐
06 B:  O  vo- o  vo- o-
       Or wa- or wa- or-
       Or wa- or wa- or-
```

```
07 A:  Dove   eravamo arriva- dunque
       Where  be.1p   arrive- so
       Where did we sto- let's see
```

A does not respond the initial FPP by B and, after a silence of 0.6 seconds and a short inbreath, B pursues a response verbally (line 3), continuing to look at the objects on the table. She does so by listing the possible alternatives in terms of where to start (indicated in her gesture at line 4). A rejects the first alternative provided by B and at line 7 she asks a question of her own for additional contextual information to help her answer B's question. Two points are particularly relevant here: as in 3.17, the pursuit starts around 0.6 seconds after the end of the first FPP, and it is performed without any speaker gaze towards the recipient. The speaker is looking at the objects that are relevant for the upcoming task (making flowers with beads),  and doing so facilitates B's suggestion of where they could begin. Her lack of gaze towards the recipient during the pursuit is therefore accountable by considering the contingencies of the task at hand.

Now let's look at two examples in which a verbal pursuit is also accomplished with speaker gaze towards the recipient. Example 3.19 shows a case in which the speaker of the initial FPP pursues a response while gazing towards the recipient and, as a result, the participants engage in mutual gaze. In this example, two friends are preparing for an exam together and B is asking A questions to check his preparation. Our focus is the gaze behavior during the silence at line 2 and the pursuit at line 3.

**(3.19) 2PEXAM-classi  42:19**
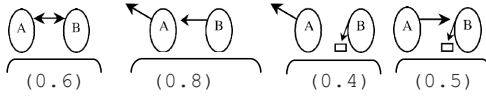


```
01 B:  Le   malattie sono divise  in cinque  classi
       The  diseases are  divided in five     classes
       Diseases are divided in five classes
```



```
02     (0.3)         (1.3)              (0.4)
                     (2.0)
```

188
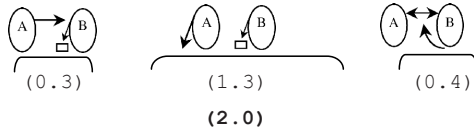
03 B:  *Le   sai    le classi di divisione*
       Them know.2s the classes of division
       **Do you know the classes in which they are divided**

04     (0.6)   (0.8)   (0.4)   (0.5)

       **(2.3)**

05 A:  *No ((scuotendo la testa))*
       No
       **No ((+ headshake))**

06     (1.5)

07 B: hhh (0.2) hhu hhu
      **hhh (0.2) hhu hhu**

Example 3.19 starts with B looking at some notes and stating that some animal diseases are divided into five classes. In the context of B asking A questions, line 1 implies "can you list the five classes?" There is no response and after a long silence the speaker moves his head up and looks towards A, with his eyes reaching A exactly at the time he starts producing the verbal pursuit at line 3. Notice that B withdraws his gaze from the recipient after a long delay in which A enacts thinking about a response. Then, following 2.3 seconds, the question receives a disprefered response.

Example 3.20 shows a case in which the speaker of the FPP is already gazing at the recipient during its production and holds his gaze throughout the ensuing gap and well into the verbal pursuit. By the time the verbal pursuit occurs, the speaker is already gazing at the recipient so both modalities are in play here. In this excerpt two friends are looking at pictures together, and at line 9, B asks how much A paid to get into the place shown in one of the pictures. Our focus is the gaze behavior during the silence at line 10 and during the pursuit at line 11.

**(3.20) 2PCOMP-pagato 9:33**

```
08        (1.0)

09 B:   Soccia  quanto   hai  pagato per entrar qua.
        Wow     how much have paid  for enter  here
        Wow how much did you pay to enter here.

10        (0.2)

11 B:   Die[ci
        Ten
   B:   Te[n

12 A:       [Poco.  °Un euro. Due euro neanche
            Little. One euro two euro neither
            [Little.° One euro. Not even two euros
```
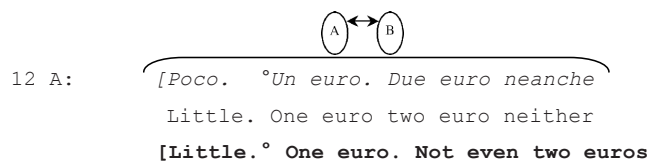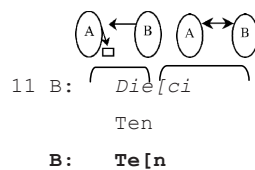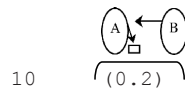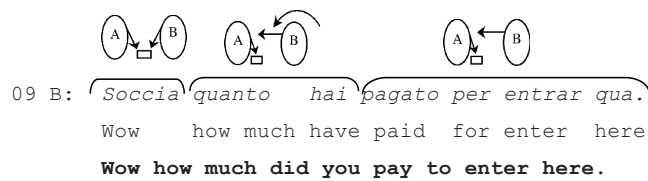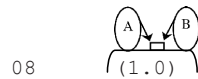
In this example the response to the FPP occurs in partial overlap with the pursuit at line 11. The pursuit is incomplete but it clearly proceeds in the same way as the verbal pursuit in example 3.17: by providing a candidate answer to his own question. Here the speaker holds his gaze towards the recipient throughout the gap at line 10, and by the time he initiates the verbal pursuit he is already looking at the recipient. This is a potential variant of what was shown in example 3.19 in terms of how a speaker may pursue a response by both producing a new utterance *in addition to* looking at the recipient. In example 3.19, the speaker looks up towards the recipient at the same time as he starts the pursuit as he was not looking at the recipient before. By contrast, in this example, the speaker gaze is held through the gap and also while producing the verbal pursuit.

In this section I have only shown examples of verbal pursuits, which are mainly accountable through the contingencies of the interaction, and examples of combinations of verbal pursuits produced with speaker gaze towards the recipient. Together with the examples of gaze pursuits shown in the previous sections, these four patterns constitute the possible ways in which a speaker can pursue a response. In the following section, I show examples of the patterns observable when a response is pursued more than once.

## 3.4.9 Multiple Pursuits

As shown in Table 3.11, in my dataset only four different patterns are observed when a response needs to be pursued more than once:

| | | |
|---|---|---|
| 1) 1$^{st}$ Pursuit Verbal Only | → | 2$^{nd}$ Pursuit Gaze Only |
| 2) 1$^{st}$ Pursuit Gaze Only | → | 2$^{nd}$ Pursuit Verbal Only |
| 3) 1$^{st}$ Pursuit Gaze Only | → | 2$^{nd}$ Pursuit Both modalities |
| 4) 1$^{st}$ Pursuit Both modalities | → | 2$^{nd}$ Pursuit Both modalities |

In what follows I provide examples that illustrate these patterns and discuss how they come to be realized by participants in interaction.

### 3.4.9.1 Two Pursuits: 1$^{st}$ Pursuit Verbal Only→ 2$^{nd}$ Pursuit Gaze Only

In example 3.21 we see that the first pursuit is done only verbally, while the second one is done only through gaze. Here there are two students preparing for an exam. A is visiting B, who is looking at his computer screen during the entire fragment produced below. Our focus is A's gaze behavior at line 8 and during the silence at line 9.

**(3.21) 2PEXAM-appunti 47:22**

```
01 A:  Eh mi devi    dare allora il::
       Eh me must.2s give then   the
       Eh you should give me then the::

02     (0.6)
```

```
03 A:   .hhh
        .hhh

04      (0.3)
```



```
05 A:  il  coso <E    hai     degli appunti?A parte il libro
       the thing and  have.2s some  notes  apart   the book
       the thing <And do you have notes? Apart from the book
```



```
06     di Trenti
       by Trenti
       by Trenti
```



```
07     (.)
```



```
08 A:  O hai      solo quelli li'
       Or have.2s only those  there
       Or do you have only those ones
```



```
09     (0.6)  (0.1)  (0.8)
             (1.5)
```
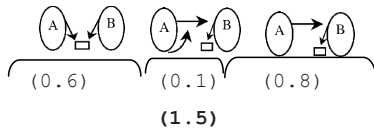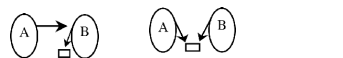


```
10 B:  Appunti non ne    ho
       Notes   not them have.1s
       I do not have notes

11     (1.2)
```

When A initiates a request at line 5, it is unclear what the request is for ("you should give me the thing"); however, he subsequently asks whether B has notes. This is a request for

192

information but also an indirect request to obtain the notes. At the beginning of the transition relevance place, but before B responds, A adds an increment clarifying that he is not asking about a specific book. Nonetheless, B neither turns towards A nor answers the question, and at line 8 A reverses the preference of the question by asking whether B only has the book A just mentioned. This is a clear pursuit of response and it is delivered without any gaze towards the recipient. A's question at line 8 suggests a reasonable inference regarding B's lack of response; however, further delay by B nevertheless hints at the possibility that B does not have any notes. Before B's response at line 10, however, A implements a second pursuit, this time with gaze only. Notice that A starts moving his eyes towards B 0.6 seconds after the beginning of the silence and that he holds his gaze towards B for another 0.8 seconds. The length of this delay already suggests that the upcoming response will be dispreferred, and indeed it is (i.e., B does not have any notes; line 10). Throughout this fragment, the recipient looks at a computer screen and does not turn towards the speaker even when he responds at line 10. This suggests that the timing of the response has not been affected by the gaze pursuit. At the same time, it is notable that after having pursued a response only verbally, a speaker first initiates a gaze pursuit rather than another verbal one even though the recipient is looking towards his computer screen. Also notable is that the gaze pursuit occurs around 0.6-0.7 seconds from the completion of the prior turn.

### 3.4.9.2 Two Pursuits: 1st Pursuit Only Gaze→ 2nd Pursuit Both Modalities

In the following example (see also § 3.4.10 for two further instances of this pattern), a response to a FPP is first pursued only through gaze and then it is pursued verbally while sustaining the gaze towards the recipient. In example 3.22 two friends are having dinner together and B, the host, is asking about a typical schedule for A's day. Our focus is the gaze behavior during lines 7-8.

**(3.22) 2GCOLL-lezione 18:06**

```
01 B:  Facciamo una giornata normale
       Make.1p  a   day       normal
       Let's talk about a normal day
```
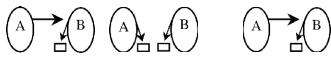
```
02 A:   Mm
        Mm
        Mm


03 B:   la   mattina ti  alzi,
        the morning you get up.2s
        in the morning you get up,


04      (0.2)


05 A:   Mm hm
        Mm hm
        Mm hm
```
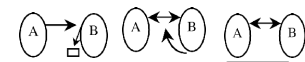


```
06 B:   poi   vai   a   lezione?
        then  go.2s to  lesson
        then you go to class?
```



```
07      (0.1)   (0.1)   (0.2)
                (0.4)
```



```
08 B:   questo periodo
        this   period
        these days
```



```
09 A:   ((gesticolando no con l'indice))  No
                                          No
        ((gesturing 'no' with index finger)) No


10      (1.1)
```

Here we see that B's request for confirmation at line 6 does not get an immediate response and B pursues it by looking up towards A. At this point they are in mutual gaze but A supplies no immediate response. B further pursues a response by specifying that she is not

asking whether she goes to class in general but whether she went to class in the morning during the last few days. At this point, A produces a negative response first through a gesture and then verbally.

In this section, we have seen that if the speaker is not looking at the recipient when s/he produces the initial FPP, s/he can pursue a response via gaze. If this fails, then a verbal pursuit may be produced, but gaze is sustained and not withdrawn, suggesting that the speaker is increasing the pressure for a response. In such cases, gaze appears to be used to hold the recipient accountable for a relevant next action.

### 3.4.9.3 Two Pursuits: 1<sup>st</sup> Pursuit Both Modalities→ 2<sup>nd</sup> Pursuit Both Modalities

The third pattern consists of two pursuits done with both modalities: producing a verbal pursuit while gazing at the recipient. In example 3.23 two friends are preparing for an exam and A asks B whether the professor expects students to know what B just asked A. Our focus is what happens at line 3 and at line 5.

**(3.23) 2PEXAM-sapere 42:49**



```
01 A:   Ma  lei vuol   sapere ste  cose   qui?
        But she want.3s to know these things here
        Does she want (us) to know these things here?
```



```
02      (1.6)
```



```
03 A:   Eh?
        Eh
        Eh?
```



```
04      (1.0)    (0.2)    (0.4)

              (1.6)
```

```
                    A    B                  A →  B              A ↔ B
                   ╱╲ ╱╲ □                 ╱─╲ →╱─╲            ╱─╲   ╱─╲
                  (A) (B)                 ( A )  ( B )        ( A ) ↔( B )
                                            →  □                □

05 A:    ⎧  Lei  son  ste   cose   qua [che  vuole  sapere?] ⎫
         She  are  these things here that want.3s to know
         **Are these  the   things [that she wants (us) to know?]**
```

```
                                  A →  B          A ↔ B         A ↔ B
                                 ╱─╲ →╱─╲        ╱─╲ ╱─╲       ╱─╲ ╱─╲
                                ( A )  ( B )    ( A )( B )    ( A )( B )
                                   →  □

06 B:                        ⎧ [Lei  vuole   sape]re quelle ⎫
                              She  want.3s to know those
                              **[She wants (us) to know those**
```

```
                      A ↔ B
                     ╱─╲ ╱─╲
                    ( A )( B )

07       ⎧ 'segnalazioni immediate'=si' questo e' importante ⎫
          signaling    immediate   yes  this  is important
          **'Immediately Inform'=yes this is important**
```
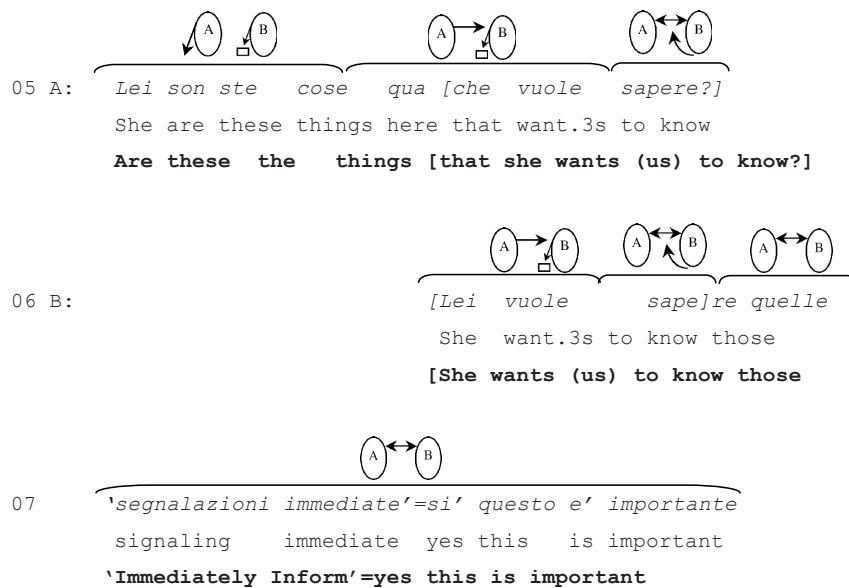
In this excerpt A's initial request for information at line 1 is not responded to for 1.6 seconds and A first pursues response minimally with "eh?"(line 3). A had already started looking at the recipient during the first FPP and sustains his gaze towards B. B, however, is looking down at some notes and very likely cannot perceive A's gaze. As a result, this first pursuit is not successful, and after waiting for another 1.6 seconds[26] and sustaining gaze towards the recipient, A produces a second pursuit (line 5), which is responded to in overlap (line 6). B's answer confirms that the professor does indeed want to know what B just asked A, but this confirmation is done first by repeating what A had just asked and then by saying "yes". As this example shows, if a speaker starts looking at the recipient during the initial FPP, usually s/he sustains the gaze throughout the following gaps and pursuits and does not abandon it as a modality for pressuring the recipient to respond.

---

26 The fact that A waits an identical amount of time before pursuing a response both times confirms the claim that speakers tend to provide recipients with the same opportunity to respond, independently of whether it is the first or the second time they are asking or pursuing the question.

### 3.4.9.4 Double Pursuit 1ˢᵗ Pursuit Only Gaze→ 2ⁿᵈ Pursuit Only Verbal

The last pattern discussed here is a deviant case. It is the only case in which, after an initial gaze pursuit, the next verbal pursuit is produced without speaker gaze. In example 3.24 a couple is having lunch and they are talking about a teaching assistant they both had. B has previously stated that this teaching assistant, Arnoldi, is quite bad (presumably as a teacher). After some silence and a topic change, A returns to B's negative assessment of Arnoldi and announces that Arnoldi studied at Berkeley (line 31-32). Our focus is the gaze behavior at line 33 and the pursuit at line 34.

**(3.24) 2PLUNCH-Berkeley 06:32**

```
31 A:  Sai    che  Arnoldi (1.0) tlk tlk e' stato
       Know.2s that Arnoldi       tlk tlk has been
       You know that Arnoldi (1.0) tlk tlk also studied
```

```
32 A:  a  Berkeley anche
       at Berkeley also
       at Berkeley
```

```
33        (0.3)   (0.2)   (0.5)   (0.3)   (0.4)
                        (1.7)
```

```
34 A:  Quindi   proprio scarso      non dev'essere
       Therefore really  insufficient not must be
       So he can't be really bad
```

```
35        (0.5)
```

```
                    ⎛A⎞ ⎛B⎞
                   ⎝__⎠ ⎝__⎠
                    ⌐⌐  ⌐⌐
36 B:   E' diventato scarso
        Has become   insufficient
        He has become bad
```

Here, A informs B that Arnoldi has studied at a very prestigious American University, and, in the context of assessing Arnoldi's qualities and intelligence, this becomes a way of countering B's previous negative assessment. At this point B could receive this announcement as news and then change her stance towards Arnoldi, but she does not respond to the announcement (note the silence in response at line 33). Following this silence the speaker pursues response with gaze and engages mutual gaze with the recipient. We know from previous examples that when participants engage in mutual gaze the expectation is that the recipient will respond promptly, but after 0.5 seconds of mutual gaze and an additional 0.3 of only speaker gaze, the recipient does not answer. The engagement of mutual gaze and then its abandonment by the recipient without responding can be understood as an indication that the recipient is not going to respond at all. After a bit more silence, during which both participants are oriented towards their dishes on the table, A further pursues a response by making explicit what the announcement at line 1 was about: providing further evidence that Arnoldi is not 'bad'. At this point, after some more silence, B responds with a quip about Arnoldi, insisting on disagreeing: probably he was good and he became bad recently.

In this example we see that the initial announcement is not just designed to be received as news but mainly to counter what has been negatively assessed by one party (the recipient). The recipient clearly withholds responding to this announcement by abandoning mutual gaze without responding, after having sustained it for 0.5 seconds. At this point, although perceivable as a further pursuit, line 5 becomes a way of making explicit A's evaluation of Arnoldi and his disagreement with B. Only at this point, and after further silence, B responds. This suggests that this pattern is occasioned by the context of disagreement and resistance by the recipient and by the fact that the verbal pursuit is not really pursuing the announcement but the stance that he was conveying through that announcement.

With the above examples in mind, in what follows I show how the findings of this chapter modify our understanding of pursuits and pursuing possibilities.

## 3.4.10 Patterns of Pursuits

In § 3.4.0 I outlined existing claims about how and when some utterances get pursued. In this chapter I showed that gaze can be used to pursue a response, just as the verbal channel can, and that, in fact, this visual channel is relied on even more often. I also showed how these two channels can be combined to be even more effective in soliciting a response. I also showed the robustness of gaze pursuits as a practice quantitatively. In what follows I show two examples of pursuits that correspond to two ways of pursuing as reported by Pomerantz (1984b), to show that they are often preceded by gaze pursuits. The following two examples show how the patterns described by Pomerantz can become, in the context of face-to-face interactions, a "second-best" way of pursuing once a gaze pursuit fails. The first example deals with the first way that Pomerantz indicated speakers pursue a response:

- If the problem is an unclear reference or term, then the speaker could offer a more understandable reference to replace the problematic one.

Example 3.25 represents a case in which a gaze pursuit fails to obtain a response by the recipient (see Table 3.11). In this example A is teaching B how to make flowers with beads and while she produces line 1 A holds a rose made with beads in her left hand and points towards it. After producing a turn, the speaker looks up at the recipient and keeps looking until the recipient looks back. They engage in mutual gaze but silence ensues. Then, A self-repairs her talk (in the transition space), at which point B responds.

**(3.25) 2PERLINE-petali 29:06**

```
01 A:  Tu puoi farne    anche per dire (.) se^i.
       You can make them also to say       six
       You can also make of them for example (.) six.
```



```
02     (0.3)   (0.2)   (0.2)   (0.5)  (0.2)
                    (1.4)
```
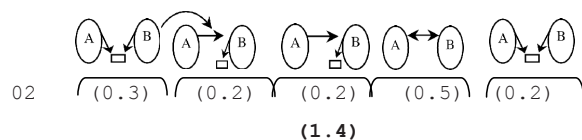
```
       (A)→,(B)  (A)→,(B)    (A)→,(B)
03 A:  Sei     p e t a l i [se vuoi
       Six     petals     if want.2s
       Six     p e t a l s [if you want


04 B:                      [Ah  beh si' si'
                            Oh well yes yes
                            [Oh well yes yes
```

At line 1 A is referring to the number of petals one can use to make a rose. The word "petal" has been used some turns earlier and here it is indexically referred to with the suffix "ne" in the word "farne" (make of them). After 0.3 seconds silence, A looks up at B who looks back at A. They sustain mutual gaze for 0.5 seconds but there is no uptake by B. They both look at the rose A holds in her hand and at this point A begins revising "sei" to "sei petali" (six petals) and then adding "if you want". But this second clause, "if you want", is overlapped with A's uptake introduced by "ah" that corresponds to the English "oh" and indicates a change of state (Heritage, 1984). B therefore produces an appropriate display of understanding only after A revises her initial utterance at line 1. This revision does not only occur after silence but after the participants have engaged, sustained and finally abandoned mutual gaze during the silence. B's production of the change of state token marks that some piece of information was missing, and, once it is offered, B understands what A meant to convey in her talk at line 1 and what sort of response was sought.

This example shows the practice described by Pomerantz (1984b). If there is no response speakers routinely orient towards the possibility of some trouble for the recipient in understanding the turn. However this example also adds further evidence for the practice described in this chapter: before participants pursue responses verbally, they can pursue a response with gaze. In particular, the speaker solicits a response by gazing during the transition relevance place. The recipient looks up too and engages in mutual gaze, but the pursuit here initially fails.

A second practice for pursuing described by Pomerantz is checking that the information required to respond is actually shared. Example 3.26 is a case in point. In this example A and B are looking at pictures together and A is describing some pictures of Stockholm. A attempts an indirect reference to something salient for B (a Swedish soccer

team that played against his favorite Italian team two days earlier). This reference initially fails and so does the gaze pursuit. At this point, A makes explicit what he was referring to.

**(3.26) 2PCOMP-Stoccolma 12:40**

```
49 A:   He hh .hh hh Questo invece  e' il lungo canale_ (0.5) tch
                    This    instead is the long channel
        He hh .hh hh This instead is the long canal (0.5) tch
```

```
50      che  arrivi    al- al parco di Djugordel.
        that arrive.2s at- at park  of Djugordel
        that arrives at- at the Djugordel's park
```
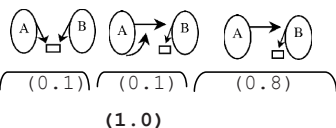
```
51      (0.1)  (0.1)   (0.8)
              (1.0)
```

```
52 A:   Che poi e' di quella squadra che  ha  giocato contro  la
        That then is of that  team    that has played  against the
        That is then that team that played against
```

```
53      Juventus °l'altra   [sera.
        Juventus   the other evening
        Juventus °the other [evening.
```

```
54 B:                  [Che gli abbiamo dato quattro pappini. [E'=
                        That it have.1p given four    slaps    Is
                       [To whom we gave four goals.        [It's=
```

```
55 A:                                                        [Hhh
                                                             [Hhh
```

```
56 B:  =quella li' °insom[ma eh ho    capito°
        That   there in sum  eh have.1s understood
     =that one °in the end eh I understood°

57 A:                    [hhu hhu hh
                         [hhu hhu hh
```

Djugordel is the name of a football team who played the night before against Juventus, an Italian football team which B proudly supports. What seems clear from the extract is that A is producing a reference to a football team by referring to a park. The problem here is not the name of the park itself, but rather the inferential process that B has to make in order to recognize the reference to the football team as something familiar to him. When A looks at B during the silence at line 51, and B not only does not respond but does not even look up towards him, A explicates this inferential process and therefore why he mentioned the fact that the long channel leads to that park. B's reaction at line 54 and 55 confirms that he understands A's talk as a further pursuit. B not only displays recognition of the reference, but he also stresses that he understood "eh ho capito" (eh I understood). Thus, this is another example in which a practice of a verbal pursuit described by Pomerantz is actually preceded by an attempt to solicit uptake using gaze.

The last two examples illustrate that, in face-to-face interaction, visible behavior is routinely deployed before (or sometimes after) verbal pursuits to pressure for a response. In what follows I show how the practices described throughout § 3.4 modify our understanding of what a speaker does when a response is not provided in a timely fashion.

## 3.5 Schemes of Pursuits

Section 3.4 examined what speakers do when a FPP does not get an immediate response. At the beginning of § 3.3, this chapter presented two ways in which a response can be secured if it is not produced in overlap or in a timely fashion:

3.
Time ····|A: FPP | Silence |··········▶
         |       |         |B: SPP |

4.
Time ····| A: FPPb | Silence | A: FPPp |··········▶
         |         |         |B: SPP |
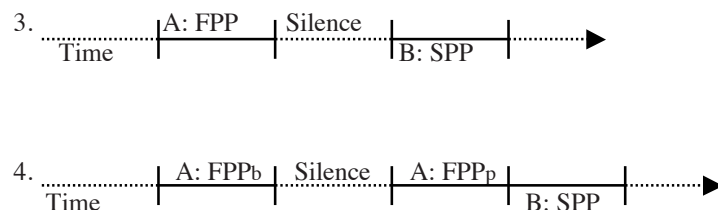
Scheme 3 illustrates the situation where a response occurs after some silence without any pursuit by the speaker of the FPP, while Scheme 4 illustrates the situation where a speaker produces a verbal pursuit after some silence, and, at that point, obtains a response. However, throughout § 3.4 I have shown multiple variants to these schemes that include both verbal and visual pursuits.

The first two variations consist of producing an "incremental modification" of the FPP as soon as the FPP reaches completion and before the recipient has delayed any response (see § 3.4.2). I showed that this incremental modification can be done verbally or simply through gaze. The schemes are as follows:

3b) ···················· | A: FPP | Incr. Modif. | B: SPP | ····▶
       Time

3c) ···················· | A: FPP | Silence | B: SPP | ············▶
       Time
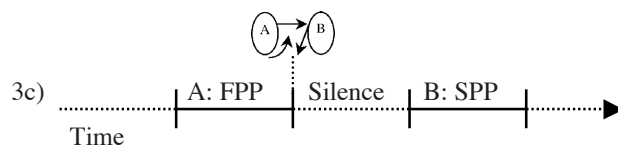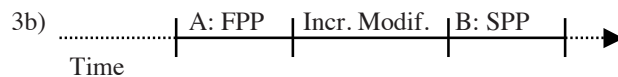
I then showed that speaker gaze can be used to pursue a response ,after some silence and that usually the recipient responds promptly to this pursuit (see § 3.4.4). The scheme looks as follows:

3d) ·········· | A: FPP | Silence | B: SPP | ············▶
       Time
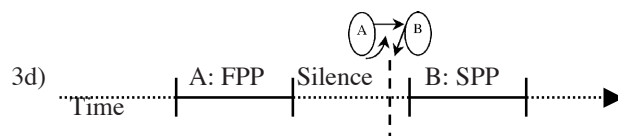
Next I showed that gaze pursuits sometimes does not obtain an immediate response by the recipient, but rather a delayed one. This is normally due to the dispreferred nature of the response (see § 3.4.5). The scheme looks as follows:

3e)

A → B

| A: FPP | Silence | Silence | B: SPP |
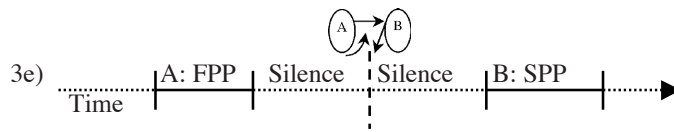
Time

I then showed that sometimes the gaze pursuit can be abandoned before any response occurs because the lack of response may suggest that no response will actually be produced or because the recipient displays 'thinking' about an answer, and therefore that the pursuit has been successful. A dispreferred answer usually occurs nonetheless after more delay and usually this is attributable to a competing activity. In these cases, the occurrence of a response is not *necessarily* caused by the gaze pursuit, yet the practice of pursuing via gaze is deployed and then subsequently abandoned (see § 3.4.5). The scheme looks as follows:

3f)

A → B        A → B

| A: FPP | Silence | Silence | Silence | B: SPP |

Time
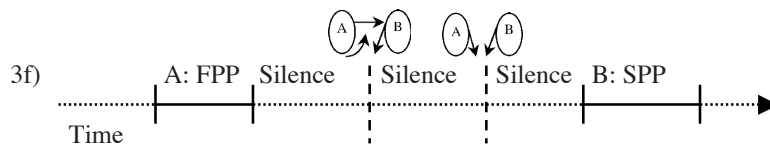
The schemes 3c-3f have not been described before and yet they are patterns that occur repeatedly in dyadic face-to-face interactions. More importantly, the patterns 3c-3f occur without the occurrence of any verbal pursuit and they seem to play an important role in the occurrence of a response. If we then consider situations in which a verbal pursuit occurs, we should add to scheme 4 the following three schemes:

4b)

A → B

| A: FPPb | Silence | A: FPPp | B: SPP |

Time

4c)

A → B

| A: FPPb | Silence | A: FPPp | Silence | B: SPP |

Time

4d) Time — A: FPP$_b$ | Silence | A: FPP$_p$ Silence | A: FPP$_{p_2}$ | B: SPP
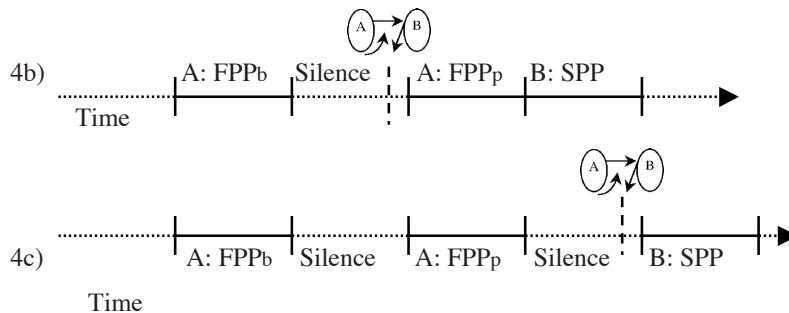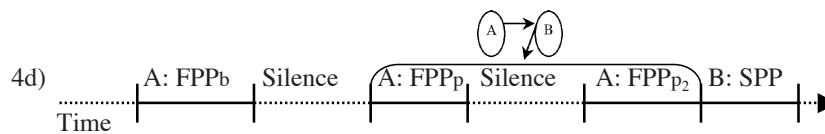
Scheme 4b represents all those cases of double pursuits in which a response is first pursued only by speaker gaze and then verbally (usually while sustaining the gaze towards the recipient). We have seen examples of this pattern in § 3.4.9.2, in § 3.4.9.4 and in §3.4.10. Scheme 4c represents those cases of double pursuits in which a response is initially pursued only verbally and the second pursuit is done through speaker gaze (see § 3.4.9.1). Finally scheme 4d represents those cases in which both pursuits are done verbally and with gaze. In only two cases in my dataset, a response is pursued three times, and in both cases the third pursuit is done with both modalities (verbally while looking at the recipient). No response in my database is pursued more than three times.

If we consider all of the above schemes one after the other we can see that our initial knowledge of pursuing patterns has been considerably enhanced. Moreover, we can see a gradation of pursuing practices and pursuing possibilities that go from an immediate verbal post-positioned modification, to a single gaze pursuit, to a combination of verbal and visual pursuits. While each occurrence has a discrete beginning (the looking at the recipient or the beginning of the verbal pursuit), the difference between schemes 3c and 3d is based on the time elapsed between the end of the FPP and the beginning of the speaker's looking up towards the recipient. Similarly, the difference between 3d and 3e is based on the time a recipient takes to respond to the gaze pursuit. From an objective point of view, we are differentiating patterns in terms of the timely occurrence of specific practices, which would mean simply adding time as a variable for our model. Yet, I have shown that the transition relevance place is segmentable into more specific points of particular interactional relevance. The occurrence of a modification of a FPP, or of a response before any silence or delay, is perceivable and has very different interactional consequences and implications than the occurrence of a verbal pursuit or of a response after some noticeable silence. In other words, once we map gaze practices onto interactionally relevant places, more fine-grained schemes can be differentiated and other types of orderly behavior can be observed. The most important point here is not that these different patterns can be observed, but rather that participants in interaction regularly differentiate them, as the multiple examples shown in § 3.4 illustrated.

Moreover, the kind of patterns that can be identified, in terms of how the different modalities can be deployed to pursue a response, appear to play an important role. The fact that no instances were found of some specific patterns (e.g., using the same modality in the first pursuit to pursue a second time, or moving from using both modalities in the first pursuit to using only one in the second) suggests that gaze and talk can have a combinatorial effect. Moreover, it suggests that gaze and talk can likely be ranked in terms of the degree of pressure they exert, where gaze may be the least pressuring, a verbal pursuit may be the next step and a verbal pursuit produced while gazing towards the addressee is an even stronger means to pursue a response.

Once we reach this level of granularity and once we see how a speaker can deploy her/his gaze to pressure a recipient for a response after a FPP, it is also possible to extend these observations to similar environments, where the relevance for a response is not as strong. In what follows I present four other environments in which we can see how speaker gaze can be used to obtain a response by a recipient.

### 3.6 Gaze Pursuits in Different Sequential Environments

Until now we have observed how speakers can use gaze to pressure for a response either during the production of a FPP (§ 3.3) or in the transition relevance place following it (§ 3.4). In all of the previous examples the participants were engaged in turn-by-turn talk (Sacks et al., 1974), alternating those turns mainly through question-answer sequences. However, there are other sequential environments in which gaze can be deployed to pursue a response by a recipient. In what follows I show what happens in two other sequential environments:

1. within a storytelling;
2. after a laughable item is produced (e.g., joke).

I then look at two additional environments in which gaze can be deployed to pressure for a response, this time during a TCU, rather than at the transition relevance place. These two environments are:

a)  after a preliminary component of a TCU;

b)  after a recognizable reference.

## 3.6.1 During a Telling

Among the ways in which participants make a response relevant is through the use of tag questions. During the course of a telling, a teller can provide references to individuals, places or events that are supposed to be part of shared knowledge or constitute the common ground (e.g., Clark & Brennan, 1991) for the successful reception of the telling. At specific points, a teller can verify whether the recipient recognizes a specific reference by using tag questions (see 'try marking' in Sacks & Schegloff, 1979). We will see in § 3.6.3.2 that, in addition to tag questions, a teller can also use gaze to elicit a response, not at just any point but while producing recognizable references.

The following example illustrates a telling during which the teller uses a tag question to verify that the recipient knows the movie he is talking about. The recipient does not respond and response is then pursued through gaze. This interaction takes place in a car and the participants are friends traveling towards a mountain resort. A is the driver and B is sitting in the middle of the back seats, with his head forward between the two front seats. A and B are talking about the movie, Pulp Fiction. Our focus is the gaze behavior during the silence at line 12.


**(3.27) 3PDRIVING-massaggio 43:56**

```
01 A:  E   poi  mi e' piaciuto molto come attore (.) anche:: (1.5)
       And then me is liked    a lot as   actor      also
       And then I liked a lot as an actor (.) also:: (1.5)

02     come cazzo si  chiama (1.0) Bruce Willis. A me::
       how  dick  cl. call.3s     Bruce Willis  to me
       what the hell is his name (1.0) Bruce Willis. To me::

03 B:  Ah che  fa    il [pugile
       Oh that do.3s the boxer
       Oh who plays the [boxer
```

```
04 A:                        [Per me ha  saputo fare il  suo ruolo
                              For me has known  make the his role
                             [In my opinion he managed to play his role


05      mol[to bene
        very   well
        ve[ry well


06 B:   [Fa      il pugile [lui
         Make.1s the boxer  he
        [Plays the boxer  [he


07 A:                         [((Nods))


08 A:  Comunque: e' incredibile_ cioe'  se tu  guardi  ogni
       Anyway    is incredible   I mean if you look.2s every
       Anyway: it is incredible_ I mean if you look at every


09      singola st- scena ogni singola storia vedi
        single     scene every single story  see.2s
        single st- scene every single story you see


10      dei paradossi incredibili tipo .hh cominciano a parlare e
        of  paradoxes incredible  like    begin.3p   to speak  and
        incredible paradoxes like .hh they start talking and
```



```
11      discutono un'ora (.) del   (massaggio dei   piedi no?
        discuss.3p one hour  of the massage    of the feet no
        they talk for one hour (.) about foot massage no?
```



```
12                                                   (1.0)
```
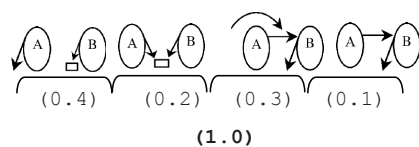


```
13 B:  Io non me lo ricordo    bene
       I  not me it remember.1s well
       I do not remember it well
```

```
14 A:  E   vabbe'. Poi dopo  c'e'   Bruce Willis che e' sto uomo duro,
       And alright then after cl is Bruce Willis that is this man tough
       And alright. Then there is Bruce Willis who is this tough man,

15     cattivo e   tutto che sta   con  la  donna piu' buona
       nasty   and all   that stays with the woman more good
       nasty and so on that is together with the sweetest woman

16     del    mondo
       of the world
       on    Earth
```

Both participants in this fragment are looking ahead at the beginning of this example. B has indicated, at lines 3 and 6, that he has seen the movie (i.e., he can remember which role Bruce Willis played). During the silence at line 12, though, B does not react to the tag question ("they talk for one hour about foot massage no?"). A has indeed started describing some of the things that he considers "incredible paradoxes" (see line 10) and the first item he provides is this one-hour conversation. This is supposed to be a recognizable reference for B, as someone who has seen the movie, but B does not respond to it. After a silence of 0.6 seconds the teller pursues response through gaze by turning his head towards the recipient and gazing at him. The recipient has clearly perceived the head turning of the driver and his gaze given that B's face is few centimeters behind A, positioned in between the head rest of the two front seats. At line 13, B accounts for why he has not responded yet: he does not remember the movie very well. This has consequences for the design of A's multi-unit turn, given that he resumes speaking at line 14 and modifies his way of describing the characters (see lines 14-15 in which A describes the character of Bruce Willis as if B has not seen the movie). This example shows an extreme case in which the teller would turn his head away from the road and his driving task to look at the main recipient sitting behind him, in order to pursue uptake rather than pursuing response verbally. Such an exaggerated move suggests that pursuing a response through gaze is not the same as pursuing it verbally. Moreover, this example is interesting because it shows a situation in which, after the response, what follows is not another question or assessment by the speaker of the initial FPP but rather the resumption of the telling. By inserting a tag question (at which point the driver is not looking towards the recipient) the speaker makes relevant the production of a response and when one is not provided promptly, response is pursued by gaze.

## 3.6.2 After a Laughable Item

The second sequential environment I investigate is that which occurs following the production of a laughable item such as a joke or a quip. The unsuccessful delivery of many jokes in interaction and the lack of general sanctioning suggests that laughing is the preferred reaction to a joke (Sacks, 1974) and so when it does not occur in response, it is noticeably absent. The following example provides some evidence that laughter can be pursued with gaze when it is missing.

In example 3.28 the speaker of the laughable turn is the driver of the car (A), who pursues a response by turning almost completely towards the main recipient (B), who is sitting in the back of the car. In the interaction preceding this extract, A informed the other two passengers that there are engineers paid to develop a pasta shape, which improves the amount of sauce that attaches to it. A clearly indicates his negative stance towards this fact first by saying that there are people who work on ridiculous issues and then by asking the rhetorical question at line 1: "which kind of problem is it?" Our focus is the gaze behavior during the silence at line 7.

**(3.28) 3PDRIVING-gin 13:10**

```
01 A:   Cioe'  che  problema e'¿
        I mean what problem  is
        I mean which kind of problem is it¿

02      E   quelli che studiano (0.4) che   reazione
        And those that study.3p       which reaction
        And those who study (0.4) which reaction

03      hanno   i   pesci=l'ho     letto sulla Settimana Enigmistica
        have.3p the fish  it have.1s read  on the Settimana Enigmistica
        fish have=I read it on the Crosswords Magazine

04      .hh [che reazione hanno(0.2) i   pesci se sottoposti a=
            which reaction have.3p  the fish  if exposed    to
        .hh [which reaction have (0.2)  fish if=
```

```
05 B:      [he he ha ha
           [he he ha h


06 A:    = ubriacatura da   gin:
           drunk          from gin
         = made drunk with gin:
```



```
07       (0.7)   (0.2)   (0.1)
                 (1.0)
```



```
08  B: [he he he he   ((See Figure 3.2))
       [he he he he
```
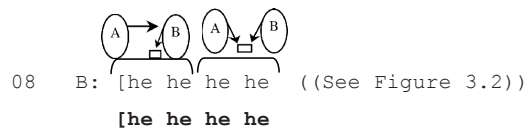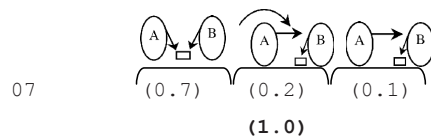


Figure 3.2. A's head turned towards B, who starts laughing (line 8)

The turn at line 1 is uttered while the speaker is looking ahead at the street and while making a "mano a borsa" gesture, often associated with rhetorical questions (Kendon, 1995). The way this turn is designed displays that it is not a real question but rather a way of displaying his negative stance. At line 2 A starts another TCU that projects the utterance of another type of job towards which he adopts the same stance. Before getting to the punch line of this funny announcement, he interrupts and begins a parenthetical sequence (Mazeland, 2007) with a turn ('I have read…') that makes explicit the source of this information. This is responded with a laughter by B at line 5 in overlap with A's resumption of his announcement ('which reaction…') that is completed at the end of line 6. B's laughter is not to the projected main gist of the telling (Norrick, 2010) concerning weird jobs, but rather towards the source

of information that he cites (the crossword magazine), clearly not the most commong type of news source.The turn that ends at line 6 does not get any immediate uptake. For 0.7 seconds nothing happens, until the driver, who is driving on the highway, turns all the way around towards B. Once A's gaze reaches B, B starts laughing. The point here is that it is not just the beginning of the turning of the head that pursues a response (or at least that this behavour displays that the speaker expects a response); rather, it is actually the engagement of mutual gaze by the participants and then the holding of the mutual gaze that achieves uptake/response. This shows that the same practice described in the previous examples can work during a silence following a funny remark.

### 3.6.3 Pursuing Uptake Mid-TCU

In what follows we will see two environments in which gaze can be used during the course of a TCU to obtain some uptake to a specific action:

a) during a silence following a preliminary component of a TCU;
b) during a silence following a recognitional reference.

#### 3.6.3.1 After a Preliminary Component

The practice of looking up towards a recipient during silence can also occur after the preliminary component of a TCU.[27] When a participant produces a preliminary component of a compound TCU another speaker has "semi-permeable" access to the turn (Lerner, 1996b). As Lerner (1996b: 241-242) puts it:

> *While distinctly different in social and interactional terms from TCU possible completion as a place of transition relevance, preliminary component completion also provides a place for another participant to begin speaking. [...] Preliminary component completion provides an opportunity for another to begin speaking even when no form of elicitation is evident and the first speaker continues on without hesitation.*

---

27 With compound TCU I refer to the type of TCU identified by Lerner (1991, 1996a) that includes a preliminary component that more or less projects the shape of the TCU as a whole and a possible form for the final component of the TCU. The most typical example would be an 'If X, then Y' construction.

In this section we will see that gaze can also work as a resource to elicit a response by the other participant in this environment. In particular, if silence follows the preliminary component, the speaker can look up towards the recipient and by doing this s/he solicits a response.

In the following example (continued from 3.24) A and B are talking about a teaching assistant and B has assessed him in negative terms. A has defended him citing his CV, and line 36 is B's teasing response to A's effort. The focus is the gaze behavior at line 42.

**(3.29) 2PLUNCH1-scarso 6:40**

```
35     (0.5)

36 B:  E'      diventato scarso
       Have.3s become   insufficient
       He became bad

37     (1.5)

38 A:  No perche' noi adesso lo vediamo li'   che  fa (0.4)che
       No because  we  now   him see.1p there that do.3s   that
       No because now we see him there doing (0.4) that

39     l'han      preso come:: come tutore sembra un po' portaborse
       him have.3p taken as     as   tutor  look   a  bit servant
       they got him as:: as a tutor he looks a bit like a servant but

40     pero' in realta' (1.0) gia'   se l'hanno preso per fare
       but   in reality        already if him had taken to  do
       actually (1.0) already if they got him to do

41     quella roba  li'_
       that   stuff there
       that thing_
```



```
42     ⌒          ⌒          ⌒
       (0.9)      (0.4)      (0.5)
              (1.8)
```
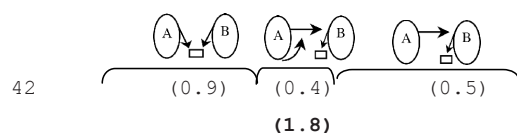
```
43        ((After 1.5 seconds of silence B makes disaffiliative facial
                                    expression see Figure 3.3))
```



```
Figure 3.3. B's facial expression at line 43.
```



```
44 A:  per curare il  corso °c'e'    un motivo°
       to  cure   the course there is a  reason
       to take care of the course °there is a reason°
```

```
45     (0.5)
```

A's turn at line 38 is responsive to B's negative stance towards the teaching assistant and to the turn at line 36 in which she teases A. In response, A provides another piece of evidence to support his positive stance towards the teaching assistant. But this evidence is provided through a compound TCU (lines 40-44). It should be noted that without taking into account the visible behavior of the participants, the silence at line 42 would be glossed as a simple pause in A's speech, which is then resumed and completed at line 44. However, by neglecting the participants' gaze behaviors we would miss the interactional events that occur during that silence. In the context of disagreement between A and B, the gist of the second component of the TCU is easily projectable after the preliminary component "but if they got him to do that thing". At this point, following Lerner (1996b), we could say that the other participant could start speaking, but she does not. On the other hand, in most of the examples shown in this chapter, by withholding speech the speaker provides an opportunity for the other participant to interject and possibly complete her/his turn. The silence is therefore created by the additional absence of talk by the recipient. There cannot be any collaborative completion given the disagreement, so what is the speaker doing by remaining silent at this

point? His use of gaze during the silence displays that he is actually trying to secure a response from B, and indeed he gets one. She does not talk but while closing her eyelids she raises her eyebrows making a disaffiliative facial expression (see Figure 3.3). This visible behavior displays that she does not accept A's account. At this point, A produces a self-repair replacing the generic "per fare quella roba li'" (to do that thing) with the specific "per curare il corso" (to take care of the course) and completes the compound TCU in a soft voice. Note that B's facial expression only makes sense by relying on the fact that the speaker (B) is looking at her.

What is striking is not only the use of gaze to solicit a response, but rather the way in which the speaker displays a clear orientation towards the possible relevance of recipient response (after a preliminary component of a compound TCU). The speaker orientation towards the possible relevance of a response at that point is further displayed by the fact that he produces the second component with a soft voice looking at his dish and that B does not respond. A had designed this turn to end at line 41 or at least to get responded to at that point and behaved accordingly, only to then complete the compound TCU once the preliminary component has failed to elicit a response aligning with A's stance towards the teaching assistant. Looking at the way gaze is deployed provides us with additional information about the speaker's orientation towards her/his own talk. In particular, it shows us that the turn was designed as a "trail off". Moreover, we can see that the speaker's turn becomes a compound TCU when the trail off is not responded to, rather than being designed from the beginning as a compound TCU and interrupted because of contingent actions.

### 3.6.3.2 After a Recognizable Reference
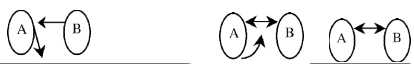
Another environment in which looking towards the recipient during a silence effectively pursues uptake is after a recognizable reference. The term "recognitional" is usually employed in CA to indicate a reference to a person or a place that invites the recipient to find who or what among the people or places s/he knows the speaker was referring to. That is to say, it is a reference designed to be recognizable by the recipient (Sacks & Schegloff, 1979; Schegloff, 1996b). The term "recognizable reference" is a more general category and includes things like technical terms and names of objects and animals in addition to names of people or places. Participants, indeed, orient towards the relevance of getting such things recognized and understood by recipients before continuing with their talk.

In previous work on reference, Auer (1984) has described a specific sequential format called "indexicality marker/continuer", which refers to practices deployed in producing a reference to people or objects that tends to elicit a continuer as a response. He explicitly states that this is *not* a classic adjacency pair because, in contrast to an offer and its acceptance, the second part of the format is optional. He suggests this is the case because "the recipient may ignore the referring party's orientation to the potential problem status of the referential item, that is, he or she may choose not to acknowledge it" (Auer, 1984: 645). When Auer describes the ways in which the reference can become the first item of this specific sequential format, he refers to the act of "try marking" (Sacks & Schegloff, 1979) and to the use of language specific demonstratives and specific techniques for displaying hesitancy like short pauses, repeated onsets, lengthening of sonorants etc. He analyzed audio-recordings of face-to-face interactions and therefore could not observe speaker gaze as an additional (crucial) element that contributes to the occurrence of a recipient response on the mentioning of a recognizable reference. In particular, speaker gaze towards the recipient is often already deployed on the production of the reference, so that by the time the reference is produced, (and in particular, if some silence follows) the recipient knows that some uptake is expected. In Auer's work, the focus is on practices that point to achieving reference as one of the main locally relevant activities for the participants. In the examples that follow, we should also acknowledge the relevance of recognizing the reference in order to understand what is going to be said next by the teller.

The response soliciting practice here presented tends to occur in the midst of multi-unit turns, and what follows is usually a re-engagement into the turn by the speaker. The following two examples show how this practice works. Example 3.30 illustrates the occurrence of this practice in this environment. The main difference with the previous examples is the occurrence of a gaze shift by the speaker during the production of the recognizable reference, and not just during the following silence. Prior to the beginning of this fragment, A began telling B about the planned visit of two guys she met on vacation. At line 1 and the beginning of line 2, the participants close off a repair on a reference to where the two guys are from.

**(3.30) 2GSOFA-biondo 14:12**


01 B:  *Vabbe' hhu hh[u*
       Alright
       **Alright hhu hh[u**


02 A:       *[Vabbe'fatto sta che c'e': Rindi (0.2)*
             Alright fact stays that there is Rindi
            **[Alright the fact is that there is: Rindi (0.2)**


03     *quello di  [Spe-  hhh]a si' che e' malato quindi =*
       that from  Spe-    oh   yes that is sick so
       **the one from [ Spe-  hhh]a yes who is sick so =**


04 B:  *((annuendo))  [Il biondo]*
                       The blonde
       **((Nodding))  [The blonde]**


05 A:  *= non son venuti .hh poi  oh Antonio non l'ho       sentito*
         not  be.3p come    then oh Antonio not him have.1s heard
       **= they didn't come .hh then oh Antonio I didn't hear from him**


06 A:  *quindi_*
       so
       **so_**
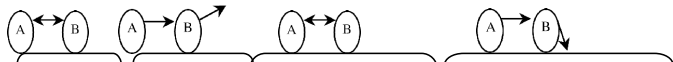

07     (1.0)
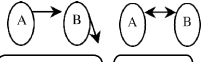

At line 2 A refers to one of the two guys by using his nickname (Rindi) and then remains silent for 0.2 seconds. While referring to Rindi she also looks at B, who is already looking at her. At this point, A makes the reference more explicit and she tries again by saying "the one from Spe-" but her talk is interrupted by B who demonstrates her recognition of the person A is talking about ("the blonde one" at line 4) (Heritage, 2007). A immediately confirms the identification of Rindi as "the blonde" ("yes") during her turn at line 3 and then proceeds with the telling. At this point, A is providing B with an opportunity to display recognition not

by using a vocal trymarker or by displaying hesitation, but by looking up and remaining silent immediately after the reference instead of continuing her telling. Given the absence of uptake, A proceeds by further specifying the reference while B produces a display of understanding in overlap with it.
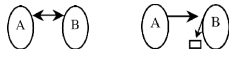
Example 3.31 shows a similar pattern, only this time the looking up by the speaker occurs on the last syllable of the recognizable reference (the Chemistry Department) and after some silence the recipient responds with an acknowledgment token. In this example, the participants are having dinner together and they have been commenting on the fact that one of their professors is a 'player' and is considered very good-looking by many students. A has just said that only men behave like that, before B counters that women would also do so. Our focus is what happens between lines 2 and 4.

**(3.31) 2GCOLL-chimica 24:17**

```
01 B:  Ma    no dai    anche una donna se tu  adesso vai
       Well  no come on also  a   woman if you now    go.2s
       Well no come on also a woman ((would do it)) if you now go
```

```
02     a Chimica
       to Chemistry
       to the Chemistry Department
```

```
03     (0.3)
```

```
04 A:  mm      hm ((annuendo))
       mm      hm
       mm      hm ((nodding))
```

```
05     (.)
```

218

```
              A ←→ B

06 B:   vieni   circondata_
        come.2s surrounded
        you get surrounded_



07      (.)

08 A:   mh mh[hh] hmm hmm
        mh mh[hh] hmm hmm

09 B:      [da]
            by
           [by]

10 B:   da milioni  di chimici che: che  vogliono solo te
        by millions of chemists that that want.3p   only you
        by millions of chemists tha:t that desire only you
```

In order to counter A's claim that only men exploit their attractiveness with students, B tries to present a scenario in which A could find herself acting in the same way. B starts by defining the environment in which A could find many good-looking male students and she does so by saying "if you now go to the Chemistry Department…" The reference to the Chemistry Department is of crucial importance because A has to recognize the reference to an environment in which most of the students are young men and they are good-looking (they had mentioned previously that some of their attractive friends study Chemistry). Before stopping talking at line 3, B looks up towards A and sustains mutual gaze until A confirms recognition of the reference and B can then continue with constructing a plausible scenario for a comparison of women's and men's behavior. In this example we again see that a speaker not only momentarily interrupts her speech to allow the recipient to confirm recognition of a specific reference (here the Chemistry Department), but also looks up towards the recipient and sustains gaze through the silence to show that a response is relevant at that point.

  In general, it is of particular importance to recognize that the practice of looking towards a recipient during a telling can indeed secure a response; however, it does so either because it occurs on a new reference to a recognizable item, sometimes verbally produced to

further solicit a response (see, e.g., Sacks & Schegloff 1979; Auer 1984), or because the gaze occurs on or after a tag question that already makes relevant a response. This is a crucial specification to the much broader and less specific claim by Bavelas, Coates and Johnson (2002) about the symbolic function of gaze during a telling: if the speaker looks at the recipient, s/he wants that recipient to respond with an acknowledgment token or a continuer. § 3.6.1 and § 3.6.3.2 provide a more specific claim about the particular environments in which gazing towards an addressee may solicit a response, a specification based on the interactional relevance of specific objects in conversation such as new references and tag questions.

## 3.7 Why Gaze Now?

Gaze pursuits contribute a partial solution to a crucial interactional problem: how can a participant get a response to what she has just said without actually asking again? Phrasing here is important. Let's start by focusing on "a response". One of the questions participants may unconsciously ask themselves in interaction is "why that now?" (Schegloff & Sacks, 1973). During a silence, once the speaker looks up towards the recipient it is very likely the recipient will ask herself/himself "why gaze now?" and also "what from me now?" It is clear, indeed, that participants in interaction are not dealing with gaze just as an epiphenomenon that happens to co-occur with talk. Nor do they treat it simply as a way of displaying attention, even though its relationship with attention might be the root of the successful use of the gaze practices here described. When deployed in certain sequential environments, recipients instead perceive speakers' gaze behavior as a way of doing specific actions in interaction. The way in which gaze pursuit is understood, as can be seen in how it is responded to, depends on the sequential environment in which it occurs. Notwithstanding the general characterization of this gaze as "response needed", gaze is not really symbolic and is instead mainly doing an indexical job. That is to say, in an environment in which some response could be relevant and/or appropriate, by looking towards the recipient the speaker appears to be asking the recipient to go back to what has just been said or done and to deal with it in an appropriate way. If we look at the kind of responses that participants get by deploying gaze in the sequential positions described in this chapter, we notice that they are not random ways of displaying uptake, but are in fact appropriate to the types of possible responses made relevant by the occurrence of a specific FPP (like a request for confirmation

220

that gets a yes/no answer or jokes that get delayed laughter). The evidence that the speaker is looking for some uptake when the gaze is deployed is provided by the cases in which gaze pursuit fails and then a verbal pursuit is performed. This makes plausible the indexical function of gaze in this context and at the same time displays how it relates to the terms used in framing the general question at the beginning of this section ("to what she has just done").

The question of which sequential environment gets a pursuit is valid also for verbal pursuits. Noticing how something can be pursued does not always tell us why something gets pursued in the first place. And even if it is true that speakers mainly pursue responses to FPPs, this chapter has also shown that other things can be pursued. The main gaze practice described here occurs during silence following a FPP, and also a joke, a recognizable reference and a preliminary component of a TCU. As Davidson (1984) and Pomerantz (1984) have shown, one possible reading of silence after a FPP is the projection of upcoming disagreement, or, at least trouble. I have also tried to show that another potential issue regarding silence after a FPP is actually recognizing that something could or should be responded to when it does not constitute a turn boundary. Gaze is certainly one of the resources participants rely on. In sum, the environment in which the practice of pursuing a response gets deployed is one in which something has failed and the progressivity (Schegloff, 1979, 2007b; Stivers & Robinson, 2006) has been halted.

The affordances of gaze (Gibson, 1979) are particularly important and lead us to the issue raised by the expressions "without being heard" and "asking again" that I used at the beginning of this section. The use of gaze allows the recipient to produce visible responses. Indeed, many of the actual responses produced by participants following its deployment are only visible or have a visible component. These responses range from nods, to gestures, to smiles, to facial expressions that display their stance towards the previous action.[28] Perhaps this is why the very occurrence of gaze in pursuing responses has simply gone unnoticed because so often it, like other visible responses, is neglected. Looking carefully at the silences in the sequential environments described here, it should be noted that what could otherwise be understood as simple pauses in the course of the speaker's turn are actually gaps, in which some pursuit occurs. Moreover, gaze pursuit is so often successful that the return to the speaker's turn after a visible response becomes an indication of its effectiveness. Gaze often goes unnoticed.

---

28 Looking at eight different types of response tokens, Gardner (2001: 251) stressed that "each does distinctive interactional work, and that it is misleading to assume that they can be lumped together as a homogeneous group of 'minimal responses'". This is also true for the visible responses I am describing here.

For a recipient, the possibility of neglecting speaker gaze and the possibility of responding just with a gesture or a facial expression allows him/her to minimize the public aspect of the pursuit. If the recipient responds after a verbal pursuit, s/he will be heard as having not answered the first question, while if the answer comes after the gaze, he will simply be heard as having delayed his answer. And, the fact that often this answer is dispreferred makes the occurrence of a pursuit even less noticeable. By allowing a visible answer, speakers also permit an easier way out for the recipients. A disagreement can be resolved without explicitly stating that the participants hold opposite views on the matter. With a gesture or a facial expression the recipient can display her/his orientation towards the previous turn without making explicit the exact elements on which s/he disagrees. This means that what actually goes on the record for that specific conversation may be quite different from what is conveyed visibly.

From the speaker's point of view, by using gaze to pursue response s/he will not be heard to be asking "again", but rather as having increased pressure for a response, without any modification. Sometimes when the gaze pursuit fails, the speaker abandons the pursuit without using verbal turns and moves on with the conversation (in my dataset the pursuits that get abandoned after the first attempt are only gaze pursuits). Another advantage of using gaze pursuits for the speaker is the possibility of asking the recipient to respond without changing the direction of preference. It has been previously shown that one way participants orient to the delay in responding is by modifying the FPP in order to facilitate an answer by the recipient (e.g., Davidson, 1984). But gaze cannot and does not do that. Rather, gaze points back to what happened and displays the relevance of response. Nothing gets changed in the structure or semantics of the initial action that should be responded to. One possible implication is that the speaker assumes that the action has been perceived and understood, but what is missing is a response. Speakers appear to treat their prior turn as "problem free". Rather than assuming trouble for the recipient and therefore orienting towards the possible dispreference of the response, it is pursuing the action as it is and it is pushing for a response to what has just been said or done.

On a more general level, we can say that what favors the use of this practice in the sequential positions previously described is the different impact and interactional accountability of visible behavior compared to verbal behavior and the fact that gaze can be used before moving one step further and using talk to ask again, usually while sustaining the gaze towards the recipient.

## 3.8 Summary

In the first part of this chapter I provided evidence that the occurrence of speaker and recipient gaze during a FPP has different effects in terms of the provision of a response and its timely delivery. Speaker gaze works to elicit a response while recipient gaze works to encourage the timely delivery of the response. If the participants engage in mutual gaze, the occurrence of a timely response is even more likely. In the second part of this chapter I analyzed how gaze can be used to pursue a missing response in the transition relevance place following a FPP. In particular, I showed how the timing of gaze towards recipients is organized, its effectiveness, its possible limits and how it relates to other ways of pursuing a response, mainly by deploying verbal resources. Finally, in the third part of this chapter I showed how the knowledge accumulated in the first two parts of the chapter accounts for practices observable in other sequential environments in which the occurrence of a response can be relevant, although it is not normally considered conditionally relevant. I then showed how gaze plays a role in obtaining responses in these environments.

Further, I provided a characterization of sequential contexts in which a speaker manages to solicit and obtain a response by looking towards the recipient during silence. Although not exhaustive, these contexts can be divided into two broad types, with more specific characterizations:

1) In the transition relevance place:
    a. After a first pair part
    b. After a joke

2) In a silence mid-TCU:
    a. After a preliminary component of a TCU
    b. After a recognizable reference

These environments have been previously characterized as distinct in terms of the sequential implications of the completion of each of these items. On the other hand, they still initiate interactional formats that make response or a display of understanding as the relevant next action. Speaker gaze towards the recipient in these environments seems to solicit not just a response but also an analysis of the previous action performed by the person looking towards

the recipient, which then gets responded to in an appropriate way. Moreover, this practice is context sensitive; that is, recipient understanding depends on the interactional environment in which the practice occurs.

The first part of this chapter showed that the occurrence of gaze during a FPP can affect the likelihood of obtaining a response. The success of gaze pursuits further underscores this claim. Finally, the fact that the same practice (using gaze to pressure for a response) can be seen to work in similar ways across different sequential environments could suggest a reconsideration of the notion of "response relevance" as something not necessarily given but potentially constructed in each utterance by design. And when the recognition of response relevance fails or no answer is produced because it is dispreferred, gaze can come into play by upgrading the relevance of a response by the recipient.

As shown in chapter 2, the usual analysis of gaze as a tool meant simply to regulate the turn-taking system is not appropriate, because the exchanging of turns is not the main goal of the conversation nor the cause of a conversation in the first place, even though changes of speakership regularly occur. Rather turn-taking is a mechanism through which actions can be performed and reacted to. The sequential environment and the specific actions performed by the person who looks up towards the other must be taken into account, as these two things play a key role in who speaks next. The work by Torres et al. (1997) that shows how Kendon's suggested rules about gaze behavior and turn-taking are not as systematic as originally claimed, demonstrates that we are still far from having a predictive model of gaze behavior in interaction if we keep focusing only on the level of turns of talk.

Looking back at the practice of pursuing a response through gaze and considering the claims about the regulatory function of gaze, it might be claimed that gaze actually works as a sanctioning look, like "the Look" described by Mardi Kidwell (2005), rather than as soliciting response. That is to say, by looking at the recipient the speaker is sanctioning her/him for not responding earlier and the recipient responds because of the sanctioning. This does not seem to be the case. There is no evidence in the behavior of recipients that speaker gaze towards them is interpreted as sanctioning their behavior. There is never any apology or account for not having responded before, apart from example 3.26, in which the account about not remembering the movie is actually due to the fact that previously the recipient had claimed to remember it by mentioning the role of Bruce Willis in Pulp Fiction. What the gaze pursuit is doing, then, is not sanctioning but rather soliciting response.

Finally, looking at the different resources and modalities a participant can mobilize to obtain a response when a response is missing, we can appreciate even more the effort a

speaker makes any time s/he is producing a sequence-initiating action: make a move and pursue it if the recipient is not cooperative. Even though few FPPs remain unresponded to, some nevertheless do. Gaze, then, can be used to minimize the likelihood of this outcome, and, in this respect, its contribution to social interaction and human sociality is substantial.

# 4 Gaze Withdrawal in Sequence Closing

*One routinely presumes on a mutual understanding that doesn't quite exist. What one obtains is a working agreement, an agreement 'for all practical purposes'. But that I think is quite enough.*
Goffman (1976: 261) Replies and responses

## 4.1 Introduction

Turns of talk commonly form sequences of action, which are often further combined to form courses of action. Once a sequence reaches possible completion, it is sometimes expanded by producing further related talk, while in other cases it is not. A current claim about *when* sequences get expanded is that "preferred responses tend to lead to closing the sequence, while dispreferred responses regularly lead to expansion of the sequence" (Schegloff, 2007b: 117). As to *how* they get expanded, there has been extensive work on the verbal resources (e.g., *oh*, *okay*) and the action types (e.g., assessments, other initiation of repair) through which participants expand a base sequence (see Schegloff, 2007b for a detailed overview of these resources). However, little is known about the role that visible resources play in the *closure* of a sequence or a course of action and about the participants' online-orientations toward the relevance of (or likelihood of) a sequence expansion.

In chapter 2 I showed how recipients differentially deploy gaze while listening to the beginnings of different sequential organizations (i.e., an extended telling or an adjacency pair sequence), so I focused on the start of sequences and courses of actions. In chapter 3 I showed how speakers can use gaze to pressure for a response during the production of a first pair part (FPP) turn and to pursue responses when they are not forthcoming. In other words, I focused on what happens to a participant's gaze during a first action and between a first and a second action of a sequence. In this chapter, I investigate how people manage the interactional development of courses of actions[1] and the role that gaze plays in them by focusing on how participants manage the possible completion of a sequence and/or a course of action. Specifically I investigate how gaze contributes to the interactional achievement of

---

[1] "Courses of actions" is used here to refer to any interactional project that requires more than one turn to be fully accomplished. So an adjacency pair (FPP + SPP) becomes a course of action in the sense that without the SPP (e.g., acceptance, granting, etc.) the FPP cannot be considered accomplished or completed (e.g., an offer, a request, etc.). Anytime a participant reacts with something to make some first action successful is referred to as "a course of action".

sequence closure. I continue focusing mainly on sequences of talk and courses of action rather than single turns and speaker-hearer relationships within turns.[2]

Interactional actions commonly make relevant some responsive behavior or at least ratification by another participant in order to be treated as satisfactory. Understanding when a speaker is starting a new course of action rather than continuing a previous one is a difficult task, for both participants and analysts, and therefore requires sophisticated analysis of the conversation. It has consequences for what a recipient can and cannot do next. What appears to be the same item produced in the same sequential position could be followed by different sequential outcomes (sequence closure *versus* more sequential talk). For example, if after a FPP such as "do you want to play football today?" I say "it is a beautiful day", I will be heard as dealing with the question in some way, and the other participants will try to understand my response in terms of the possible relevance induced by its contiguous production (Grice, 1975; Levinson, 1983; Sacks, 1987; Schegloff, 1968; Sperber & Wilson, 1986). If, for example, I am only talking about the weather and I am not dealing with the prior speaker's FPP, then that same speaker may do something in order to deal with the possible misunderstanding (e.g., initiate repair).

We may wonder about when and why some sequences are expanded and others are not and whether sequence completion and silence is an interactional achievement or the natural byproduct of selfishly motivated behavior (e.g., the participant who should speak next does not take the floor or suddenly changes topic). For participants in interaction, it is crucial to understand and detect whether and when they need to say or do something in particular. These "degrees of freedom"[3] are clearly not only related to the turn-taking structure (e.g., whether the turn is possibly complete) but also to sequences of talk (e.g., whether what somebody has just said is a FPP, SPP or something else that does not make conditionally relevant further talk, but allows for it) and to sequences of actions. So, for example, if I make a request, I can start with a pre-pre turn (Schegloff, 1980) (e.g., Can I ask you something?), then I can move to a pre-request (Sacks, 1992 [1964-72]: 685-692 et passim, Vol. I; Terasaki,

---

[2] As the completion of a course of action usually co-occurs with the completion of a turn and the completion of a sequence, focusing on gaze during courses of action necessarily requires analyzing gaze behaviors occurring *during* turns. However, such turns will not be considered in isolation or in terms of how the rights to produce a next bit of talk are allocated, but in terms of their contribution and position within a larger structure: a course of action.

[3] This notion is not being used to refer to what participants "are capable" of doing (because in theory, one is capable of saying anything at any time) but to what participants will be "heard to be doing", for which they can be held accountable.

2004 [1976]) (e.g., Do you have plans for tomorrow evening?) and finally I can initiate the request (e.g. Could you give me a ride to the airport?). The possibilities of blocking or derailing a request are clearly oriented to by the requester who gives the other participant a chance to pre-empt her/his request with, for example, an offer. Understanding where we are in a course of action and the degrees of freedom we have at any interactional point provides us with tools to project what is relevant next. If somebody is telling us a story that has been launched with "something terrible happened to me yesterday" and in the midst of it we say "oh, that's really terrible", we might be misunderstanding where we are in the course of action that the teller had projected and we may also show that we mis-projected the place and time in which a display of understanding was relevant. Our ability to parse behavior and courses of action is summarized nicely by Byrne (2006: 478): "When we notice someone engaged in activity, we see not only how their body moves and what effects those movements are having on other things, but we also see what it *means*. The meaning of action includes *what is likely to happen next*, as a consequence of what has been done already; and what overall result is to be expected from the activity, in short, *why it is being done*". In this chapter, I analyze gaze behavior along the same lines.

My claim here is that gaze can be used to communicate where we are in a course of action when we approach a sequence boundary, at least in dyadic conversations. That is, *ceteris paribus*, whether we are ready to close the sequence and treat what just occurred or is occurring as a possible last component of the action-in-progress or whether we expect that action-in-progress to develop further.

This chapter will focus in particular on sequence expansion and sequence completion using an analysis of the relationship between gaze behavior and talk-in-interaction in naturally occurring data. I focus on what happens at possible sequence boundaries and I show how sustained gaze by even one of the participants in the transition relevance place (henceforth TRP) displays an orientation toward the relevance of more talk or general uptake by the other co-participant. On the other hand, gaze withdrawal at possible sequence completion displays an orientation toward the possibility of ending the sequence. This withdrawal allows both participants to re-orient toward other ongoing activities as primary in a local hierarchical structure of activities. It also marks what just preceded as a possible closure of the prior sequence of talk and the prior course of action or even a number of successive such sequences. In other words, it marks what preceded it as the possible

completion of an interactional project. This chapter, therefore, furthers our understanding not only of gaze but also of the sequential organization of social action.

In what follows, I first provide qualitative and then quantitative evidence for the claim that gaze withdrawal at possible completion of a sequence constitutes a bid for closing the course of action and displays an understanding of the current interactional project as complete. The first set of examples illustrates the practices concerning the role of gaze in sequence closure or sequence expansion. Then, the results of statistical analyses of a large corpus of data will be reported in order to show the validity and robustness of the claims. Further, all of the cases that do not appear to follow the predicted pattern will be examined and accounted for and the normativity of the gaze practices here investigated will be discussed. Finally, the implications of this work for research on visible behavior and human cognition in interaction will be addressed.

## *4.2 Background*

In chapter 1, I discussed existing claims about the relationship between gaze and turn taking (see, e.g., Duncan & Fiske, 1977; Goodwin, 1981; Kendon, 1967). In chapters 2 and 3, I argued that these claims are problematic because gaze seems to be organized around the actions that participants perform and the sequential organization of those actions. In this chapter, I address the issue of closure of courses of action and in particular the role that gaze plays in it.

Schegloff and Sacks' article entitled *Opening up Closings* (1973) was one of the first works to show how larger activities can be brought to closure, and it provided one of the first detailed analyses of the interactional resources deployed to accomplish such closure. In this work the authors described the details of closing a phone conversation by outlining not only one of the first accounts of the sequential organization of talk-in-interaction but also demonstrating the importance of understanding social practices as solutions to interactional problems. By showing that closing a conversation is usually not done abruptly, by dropping the telephone or by running out of the room, but is instead negotiated through sequential talk, Schegloff and Sacks showed that reaching social agreement on when something is over is complex, orderly and normatively organized. While dedicating most of their analysis to telephone conversations and "continuously sustained talk", the authors suggest that in face-to-face interactions of "continuing states of incipient talk", not every completion of a

sequence represents the end of the conversation. Continuing states of incipient talk occurs in situations such as two individuals eating together, or sitting in close proximity for an extended amount of time (e.g., people studying in the library, individuals sitting close to one another in situations such as travelling together or sitting in a waiting room). In these situations talk could emerge and fade and long silences can occur between spates of talk without the need for the participants to produce greetings or farewells (i.e., without the need of performing more ritualistic opening or closing of conversations). In these situations, the possibility of conversation remains open until the other participant(s) physically leaves the shared environment. Consequently, not every beginning or end of a sequence requires the deployment of specific verbal resources to mark them as such. Nonetheless, sequence closure is sometimes necessary to be able to re-engage competing activities or to move to a new course of action.

In further work aimed at unveiling how participants in interaction expand a sequence of talk in conversation and bring it to closure, Schegloff (e.g., 2007b) distinguished "minimal" and "non-minimal post-expansions":

> *The import of "minimal", however, is not limited to an arithmetical count of number of turns which happen to follow the second pair part [...] [but] rather, that the turn which is added is designed not to project any further within-sequence talk beyond itself; that is, it is designed to constitute a minimal expansion after the second pair part. It is designed to move for, or to propose, sequence closing (a move which may be aligned with by recipient, or not). (Schegloff, 2007b: 118)*

He terms these minimal post expansions "sequence closing thirds" which can take different forms or combinations, the most common of which are, in American English, "oh", "okay" and assessments. There are also what he calls "post completion musings", or "postmortems", which are utterances that can occur after the apparent completion of a sequence and yet they do not constitute the beginning of a new sequence but rather a sort of delayed analysis or assessment of the prior sequence and are usually produced as "outlouds" (Goffman, 1981c). Among the types of non-minimal post expansions perhaps the most common are the other-initiation of repair and related newsmarks (e.g., "is he?", "really?") and challenges of the SPP, all of which work to modify or clarify what was said or done with the SPP. There are also other non-minimal post-expansions that elaborate or modify what was done with the FPP of the base sequence, which Schegloff (2007b: 162) terms "first pair part reworking post

231

expansion". Finally, of particular relevance for the analysis presented in this chapter is what Schegloff (2007b: 181) calls "sequence closing sequences", which are sequences designed to close long sequences or topics. They are usually composed of three turns:

1) a turn that proposes the possible closing of the sequence or topic-in-progress [e.g., repetitions of words used in the launching of a sequence or a topic, summaries, assessments ("that's too bad", "it's too boring"), idiomatic or aphoristic formulations of the upshot of the sequence ( "they have a problem") (see Drew & Holt, 1998)];

2) a turn that collaborates in closing down the topic by affiliating with the stance displayed in the prior turn (e.g., a preferred response, such as "yeah");

3) a third move by the initiator of the sequence closing sequence that ratifies the recipient alignment with the closing proposal, usually by producing a closing token or assessment (e.g., "it was too much", "I don't know").

All of these different types of expansions and the resources to accomplish them show that achieving a sequence's closure is not a simple matter, and, it is often a cooperative process. This is true even in the case of foreshortened and "virtually unilateral" sequence endings, as they require the other participants to at least accept and ratify the closure without insisting on resumption of the previous sequence. The most important aspect of sequence closure (although, to some extent, the least obvious one) is that closure does not simply happen, but rather is achieved by participants, who deploy specific and recognizable forms of closure. The issue of the recognizability of forms of closure will be further addressed later in the chapter.

Keeping these aspects of sequence closure in mind, one important question is what role visible behavior plays in the recognizability of specific forms of closure. Besides Goffman (e.g., 1963) seminal work on how participants enter sustain and exit "focused" and "unfocused interactions" through their whole body, the first scholar who clearly introduced the idea that visible behavior can be deployed to mark shifts between units of behavior during an interaction was Scheflen (1964). By looking at how participants in psychotherapeutic interactions shift posture while talking, he suggested that visible behavior can provide a side view of what participants are projecting or how they understand what it is currently being done in the interaction. In a similar fashion, Schegloff (1998) showed how the postural

configurations of participants in a face-to-face interaction can display their current allocation of attention and commitment to possible competing engagements (e.g., talking to someone while also eating or working with a computer) in the surrounding environment. More recently Mondada (2006) has shown how participants in a meeting in an architect's office can multimodally project completion of a turn and project which next action they will produce themselves, though shifts to a new activity might be resisted by the other participants. Participants can project their incipient shift to a next activity gesturally and by moving objects on the table (plans of buildings) and as such modifying the current "contextual configuration" (Goodwin, 2000a). By placing these actions when turn completion is approaching, they can publicly display what they are about to do next and as such allow other participants to pre-empt or align with the suggested shift to a new activity. This idea of visible behavior as a tool for displaying participants' understanding and processing of conversational units will be central to the argument put forward in this chapter.

If we focus more specifically on the previous research conducted on gaze, one major work of interests is Goodwin's (1981) seminal *Conversational Organizations: interaction between speakers and hearers*. In this book he talks about gaze behavior as a way of displaying attention and (dis)engagement in the conversation, or more precisely, a display of the type of participation framework the participants are engaged in. From this perspective, looking away is noticeable and potentially sanctionable because it displays diminished engagement in the conversation. On the other hand, engagement in a competing activity (e.g., eating) provides a ready account for looking away rather than at the interlocutor's face and thus makes it less sanctionable. However, this claim also implicitly suggests that participants are relatively free to remove their gaze from co-participants, provided that they direct it toward another activity in which they are engaged. This dissertation has already shown and will continue showing in this chapter how this claim appears to be problematic. Indeed, as chapter 2 shows, some activities require more sustained gaze by the recipient toward the speaker (e.g. tellings) or by the speaker toward the recipient (e.g. questions) than others, suggesting that this relative freedom actually depends on the gaze expectations associated with the ongoing course of action. Moreover, this chapter will show that different sequential positions might indicate that gaze withdrawal should have different import, in particular with respect to the organization of gaze withdrawal at sequence possible completion. Finally, participants appear to calibrate their looking toward competing activities so that they do not disrupt the progressivity of the talk; that is, for example, they might look toward a glass and pick it up to drink when they expect they won't have to produce the next turn at talk (see,

e.g., example 4.7 in this chapter), so that drinking will not create a silence when talk would be otherwise expected.

Few more words about this groundbreaking work, however are necessary. In chapter 3 of *Conversational Organizations*, Goodwin presents claims that will strongly resonate with what will be shown later in this chapter, and yet differ from my claim in quite crucial ways. He claims, in particular, the following:

1. "Displays of mutual disengagement characteristically occur during lapses in the conversation.**"** (1981: 98)

2. "A speaker might use presence or absence of gaze toward recipient to display whether or not a next utterance is expected from the recipient. However, speakers look away from their recipients quite frequently during talk without in any way proposing that their recipient may/should start to disengage from the talk. [...] Recipients do not attend to the gaze withdrawal as an isolated event, but rather analyze it with reference to other activities [((not linguistic ones, but physical ones))] the speaker is performing at that moment." (1981: 104-105)

3. "[If a] speaker continues to gaze at recipient after his talk is brought to completion; recipients treats such gaze as proposing the continued relevance of conversation and puts aside another activity he was about to perform." (1981: 108)

The fundamental difference between his claims and mine consists in the conception of what gaze withdrawal does and to which conversational structure gaze is mainly related to. In his conception, gaze withdrawal displays diminished or complete disengagement with the conversation and usually allows engaging with a competing activity (e.g., smoking, eating). The meaning of gaze toward the other participant is that another utterance is expected from the other participant (quote 2), while a participant usually is not expected to continue gazing at the recipient when the former has brought the talk to completion (quote 3). Finally, Goodwin says that "most mutual disengagements occur during lapses in the conversation" (quote 1). One can easily notice how he refers to "talk completion", "next utterance", "disengagement with the conversation" and not to any specific structure in the conversation that the participant can systematically rely on to make sense of gaze withdrawal (see quote

234

2). To make sense of gaze withdrawal he suggests looking at what other activities the other participant might be engaged in, not to any specific structure present in the talk (see quote 2).

What I will show in this chapter is that gaze withdrawal can be deployed to coordinate sequence closure and achieve closure of a course of action. As such, coordinated gaze withdrawal contributes to generating the lapses in the conversation and it does not simply occur during lapses in the conversation, as a consequence of them. Moreover, it does not simply indicates that a next utterance is expected, but rather than some sequence expansion is expected, so sequentially appropriate next talk, not just continuation of the conversation. And the direction of both participants eyes (in a dyadic interaction) is not relevant vaguely when "the talk is brought to completion" (whatever this might mean), nor simply during the completion of any turn at talk, but more specifically in relation to moments in which sequence possible completion and possible completion of a course of action co-occur.

One additional piece of evidence concerning the different approaches here comes from a look at his transcripts, when compared to the ones you have seen in this dissertation and the ones you will see in this chapter. In his transcripts, Goodwin (1981) only shows a couple of utterances and what is happening to the participants' gaze at that specific point in time, but the reader does not know what had been said and done before both with talk and with their bodies, how the participants got there, and often what the course of action was. No detailed analysis is made, in that chapter, of what is going on in the talk, in terms of social actions and their sequential organization. To summarize, what is missing from Goodwin's account is a detailed description of how gaze behavior can be interpreted in relation to what is going on through the talk, rather than in relation to what else is going on in terms of competing activities. What is missing is a focus on the sequential organization of talk and its role in structuring what participants are doing with their talk and with their bodies. With their eyes, participants do not just display engagement and disengagement "with the talk", but with the specific course of action sequentially developed both through talk and visible behavior. And the "closure" that they are coordinating is not simply the closure of a topic, or of a stretch of talk, but rather the closure of a sequence or sequences of social actions, unified by a specific interactional project. So much so that it is not necessarily the case that mutual gaze withdrawal leads to silence and lapses (although often they do so) or to re-engagement in some competing activity. Sometimes it simply allows re-engagement in another interactional project developed through talk that had been interrupted or derailed to accomplish something else, also through talk (see, e.g., Example 4.4). Notice, however, that while I emphasize the importance of relying on a more fine-grained understanding of what participants are doing

through their talk, to establish more precisely at what level of order gaze is organized, my claim concerns the organization of social action (accomplished in whichever modality) in face-to-face interaction and is not limited to talk-in-interaction, as it might otherwise appear.

These specifications will become more meaningful as we advance through the chapter, however, it is of absolute importance emphasizing that they do not diminish the insightfulness and in many ways correctness of Goodwin's observations concerning the correlation between gaze withdrawal with displays of minimal engagement and potential attempt to closing an exchange.[4]

Expanding on their previous work on gaze, Goodwin & Goodwin (1987) talk about gaze withdrawal as a way of displaying diminished participation in the activity, and discuss assessments as a resource for closing topics and mutual orientation. Interestingly, though, in all examples shown in that paper, gaze withdrawal is not followed by immediate topic closure but rather by further talk (at least an additional TCU) either by the participant who withdraws gaze (if s/he was the speaker) or by the current speaker (if the person withdrawing gaze is the recipient). As will be shown later in this chapter, gaze withdrawal is a resource for making a bid for closure, or for displaying a specific understanding of the ongoing development of the course of action. By bidding for closure, participants display diminished participation in the activity.

To summarize, this chapter reports the initial investigations on conversational closure by looking at how sequences of talk and courses of action reach possible completion, when and how actual completion is achieved and what visible behavior contributes to it. In order to understand what happens at sequence completion, we first need to acknowledge that sequence completion also constitutes turn completion, and, as such, the issues of transition relevance and speaker transition need to be addressed.

---

[4] While the level of granularity differs, Goffman's focus in *Behavior in Public Places* (1963), in particular in terms of how participants display involvement in an interaction, is quite similar to the one Goodwin attempted to pursue in *Conversational Organizations* (1981).

## 4.3 Transition Relevance Places

In the previous chapters, turn constructional units (TCUs) and TRPs were discussed as the building blocks of turns and the place in which transition to a next speaker is possibly relevant, respectively. The notion of TRP, and therefore possible turn completion, is particularly relevant for the systematicity of the behavior that will be presented in this chapter. As discussed previously, a TRP is a delimited time period and not a singular moment in time. According to Schegloff (2007b: 4):

> As a speaker approaches the possible completion of a first TCU in a turn, transition to a next speaker can become relevant: if acted upon, the transition to a next speaker is accomplished just after the possible completion of the TCU-in-progress. Accordingly, we speak of the span that begins with the imminence of possible completion as the "transition-relevance place".

The fact that turn transitions do not only start at possible completion but also as a speaker approaches possible completion is empirically demonstrated by the timing of overlapping speech (see, e.g., Jefferson, 1973; 1983, 1986, 2004c; Schegloff, 2000a). The fact that transition relevance is not a point in time but rather a period of time becomes particularly important in terms of the projectability of possible turn and sequence completion and what participants can do during a TRP. In what follows, I analyze TRPs that lead to sequence closure and places in which, through gaze, participants leave open or even push for sequence expansion, although verbally the sequence could have reached possible completion. By doing so, this chapter shows how gaze behavior can be used to modify the interactional import of a TRP and thereby affect the development of a course of action.

## 4.4 Possible Sequence Completion

This section shows that when both participants avert their gaze from the other before or at possible sequence completion, the sequence is treated as complete and what follows is nearly always another course of action. On the other hand, when participants continue looking at each other once they have reached possible sequence completion, the sequence is usually treated as not yet complete, and it is consequently expanded until both participants avert their gaze or re-orient toward other activities at the next sequence possible completion.

Some categorical and terminological choices need to be explained here before proceeding with the analysis. First, if our focus is what happens at possible sequence completion, then a further specification is needed: what counts as a sequence expansion? In this chapter I will consider two types of expansions as constituting sequence expansion:
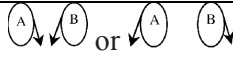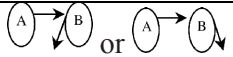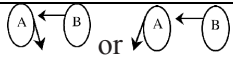
1) Prior speaker self-selects and adds a turn to the sequence;

2) Current speaker self-selects and adds another TCU to the turn that already reached possible sequence completion.

Both expansion types will be considered expansions of a sequence and it will be argued that gaze behavior has a possible impact on their (non-)occurrence.

Second, given that we are dealing with the behavior of two individuals at possible sequence completion, it is clear that the logical possibilities are not only 2 (i.e., no gaze or mutual gaze) but rather 4, including situations in which only the speaker or only the recipient look at the other (in the table, A= Speaker and B= Recipient): [5]

---

[5] Although I am using numbers to indicate the possible combinations, these four configurations should not be understood as ranked options.

*Table 4.1* List of possible gaze configurations at possible sequence completion, their symbolic representation, their meaning and examples that show how they work.

| Configuration | Symbolic Representation[6] | Meaning | Examples |
|---|---|---|---|
| 1 | (A) (B) or (A) (B) | Participants not looking at each other | 4.1-4.5 |
| 2 | (A)↔(B) | Both participants gaze at each other | 4.6 and 4.7 |
| 3 | (A)→(B) or (A)→(B) | Only current speaker continues gazing while recipient is looking away | 4.8 and 4.9 |
| 4 | (A)←(B) or (A)←(B) | Only current recipient continues gazing while speaker is looking away | 4.8 and 4.10 |

In what follows I will show how participants reach one of these gaze configurations at each possible sequence completion; that is, what happens to their gaze orientation before, and at, sequence completion. The four possible configurations represent a specific state in a specific interactional moment and the way in which participants get there might have important consequences in terms of how their gaze behavior is interpreted. Conversation is a dynamic process and any picture taken at a specific moment in that process, while valuable and crucial for quantification purposes, may also result in an oversimplification of the data. By showing how participants achieve these configurations, I hope to capture both the systematicity of gaze behavior at sequence completion and the complexity of the environment that participants quickly and efficiently navigate in order to implement appropriate gaze behavior at the right point in time.

One way of operationalizing these four possible configurations is distinguishing participants not simply in terms of A or B but in terms of their current participation role at the time the sequence reaches possible completion: that is, whether they are speaker or recipient

---

[6] While there is only one way of looking at one's face in terms of gaze direction, there are many ways of "not looking" at someone, including closing the eyes or looking at some object in the surrounding environment. For simplicity, in Table 4.1 I provide only two of the many symbolic representations of "not looking at the other participant". Other ways can be found among the gaze symbols in Appendix B.

of the last turn. Indeed, it is clear that these distinct participation statuses offer different options for participants in terms of sequence expansion: the current speaker can expand the sequence by adding another TCU to her/his turn, while the current recipient can self-select and produce a new turn of talk that expands the sequence.

In all of the examples that will be presented, the sequences will end with gaze withdrawal by both participants (or, at least with participants not looking at each other, if they withdrew before sequence completion). This provides additional evidence for the claim that the function of gaze configuration 1 is bidding for closure. In some examples, multiple sequence possible completions will be shown, which means that more than one gaze configuration might be implemented by the participants during the course of action presented.

Before presenting some examples of these four configurations, one final point is in order. When participants reach possible completion of a sequence, they might also reach a point in which the course of action underway could be considered completed. If, for example, someone requests information, once this information is provided the course of action could be considered complete (Schegloff, 1968, 2007b). Often the successful completion of a course of action entails the occurrence of multiple sequences of talk linked together by the more general gist of what the participants are doing together (e.g., booking a flight on the phone, Lee, 2008). In addition, there are courses of action that are not continuously sustained and can be momentarily abandoned, only to be returned to later in the conversation (e.g., instruction giving, gossiping, issuing multiple related requests). If this is the case, one should consider each "episode of a course of action" as a place that needs to reach some closure before the conversation can move to another episode or to other courses of action. The focus of this chapter, then, will be the achievement of closure of "episodes of courses of action". And, these courses of action can have just one episode, clearly delimited as a unit in the conversation, or can be constituted by multiple episodes that take place in different moments of the conversation.[7]

Apart from greetings and farewells, a look at the structure of conversational exchanges in face-to-face interaction shows that many other sequences are complete with the production of just two turns of talk, one by the initiator of the action and one by the person responding to it. And yet, it often happens that the sequence gets expanded. How this

---

[7] Notice here that I am treating "episodes of courses of action" as a unique category, whereas Schegloff (2007a, 2007b) has otherwise divided and labeled these as "base sequence (adjacency pair)s", "expanded sequences", "sequences of sequences" and "larger threads". The reason for subsuming them under one label is that participants' gaze behavior seem to treat and contribute to their closure in the same way, without requiring further distinction. Evidence for this claim will be provided later in the chapter.

happens will be discussed in the following section. A closer look at courses of action that are accomplished through only two turns of talk shows something systematic happening around their possible completion: participants withdraw their gaze from one another and tend to enter silence following turn completion. What follows is usually some engagement in another course of action or (re-)engagement in another course of action.

## 4.4.1 Configuration 1: Gaze Down at the Possible Completion of a Sequence

This section shows how configuration 1 in Table 4.1 can be achieved, namely: a course of action reaches completion after a SPP and, at possible sequence completion, the participants are not looking at each other. It also illustrates how participants' gaze behavior reveals their projection of possible sequence completion and how they interactionally manage the potential completion of the sequence.

The most obvious possible reason why participants are not looking at each other at possible sequence completion is that they had already not been looking at each other during the sequence. Example 4.1 illustrates a two-turn sequence produced while two young men are looking at pictures A took during his holidays. While flipping through them, B sometimes stops and asks questions about the pictures. In this excerpt, B asks A about an object in the picture and points to it with his index finger.

**(4.1) 2PCOMP-carta 10:17**
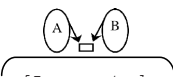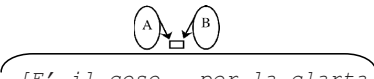
```
01      (1.0)


                     Ⓐ  Ⓑ
                       □
02 B:   Cos'e' quella roba  la'   ((indicando))
        What is that  thing there
        What is that thing there   ((pointing to the picture))


                  Ⓐ  Ⓑ
                    □
03       (0.6)
```
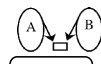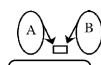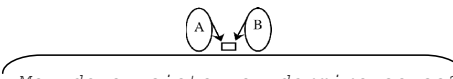
```
                              A      B
04 B:    [La carta]
         The paper
         [The paper]

                              A      B
05 A:    [E' il coso   per la c]arta
         Is the thing for the paper
         [It is the thing for the p]aper

                              A    B
06       (0.3)

07       ((B turns picture album and they start looking at new picture))

                              A    B
08       (1.3)

                              A      B
09 B:    Ma   dove  siete  a  dormire scusa?
         But where are.2p to sleep    excuse me
         But where are you sleeping excuse me?
```

B's request for information (line 2) gets answered by A (line 5) at the same time as B offers a candidate answer himself. The speaker of the initiating action (B) continues looking at the picture for 0.3 seconds after the answer and then shifts his gaze toward the next picture. He then asks A the question at line 9 and starts another sequence of talk unrelated to the previous one. All of this is done without any gaze between the participants. Indeed, they sustain their orientation toward the pictures they are both looking at, and, at possible sequence completion, neither participant looks at the other and what follows is a new sequence.

If participants do not look at each other during the course of a sequence it is not surprising that they do not look at each other once that sequence reaches possible completion. Example 4.2 shows a case in which participants engage in mutual gaze and withdraw from it before reaching possible completion of the sequence, but still at a place in which the possible

242

completion is projectable. The result is that both participants are not looking at each other once possible completion is actually reached.

In the following excerpt, two young men are sitting in B's room and B is helping A prepare for an exam. At the beginning of this fragment, B is looking at a computer screen apparently searching for useful information for A. B has just said that the book called "Treetti" is useful for one subject that they have to study (radiography), but not for the one they have been talking about (ultrasound) (lines 1-2). B then suggests, however, that it is the book that they have to use to prepare for the exam (line 8). Our target is the sequence that starts at line 6 and ends at line 10.

**(4.2) 2PEXAM-letto 45:49**

```
01 B:   pero' il  Treetti va   bene per radiografia.
        but   the Treetti goes well for radiography
        but Treetti is good for radiography.

02      Per ecografia  non dice niente
        For ultrasound not says nothing
        Concerning ultrasounds it does not say anything

03      (0.3)
```



```
04 A:   mm
        mm
        mm
```



```
05      (0.2)
```



```
06 A:   Ma  l'hai      letto tutto il  Treetti¿
        But it have.2s read  all   the Treetti
        But have you read it all the Treetti¿
```

07    (0.2)   (2.2)

       **(2.4)**

08 B:  *No   del Treetti devi fare le prime: Cioe'   le prime pagine*
       No   of  Treetti must do   the first I mean the first pages
       **No  of Treetti you must do the firs:t I mean the first pages**

09     *che  poi  sono (1.0) (0.6)   (0.2) quelle che: son*
       that then are                       those  that are
       **that then are         (1.8)        those tha:t are**

10     *schematizzate  sugli appunti di Marini.*
       schematized    on the notes of Marini
       **schematized on Marini's notes.**

11     (7.0)
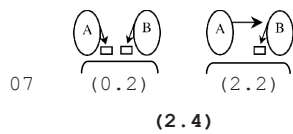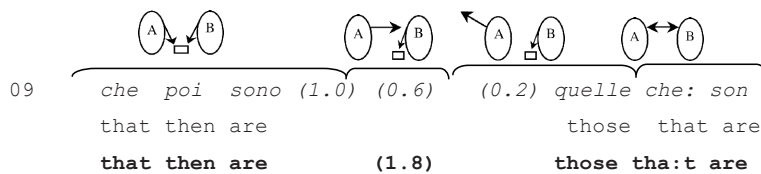
12 B:  *Pero'   cioe'  e' bella radiografia.*
       However I mean is nice  radiography
       **However I mean radiography is nice.**

A's question at line 6 is topically coherent with B's prior talk but constitutes a new sequence given that it is an indirect question about how much of the book should be studied. The occasion for their meeting, and the general activity they are involved in (exam preparation) are the background for understanding B's answer to A's question and why this is not treated as problematic. B does not treat A's turn at line 6 as a request about how B knows about the information he just provided about radiography and ultrasound, but rather as a return to business and, as such, as a request for relevant information in terms of preparing for the exam.

The participants are not looking at each other during B's production of the request for information at line 6, but B turns toward the recipient during the silence, after it is clear that the answer will not be provided promptly. During the silence at line 7, B looks at the recipient to pursue uptake (see chapter 3). B's response at line 8 is type conforming in its beginning (i.e., he answers 'no') (Raymond, 2003) but it is also dispreferred. At the same time, a simple "yes" or "no" by B would not display appropriate recognition and understanding of the action line 6 implements. A does not look away after the "no"[8] but rather keeps looking until the initiation of self-repair at line 8 "I mean". The turn at line 8 is syntactically possibly complete after "pages", though pragmatically saying that A has to study only the first pages of the book is not really helpful unless the specific pages are specified. Mutual gaze occurs when the most important information of the turn is delivered: the pages to be studied are those presented in the schema in the professor's notes. Once this information is underway, both participants withdraw from mutual gaze before B utters the last few words of the turn (Marini's notes), yet when those last words are easily projectable by A. Marini is the name of the professor who teaches the course A has to prepare the exam for. By the time he hears "the first pages, that are the ones that are schematized" he can recognize the reference by B, which in this case can only be the professor's notes. The speaker of the sequence initiating turn is the first to look away and the other immediately follows. They both look away before reaching the actual completion of the turn but at a place in which the point of completion is projectable. In this example, then, we see that both participants are looking away before reaching possible completion of the sequence and what follows is not an expansion of the sequence but rather a long silence and then a completely different sequence.

Example 4.3 shows a sequence in which the participants do not look at each other during the sequence initiating turn and they engage in mutual gaze only in the responsive turn. Indeed, the last person looking withdraws her gaze exactly at sequence possible completion and they enter a silence until a new course of action begins. In this excerpt, two participants co-complain and then visually display to each other that the sequence is complete. The participants are sitting facing each other and they have been studying together for a while. A few minutes earlier B asked A to stop studying for the day because she has a

---

[8] It will later be shown that if the sequence initiating action is simply a request for confirmation, then the speaker who sought confirmation will systematically look away after obtaining a confirming answer. In this case, the answer is disconfirming and the turn at line 6 is not a request for confirmation but rather a request for information. The participants' gaze behavior, then, conforms to the action implemented and to the relevance of further explanation by the current speaker.

headache. Before the beginning of this fragment, the two of them were joking about what A has done while B was on the phone with her mother. The target sequence is the one between lines 9 and 11.

**(4.3) 2GSTUDYING-testa  29:07**

```
01 B:  hh [  ha                 ha   ]
       hh [  ha                 ha   ]

02 A:     [Q(h)uello che  devi    leg]gere
           That      that must.2s read
          [You have to re]ad

03     (0.4)

04 A:  hh hu
       hh hu
05     (0.9)

06 B:  .hhh
       .hhh

07     (0.7)
```
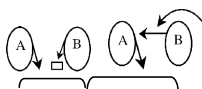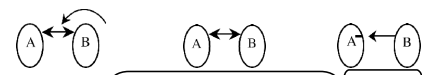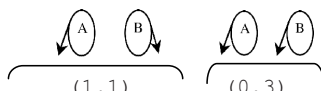
08 B:  *Okay quindi   (0.3)*
       Okay therefore
       **Okay therefore (0.3)**
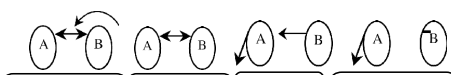
09 A:  *.h No io veramente son scarburatissima ho     un mal di testa*
           No I   really    am  out of steam   have.1s a  bad of head
       **.h No I am really out of steam I have a headache**

10      *anch'io*
        also I
        **as well**

11 B:   *Anch'io  c' ho      mal di testa*
        Also I   cl have.1s bad of head
        **I have a headache too**

12          (1.1)        (0.3)
              **(1.4)**

13 A:   *Poi  ho     detto sara'   il raffreddore sara': il-*
        Then have.1s said  will be the cold       will   the
        **Then I said it will be the cold it will be: th-**

14      *i   postumi dell'influenza, (0.5)*
        the effects of    flu
        **the side effects of the flu, (0.5)**


At line 8 B responds to the joke initiated by A before the beginning of this extract. The turn at line 9 starts with an inbreath and a "joke-to-serious" 'no' (Schegloff, 2001), which marks the upcoming talk as not part of the joke but rather as something serious. A continues the turn by producing two complaints about how she feels, and B responds to the second one. Then silence follows, and the sequence could be complete at this point. At line 13, however, A starts a telling about what she thought could have been the cause of her not feeling well.

Focusing on lines 9-10, we see that the recipient, B, is not looking at the speaker during the two complaints and only begins to turn after the TRP at line 9 "mal di testa" (headache) and immediately before responding at line 11. During the responsive turn at line
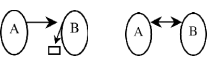
11, the recipient, A, looks at the speaker, B, for most of the time and closes her eyes only when the trajectory of the turn is clearly projectable (the speaker repeats the same words used by A during the second complaint). The current speaker, B, is the last one who looks away and she does so as soon as she completes the last word of her turn. Similar to the previous examples, the participants are not looking at each other once the possible completion of the sequence has been reached and the sequence is not expanded. Also, as in the previous example (4.2), the gaze withdrawal is not simultaneous but rather serial: first, the speaker of the sequence initiating turn withdraws gaze, then the recipient of that turn does so.
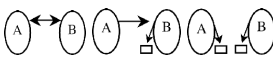
In example 4.3 a bit of silence follows the completion of the sequence and only later talk is resumed. If the sequence is not produced as a completely new course of action but rather as a side sequence, what usually follows the completion of the sequence is not a lapse, but rather a resumption of the previous course of action. The next example (4.4) shows how participants coordinate their gaze withdrawal approaching the possible completion of a side sequence. In this example, two girls are chatting while sitting at a table before going out for a drink. They have not seen each other for a few weeks and are catching up. Before the beginning of this fragment, B, who is visiting A, counters A's prior statement about the sports at the Olympic games (i.e., that they are boring and the sports all look the same). The sequence gets derailed by a change in A's posture that attracts B's attention. B asks whether A is trying to do some "zen" and they laugh about it, especially when A makes explicit that she doesn't even know what "doing zen" means. The excerpt begins with A's turn at line 9, which follows laughter by both participants. Notice the pattern of gaze behavior at lines 10 and 11 and what happens at the possible completion of the sequence at line 11.

**(4.4) 2GC-zen 30:08**



```
                            A ──► B
09 A:  (Tu  non s(h)ai) zen.  [(my) zen
        You not know.1s zen    my   zen
       (you don't k(h)now) zen. [(my) zen
```

248

```
                                      A → B        A ↔ B
                                           ▢
10 B:                               [Mi  fai   finire¿
                                     Me  make.2s finish
                                    [Do you let me finish¿


         A ↔ B  A → B A ▢B ▢B
                   ▢
11 A:    S c u s a h=
         Sorry
         Sorry h=


                    A ▢B    A ← B    A ↔ B    A → B
                      ▢        ▢
12 B:    =hhu allora ginnastica ritmica sono tutti diversi,  basket
              alright gymnastice rytmic are  all   different basketball
         =hhu well then gymnastics they are all different, basketball


                              A ↔ B
13       cosa e' tutto uguale¿ E' una partita di basket.   Pallavolo
         what is all   same   Is a  game   of basketball Volleyball
         what is all the same¿ It is a basketball game. Volleyball
```

At line 10 B sanctions A for not letting her continue talking about the Olympic games. A offers an appropriate response at line 11 by apologizing, which demonstrates recognition of the action performed by B's talk as a rebuke. The apology is followed by a minimal laugh token and B reciprocates this laughter by producing a laugh token at the beginning of her talk (line 12). The laugh tokens arguably display A's and B's stance toward B's possibly confrontational move (i.e., the sanction). Specifically, both of them seem to treat it as non-serious, mitigating the possibility of further conflict. Thus, it becomes a meta-comment about the prior sequence.

The course of action (sanction-apology) is complete after "scusa" (line 11), and we see again a correspondence between the completion of the sequence and gaze withdrawal by both participants. At the beginning of the sanction, A and B engage in mutual gaze, which extends through the TRP at the end of B's rebuke. During the last syllable of "scusa" (sorry), B lowers her gaze and looks down at the table. Immediately after this move by B, A lowers her gaze too, and this occurs in the TRP during the laugh token. In this example, the

withdrawal of gaze does not lead to a lapse, but instead to a re-engagement into a larger course of action that had been momentarily suspended. Still, the gaze withdrawal occurs at sequence possible completion and the side sequence is indeed treated as complete.

In example 4.3, silence follows the completion of the sequence and a completely new course of action is produced next. Example 4.4, instead, shows a side sequence at which completion participants re-engage the previously interrupted course of action. The following example (4.5) shows that participants can withdraw their gaze from the other at sequence completion, though the current speaker adds a TCU to the turn that is topically related but not action-related to the sequence just completed. The participants in example 4.5 are the same as in example 4.3, though in this excerpt they are talking about their friends. B has just finished her glass of iced tea while A is telling a story. At line 2, B makes explicit her desire to get another glass (the bottle is sitting in between A and B on the floor). The target sequence for the relevant gaze behavior is between lines 5 and 7.

**(4.5) 2GSTUDYING-vuoi 36:28**

```
01 A:   E' un trattore che  ti  passa   so[pra]
        Is a  tractor  that you pass.3s above
        (She) is a tractor that runs ov[er you
```



```
02 B:                                   [Ne] vorrei un altro po'
                                         It want.1s a  other bit
                                        I would like another bit of [it]
```



```
03      (0.3)
```
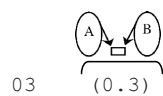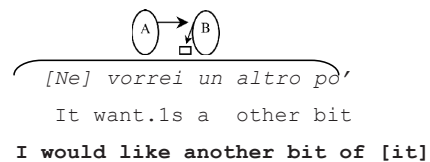


```
04 A:   Pr[ego  prego]
        Please  please
        Pl[ease please] ((please go ahead and help yourself))
```

05 B:   *[Tu lo vuo]i?*
        you it want
       **[Do you wan]t it?**

06    (0.1)   (0.1)
       **(0.2)**

07 A:  *hh No grazie °Bevi    pur te°*
        No thanks drink.2s    you
     **hh No thanks °Please drink°**

08    (0.7)

09 A:  *tlk! E  nel senso che  tipo quando io mi son lasciata con  Marco*
        And in  sense that like when  I  me  left.1s   with Marco
     **tlk! And meaning that like when I broke up with Marco**

At line 2, B requests more iced tea and at line 4, A grants permission for her to help herself. During the silence at line 3, B lowers her head and torso by moving toward the bottle of tea on the floor and looks in that direction. The turn at line 4 is overlapped by the offer at line 5 during which A looks at the speaker. During the following silence, A looks away while B starts looking up toward her and they never engage in mutual gaze. A rejects the offer at line 7 orienting first toward the format of the prior turn (as a question making relevant a response) and then to the action (an offer) it is a vehicle for (no→yes-no question but also declining the offer, thanks→offer) (Schegloff, 1984, 1995, 2007b). At this point the offer sequence is possibly complete and indeed the recipient of the last turn (B) turns her head toward the right, looks away and picks up the bottle of tea. The lifting of the bottle of tea co-occurs with the TCU produced with soft voice that A adds to her turn: "bevi pur tu" (please you drink). The fact that it is produced sotto voce, and that it is again a go-ahead for B who is already picking up the bottle, makes the turn either redundant or potentially a re-doing of line 4 (which had

251

been produced in overlap). Given that A is the host and B the guest, the TCU produced sotto voce saying "bevi pur te" might be a way of enacting being a proper host by insisting that the guest goes ahead and helps herself with the tea. By the time the offer is rejected, both participants are looking away and what follows is an invitation to proceed with the drinking. At possible sequence completion, then, both participants are looking away from each other and the offer sequence is complete. In this specific case what follows is not immediate re-engagement into the story telling that had been interrupted by the side sequence, but rather a redoing of the granting of the request (line 7), that had been previously overlapped (line 4). Then a bit of silence follows and the story telling is resumed.

The five examples here presented have in common that when the sequence reaches sequence possible completion, neither participant is looking at the other and what follows is not an expansion of that sequence but rather a new course of action or the resumption of a course of action that had been interrupted by the very occurrence of the target sequence. As has been shown, the way in which participants get there can differ. Later in the chapter I address the issue of who usually initiates the withdrawal and why this does not occur simultaneously by both participants. For the moment, the evidence presented shows that when both participants coordinate their gaze behavior so that neither is looking at the other at sequence possible completion, both parties treat the sequence as complete. I now provide additional evidence for this claim by presenting examples of what happens when participants keep looking at each other at possible sequential endings.

## 4.4.2 Configuration 2: Gaze Up by Both Participants at Possible Completion of a Sequence

The previous section showed that when two-turn sequences do not get expanded, either before or at possible sequence completion, both participants withdraw from looking at each other. The claim is that looking away before or at completion of a responsive turn displays an orientation toward the possible completeness of the sequence and therefore constitutes a bid for closing the sequence. If this claim is correct, then two patterns of behavior should be observable:

1. If participants continue looking at each other at possible sequence completion then the sequence gets expanded;

252

2. when an expanded sequence finally reaches completion, we should still see that participants withdraw their gaze by the time of actual sequence completion.

The following cases exemplify these patterns and the second gaze configuration: both participants keep looking at each other. In example 4.6, the participants are looking at pictures together and at the beginning of this fragment they are looking at a picture of the dome of St. Peter's church in Rome. The target here is the gaze behavior at possible sequence completion at lines 12, 13 and at the end of line 14.

**(4.6) 2PCOMP-pagato 9:33**
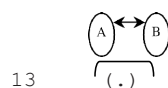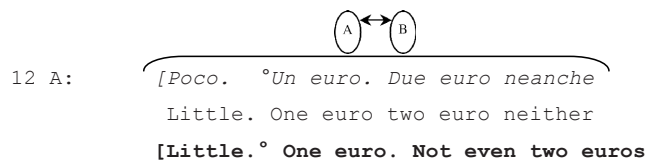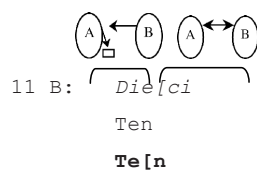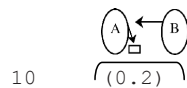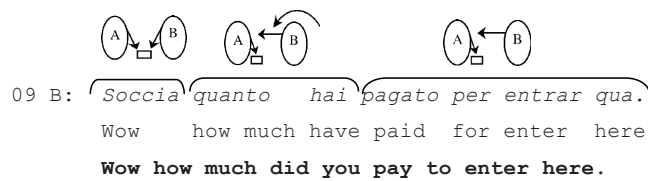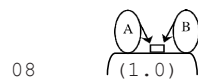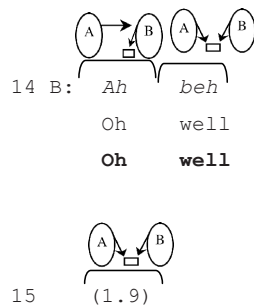


```
08      (1.0)
```



```
09 B:   Soccia quanto   hai pagato per entrar qua.
        Wow    how much have paid  for enter  here
        Wow how much did you pay to enter here.
```



```
10      (0.2)
```



```
11 B:   Die[ci
        Ten
        Te[n
```



```
12 A:      [Poco.  °Un euro. Due euro neanche
           Little. One euro two euro neither
           [Little.° One euro. Not even two euros
```



```
13      (.)
```

```
14 B:  Ah      beh
       Oh      well
       Oh      well
```



```
15     (1.9)
```

At the beginning of this extract, both participants are oriented toward the pictures on the table (see line 8). At line 9, B first produces an appreciative outloud (Goffman, 1981c) "soccia" (wow), most likely about what he is seeing in the picture (the dome of St. Peter's church in Rome). Then he asks how much it costs to visit the church. However, the recipient does not look at the speaker at all during this turn. While producing the beginning of it, the speaker shifts his gaze and head orientation from the pictures to A. At line 9, B turns his head and looks at A while A continues looking at the picture through the following silence and only looks at B when he begins to answer at line 12. Similarly to example 4.2, here too the answer is further refined during its delivery (it goes from "little" to "one euro" to "not even two euros") and the participants sustain mutual gaze throughout the turn. The first point of possible completion is after "poco" (little)[9] and the second one after the next bit "1 euro". The participants keep looking at each other after both of these. Speaker A adds another TCU to his turn, which becomes the next point of sequence possible completion (not even two euros) and still the participants sustain the mutual gaze. They keep looking at each other also during the following silence until B starts line 14 that constitutes a sequence closing third turn (Schegloff, 2007b), the most minimal expansion of a sequence. During the turn at line 14, the first person to look away is the person who started the sequence at line 9 and then A looks away immediately afterwards, at turn completion and they re-engage the activity that had been momentarily interrupted (looking at pictures). As was the case in the previous examples, the gaze withdrawal is not simultaneous and the first person who withdraws from mutual gaze is the person who started the sequence and the other immediately follows, thereby treating the sequence as complete.

---

[9] Like example 4.2, pragmatically it is not particularly informative to respond with "little" to a question asking "how much". However, the out-loud and the question at line 9 may indicate an expectation of an answer such as "a lot of money". If this is the case, then "little" is potentially sufficient to counter the assumption that seems to be built into the question. The fact that A produces a more specific answer, on the other hand, suggests that either he is treating the first TCU of his turn as not adequately informative or he is unsure and therefore deals with both possibilities.

254

Example 4.7 shows a similar pattern of participants' sustained gaze followed by gaze withdrawal before completion, only in this case, one participant looks away before approaching possible completion and the other waits to withdraw until the last syllable of the last word of the turn. Moreover, this sequence is expanded twice before it is actually treated as complete. Before entering the 15-seconds lapse (line 1), the participants were discussing the location of a cinema and the fact that this was news for A. During the lapse, both A and B re-engaged eating, the other ongoing activity. They did not look at each other during this silence and instead remained oriented toward their dishes until 0.6 seconds before B starts speaking. B looks at A before speaking and continues looking at him into her turn. The focus here is the participants' gaze behavior at the end of the turns at lines 4, 6 and 8.

**(4.7) 2PLUNCH1-lezione 4:28**

```
01          (14.4)      (0.6)
                (15.0)

02 B:   Che      lezione    hai
        Which    lesson     have
        Which lesson do you have

03          (0.3)   (0.5)
                (0.8)

04 A:    tch!    (0.3)   Trigonometria
                         Trigonometry
         tch!    (0.3)   Trigonometry

05          (0.7)
```

255
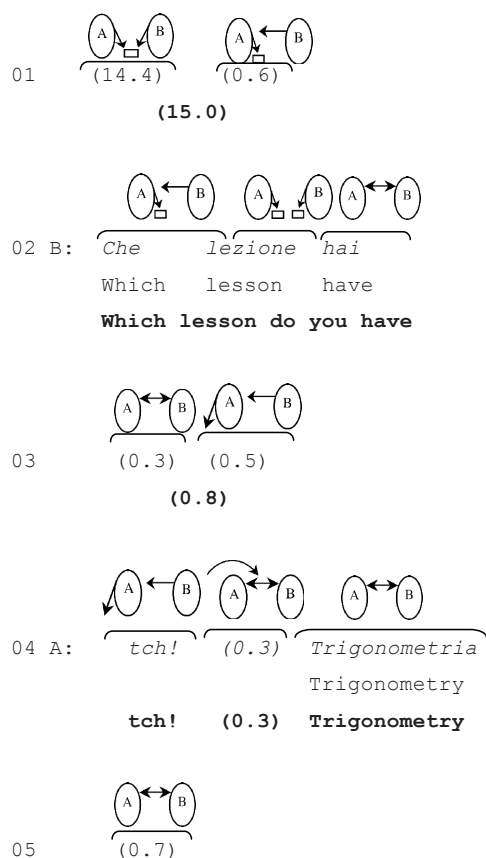
06 B:  *Che palle.*

That balls

**How boring.**



07     (0.9)



08 A:  *Laboratorio.=  E' l'ultima volta prima dell'esame.*

Laboratory     Is the last time  before the exam

**Laboratory.=It is the last time before the exam.**



09     (3.5)



10 B:  *E   l'esame    cos'e'?*

And the exam   what is it

**And what is the exam about?**

While B inquires about a lesson A has to attend in the afternoon, A wipes his lips with a napkin. Before answering, he puts down the napkin while gazing at B. Then, he turns to the right averting her gaze, produces a click sound (tch!) while apparently cleaning his teeth with his tongue and then turns back to her and answers the question: "Trigonometry". At the end of line 4, the sequence is possibly complete but both participants sustain mutual gaze. A then moves his right hand toward his glass on the table and B continues taking a bite of bread. If the sequence were complete, it would be possible for them to *just* do these actions, withdraw from mutual gaze and fully orient toward these competing activities. However this is not what happens. During the silence at line 5 B chews his food and then offers a third position assessment of A's response: "Che palle" (How boring). Thus, at a possible sequence boundary participants are in mutual gaze and one expands the sequence with a minimal post-expansion. At the end of line 6, the sequence again reaches possible completion. However, at

256

that point, and during the silence at line 7, B and A again sustain mutual gaze and what follows is another expansion of the sequence, this time by A.

During the sustained mutual gaze across the silence at line 7, A brings a glass of water toward his mouth. At this point, he could be drinking, and therefore display that he cannot, at least not in that moment, produce further talk, or he could hold the glass and say something else. This sustained gaze displays an orientation toward the sequence as not yet complete and invites sequence expansion by the other participant. Moreover, B's assessment at line 6 is a negative one and invites a second assessment that affiliates or disaffiliates with it (Pomerantz, 1984a). At that point, A does not lower the glass nor does he start drinking; instead, he holds it just in front of his face and produces the two TCUs at line 8. Both TCUs account for the fact that he has to go to this class. The first TCU ("laboratorio", laboratory) refines the responsive turn at line 4, in the sense that it further specifies what the lesson will be about, but both TCUs deal with the assessment produced by B. They appear to be either a way of disaffiliating with the idea that the lesson is going to be boring or a way of justifying his going, even if it is boring.[10]

At the end of the turn at line 8, the sequence is yet again possibly complete. Both participants have withdrawn from looking at each other; A starts drinking and what follows is a lapse of 3.5 seconds and then a question that initiates a new sequence. The silence at line 9 differs from the previous silences in the sequence in that neither participant is looking at the other nor at a common object. Rather, both look toward the next relevant object that can also stand for the next relevant activity: A looks at the glass and starts drinking. B looks down at her plate and puts the fork with the food into her mouth. In this example, at different points of possible completion of the sequence, the participants continue looking at each other, and what follows is not a new sequence, but an expansion of the current sequence. In this way, the participants display an understanding of the interactional project as not being over until the completion of A's turn at line 8, when both participants have finally withdrawn from looking at the other.

Contrary to what we observed in all of the previous examples, in this case the speaker of the current turn (recipient of the sequence initiating turn) withdraws his gaze after the first

---

[10] A piece of data from the same interaction shows that the latter is the best explanation for A's turn at line 8. Approximately 30 minutes after this fragment, when A is preparing to leave, he says that he does not feel like going to class and B immediately suggests he does not go and stays with her. A's response is that B should convince him to go and not insist that he does not go to class. This explains what could have driven A in responding with the turn at line 8 rather than with an assessment: pre-empting a recognizable course of action (i.e., B's attempt to convince A not to go to class and stay with her instead).

TCU, while the recipient withdraws her gaze later, on the last syllable of the last word of the turn. We will later see in which cases it tends to happen that the recipient of the sequence initiating turn is the first to withdraw gaze (usually in the case of disagreements or when a participant wants to end the course of action unilaterally).

The two examples presented in this section provide further evidence for the claim that, by withdrawing their gaze from the other participant before possible sequence completion, interactants treat the sequence as complete. Specifically, we saw that if interlocutors continue looking at each other at possible sequence completion, the sequence typically gets expanded.

## 4.4.3 Configurations 3 & 4: Gaze Up By One Participant at Possible Sequence Completion

As indicated in § 4.4, in addition to the two possible configurations previously discussed, in which both participants do the same thing (i.e., both withdraw from looking toward the other or both maintain looking toward the other), it is possible that only one of the two participants continues looking while the other has already looked down or away once possible completion of the sequence is reached. In what follows, I show examples of these gaze configurations (3 and 4, respectively). Here we ask whether sequence expansion occurs only when *both* interactants treat a sequence as incomplete or whether *one* interactant is sufficient.

If just one participant shows an orientation to the sequence as not yet over, the sequence normally gets expanded. Example 4.8 shows that mutual gaze is not required for the occurrence of sequence expansion and it shows examples of configurations 3 and 4. Here a multi-unit turn is in progress. The sequence in question is a departure from this turn-in-progress [an "oblique sequence" (Koenig, 2007)] that could potentially derail the development of the multi-unit turn. At line 1, A is explaining to B how to prepare a pasta recipe and how much pasta he will need. Lines 1 and 2 are part of the multi-unit turn that constitutes A's explanation. Our target is the gaze behavior between lines 3 and 9.

258

**(4.8) 2PPLAN-settanta 10:10**
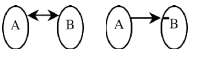
01 A:  *beh   vabbe'   in quaranta sono poi quattro kili*
      well  alright in forty     are   then four     kilos
      **well alright for forty ((people) it is then four kilos**

02  *di pasta non e' che siano p::oi tan[ti*
      of pasta not is that are   then   many
      **of pasta which is not then so:: mu[ch**

03 B:  *[°Quanti erano quand'eravamo*
       How many were when were.1p
      **[°How many were when we were**

04  *via  eran    di piu'? Settanta sessanta [settanta persone°]*
      away were.3p of more  seventy  sixty    seventy  people
      **away there were more? Seventy sixty [seventy people°]**

05 A:  *[Erano sessant:otto]*
      Were.3p sixty-eight
      **[There were sixty:eight]**

06  *mi sembr[a*
      me seem.3s
      **it seem[s to me**

07 B:  *[mm*
      mm
      **[mm**

```
        ⌢A→B⌢ ⌢A↔B⌢
08      (0.3)   (0.1)
            (0.4)


        ⌢A↔B⌢ ⌢A↙B←⌢
                 □
09 A: ⌢°Sessanta intorno°⌝
        Sixty    around
      °About sixty°


        ⌢A↙ B↓⌝
          □
10      (0.6)


        ⌢A↙ B↓⌝ ⌢A↙←B⌝          ⌢A↓←B⌝
          □       □
11 A: ⌜E::::  (0.6)⌝ ⌜(2.2)⌝ non so    quanto  sia utile perche'⌝
      An:::d                 not know.1s how much is  useful because
      An:::d (2.8) I don't know how much is useful because


12      non ho    mai fatto altrimenti pero' loro ci mettevano:
        not have.1s ever made otherwise  but   they cl. put.3p
        I have never done it differently but they were puttin:g
```
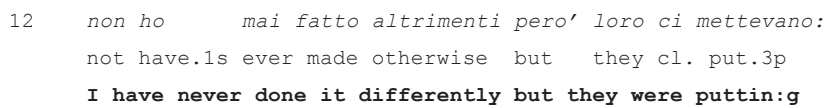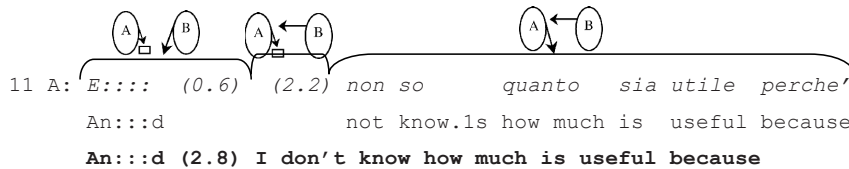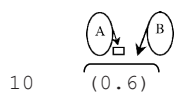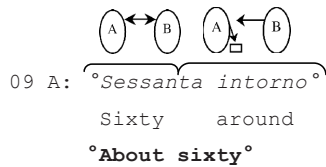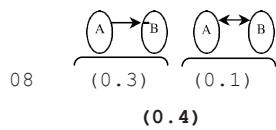
At lines 3-4 B starts an oblique sequence, asking how many people were present at the scout camp in which A and B worked together in the kitchen. B provides a candidate answer to his own question, and then at line 5, A answers. At the end of line 5 the sequence is possibly complete, although B continues looking at A and A looks back at him while adding "mi sembra" (it seems to me). On the last sound of "sembra" and while producing the acknowledgment token at line 7, B, the speaker of the sequence initiating turn, closes his eyes while A keeps looking at him.

Again, at the end of line 6 (and 7, given that they overlap) the sequence could be possibly complete but one participant, A, continues looking at the other. During the silence at line 8, B looks again at A, who adds, with soft voice, an alternative SPP to the one he produced at line 5. At the end of "sessanta" (sixty), A, the current speaker, looks down toward the menu he has on the table, while the recipient, B, looks away only at the completion of the turn. During the silence at line 10, both participants are looking away from each other (A toward the menu and B downward as he bites his nails). Subsequently, A

resumes his multi-unit turn starting with the conjunction "e" (and) and resetting the pitch to the level he had at the end of line 2 (see e.g. Couper-Kuhlen 2004 on the function of pitch reset). Two things should be noted about this example:

1) at possible completion of the sequence only one person sustains gaze, and, in this case, it is the same individual who adds a turn to the sequence (A);
2) even in the case of obliques, participants seem to orient to the possible completion of the sequence by withdrawing their gaze (in this case, B does it twice) only to look up again once the extended telling is resumed (see chapter 2).

Example 4.9 shows a case in which the speaker of the SPP continues looking at the other participant once possible sequence completion has been reached and the sequence gets expanded (gaze configuration 3). The participants are sitting facing each other and A, the host, has just sanctioned B for leaving possible stains on the tablecloth. At line 1, B provides an account in response to the sanctioning and line 3 is another sanction, this time for shaking the knife. The target is the sequence at lines 5-8 and in particular the gaze behavior at completion of the turn at line 6.

**(4.9) 2GC-paura 30:44**

```
01 B:  Appunto non te  la voglio  sporcare allora tiro    su cosi'
       Indeed  not you it want.1s stain     so     pick.1s up so
       Indeed I do not want to stain it so I pick it up like this
02     (0.3)
```
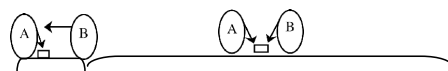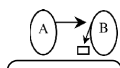


```
03 A:  Si'   no non agitare il  coltello
       Yes   no not shake    the knife
       Yes   no do not shake the knife
```



```
04     (0.3)   (0.6)
          (0.9)
```

261
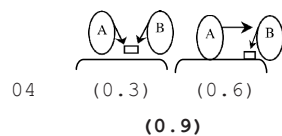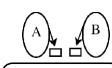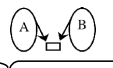
```
                    Ⓐ↔Ⓑ              Ⓐ→Ⓑ
                                        ▫
05 B:   Hai    p(h)aura che  te  lo [lanci]
        Have.2s fear      that you it  throw
      Are you af(h)raid that I will [throw] it to you

                                      Ⓐ→Ⓑ
                                         ▫
06 A:                               [hh s(h)i'] ((annuendo))
                                     hh yes
                                  [hh y(h)es] ((nodding))

            Ⓐ→Ⓑ
               ▫
07 B:   °Senti  sta s c[(h)e m]a°
        Listen this idiot
      °Listen to this id[(h)io]t°

              Ⓐ→Ⓑ
                 ▫
08 A:       [tzac] ((facendo gesto di tagliare))
             tzac
          [tzac] (( doing gesture of cutting))

       Ⓐ  Ⓑ
       ▫  ▫
09    (0.4)

       Ⓐ  Ⓑ            Ⓐ←Ⓑ                Ⓐ  Ⓑ
       ▫  ▫               ▫                ▫  ▫
10 A:  .hh ↑Beh  insomma e  quindi  dopo un po' mi annoio
            Well  so   and therefore after a bit me bore.1s
      .hh ↑Well  so and therefore after a bit I get bored
```

A is looking at the knife B is holding in her hand while she sanctions B at line 3. A keeps looking at the knife during the first part of the silence, and only looks up toward B, smiling, after 0.3 seconds. At line 5, B teases A and produces the question with a smile and a laugh token. A responds with laughter, a nod and "yes" (which deals with the question format of the previous turn). At this point, the sequence is possibly complete, but A, the speaker of the responsive turn, continues looking at B. B produces a negative assessment (joking) with soft smiley voice while A mimics the act of cutting with a knife and produces the sound at line 8.

At the completion of B's turn at line 7 A looks down and then looks at her right wrist (see line 9) until she resumes the topic of how much she does not like watching sports on TV, which was the topic of the conversation between A and B before A sanctioned B for leaving stains on the table cloth. The resumption of the prior topic is marked by a pitch reset which, like in example 4.8, marks the difference with the talk just concluded (Couper-Kuhlen 2004). In this example, we see again that at possible completion of a sequence (line 6), when the current speaker continues looking at the addressee, the sequence gets expanded. At the same time, the same participant (at this point the recipient) looks away by the time the sequence reaches next possible completion (line 7) and what follows is resumption of a previous course of action as opposed to more talk regarding the same sequence.

Example 4.10 shows what happens if only the recipient of the turn that brought the sequence to possible completion keeps looking at the other participant once possible completion is reached (gaze configuration 4): namely, the sequence gets expanded. Here, the participants resume talking about the exam that A has to take soon. In example 4.7, which occurred approximately 1 minute earlier, B had asked about the lecture that A has to attend in the afternoon. During the silence at line 1, both participants are oriented toward their dish and are eating. Lines 2 and 3 are produced while both participants maintain the orientation toward their food. Our target is the gaze behavior during the sequence that starts at line 3 and ends at line 14.

**(4.10) 2PLUNCH1-esame 5:22**

```
01     (4.5)

02 A:   No appunto quindi [e::]
        No indeed  so    an::d
        No   indeed   so [an::d]


03 B:             [L'es]ame e' scritto,
                   The exam is written
                  [The e]xam is written,
```

263
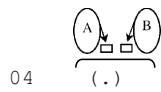
04          (.)
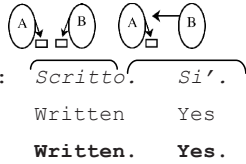
05 A:   *Scritto.    Si'.*
        Written    Yes
        **Written.    Yes.**
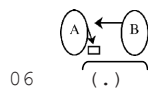
06          (.)
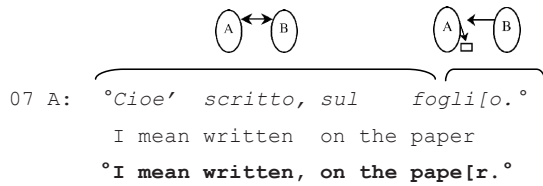
07 A:   *°Cioe'   scritto, sul    fogli[o.°*
         I mean written   on the paper
        **°I mean written, on the pape[r.°**
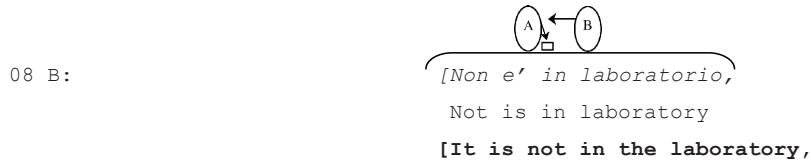
08 B:                        *[Non e' in laboratorio,*
                              Not is in laboratory
                        **[It is not in the laboratory,**

09          (0.5)
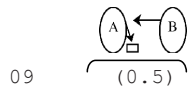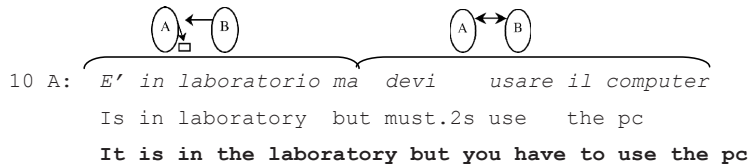
10 A:   *E' in laboratorio ma  devi    usare il computer*
        Is in laboratory  but must.2s use   the pc
        **It is in the laboratory but you have to use the pc**
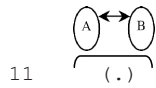
11          (.)
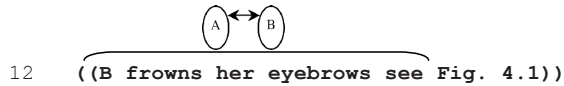
12      **((B frowns her eyebrows see Fig. 4.1))**

264

Figure 4.1. B frowning.

```
13 A:   Pero' utilizzi il computer  non per fare l'esame
        But   use.2s   the pc      not to  do    the exam
        But you use the pc not to do the exam

14      ma  per tirarti giu' per fare i   calcoli    sostanzialmente
        but to  take    down to  make the calculations essentially
        but to take down to make the calculations essentially

15      (11.0)

16 B:   hm, ieri     sera    hai     parlato con  il  prof:
        hm  yesterday evening have.2s talked  with the professor
        hm, yesterday evening did you talk to the professor:

17      (0.8) giovanissimo¿
              very young
        (0.8) very young¿   ((the very young professor))
```

At line 3 B interrupts A's turn and requests confirmation regarding the exam's format. This is a sequence initiating turn that A first confirms with a non type-conforming answer at line 5 (written), followed by a type-conforming one ("si'"). Focusing on the participants gaze behavior, we can see that B looks up toward A at the end of the word "scritto" (written) and keeps looking at B after the completion of the next TCU and into the following silence. Therefore, it is the recipient of the SPP at line 5 ("written yes") who continues looking at the other's face at possible completion of the sequence and engenders sequence expansion. First, we see that A also looks up and adds a TCU at line 7 ("I mean written on the paper") that

clarifies his response. At completion of this TCU, B continues looking at A and starts a next turn at line 8 that continues asking about the exam format. Again, possible sequence completion at the end of line 7 is passed with the recipient of that turn continuing to look at the current speaker. The FPP of a post-expansion (line 8) produces a candidate understanding about the exam not being in the laboratory, and at line 10 disconfirms it, indicating that it will be in the laboratory. B continues looking at A during the gap at line 9 and during line 10, where A looks up toward B as well so that they engage in mutual gaze. This means that at completion of line 10, which is yet another possible sequence completion point, the participants continue looking at each other. As we can see, the sequence is expanded yet again. At line 12, B furrows her eyebrows producing a facial expression comparable to a verbal repair initiation (by displaying a problem in understanding it seems to invite further clarification). At lines 13-14 A clarifies his previous utterance and again brings the sequence to possible completion. The last participant who looks away is the current speaker, A (who was the recipient of the sequence initiating turn), and he does so on the word "calcoli" (calculation) which represents the first projectable sequence completion in the turn. The following word, indeed, mitigates the idea that during the exam the only thing computer's could be used for is "to make the calculations" by saying "essentially". So we see that by the time the participants reach the next possible sequence completion both of them have withdrawn their gaze and what follows is a long lapse during which participants re-engage eating. At line 16, B starts a completely new sequence and course of action and asks A about what he did the previous evening at the end of a lesson he attended.

To sum up, in this example there are multiple possible completion points for the sequence: at line 5; at line 7; at line 10; and at line 14. However the sequence is only actually completed at the end of the turn at line 14. In terms of gaze behavior, we can see that while at line 5 and 7 B kept looking at A, and at the end of line 10 both participants kept looking at the other, at line 14 both participants have withdrawn their gaze before reaching actual completion and what follows is a lapse and then a new course of action.

In this section I have outlined how the gaze behavior of one participant treats the sequence as not yet complete and usually engenders sequence expansion. By the time both participants withdraw their gaze approaching or at the next point of possible completion of the sequence, the sequence is treated as complete.

266

## 4.5 Gaze Behavior's Relationship To Courses of Action

Example 4.10 showed us something very important: in terms of gaze behavior, participants treat the completion of a sequence in a "sequence of sequences" (Schegloff, 2007b) or of the first sequence of a topic proffering as a sequence requiring expansion. Indeed, in such cases participants continue looking at each other, therefore orienting toward the relevance of producing additional related talk. If this holds for other cases, it might tell us something about the similarities between a course of action realized through a single sequence and a course of action realized through multiple sequences. In this section, I further investigate this issue.

The sequential organization of talk allows the interactive accomplishment of action in interaction. We know that a base adjacency pair develops when participants attempt to perform some action and we know that such a pair can be repeatedly expanded and still understood as one course of action. However, Schegloff (2007b: 195) stresses the following point:

> *Although successive sequences may often be less closely linked than successive turns are, there can be particular ties of relevance between several sequences that serve to extend our sense of coherence and organizational relatedness of a stretch of talk beyond the boundaries of a single base sequence and its expansions.*

These successive sequences can be deployed to perform a course of action that requires the accomplishment of many specific steps before reaching actual completion.

In the same way in which I argued against the idea of gaze being simply related to the organization of turn-taking, I would like here to argue against the idea of gaze as simply organized around sequences. Instead, I suggest that what drives participant's gaze behavior is their understanding and recognition of *action* in interaction. Given that each sequence can be used to accomplish action in interaction, a completion of a sequence often ends up also constituting the completion of the course of action the participants were involved in. Notice that the completion of a sequence also constitutes the possible completion of a turn, the completion of a TCU, the completion of a word, the completion of a syllable, etc. Of all the things that reach completion I will focus here on places where a difference between completion of a single sequence and completion of a course of action can be observed. As mentioned in § 4.4, I am particularly concerned with the completion of "episodes of courses

of action", which can be distinguished by their contiguity in the conversation. As Schegloff (2007b) pointed out, there are stretches of talk in which the completion of a sequence does not represent the completion of the larger recognizable course of action. If we want to be able to distinguish between sequences and courses of action in terms of what is driving the gaze withdrawal of the participants, then we should analyze what happens to gaze in these sequences of sequences.

In what follows, three types of evidence will be provided in support of the claim that gaze is organized by reference to the courses of action participants produce and not just to sequences of talk:

1) Participants do not treat the completion of each sequence as a point of actual completion if it is a part of a recognizable larger course of action. Instead, participants sustain an orientation toward the other participant's face that displays an understanding that the course of action is still ongoing. Once the course of action reaches possible completion, both participants withdraw from looking at each other and the course of action is normally treated as complete.

2) If one participant looks at the other while trying to accomplish a course of action, once this course of action is clearly no longer accomplishable or is neglected by the other participant, the first participant withdraws her/his gaze, therefore orienting toward the course of action and not necessarily to sequential structures.

3) If a participant plans to shift to a new course of action s/he normally looks away first, marking the possible end of a prior course of action before starting a new one.

Examples 4.11-4.13 provide evidence for the first point, examples 4.14 and 4.15 provide evidence for the second point and example 4.16 provides evidence for the third point.

## 4.5.1 Completion of Sequence vs. Completion of Course of Action

In considering the first point, the following three examples show that participants continue looking at each other during the development of multiple sequences that constitute one course of action, and they withdraw their gaze once the course of action is recognizably complete. Before the beginning of example 4.11, A is giving B advice about what to cook for a boyscout camp. The instruction-giving is momentarily interrupted by A, who starts a side

course of action at line 2. The focus is the participants' gaze behavior at the end of lines 3, 6, 9, 11 and 12.

**(4.11) 2PPLAN-conosciuto 13:05**

```
01     (0.4)
```

```
02 A:  uhm  dov'e'   che  ci siamo conosciuti [noi a]
       uhm  where is that cl are  met            we   at
       uhm  where is  it  that  we first met [at]
```

```
03 B:                                  [A Sant']Eustazio
                                         At Sant'Eustazio
                                        [At Sant']Eustazio
```

```
04 A:  Tra-  Sant'Eutizio ((annuendo)) ti ricordi     che  ci
       Between Sant'Eustazio            you remember.2s that cl
       Betwe- Sant'Eutizio ((nod)) do you remember that
```

```
05     facevan mangiare pomodoro con  la  cipolla?
       make.3p eat       tomato    with the onion
       they were making us eat tomato with onion?
```

```
06 B:  Si'
       Yes
       Yes
```

```
07     (.)
```

269

08 A: *E   col  prezzemolo,*
And with parsley,
**And with parsley,**

09 B: *Si' ((annuendo))*
Yes
**Yes ((nod))**

10 A: *Ottimo.   Quindi se avete   della cipolla in piu' mettetela*
Excellent. So    if have.2p some onion    in more put.2s it
**Excellent. So if you have some onions left put it**

11   *insieme  ai  [pomodo]ri. E' molto buona*
together with tomatoes   is very  good
**together with [tomato]es. It is very good**

12 B: *((annuendo)) [°Si'°]*
Yes
**((nods))    [°Yes°]**

13     (1.1)

14 A: *.hh Allora riso al   pomodoro .hhh e' il risotto.*
Then   rice with tomato      is the risotto
**.hh Then rice with tomato ((sauce)) .hhh it is the risotto.**

A's question at line 2 is answered at line 3, but at possible completion of the sequence, B continues looking at A. A receives the piece of information (the place where they first met; line 4) and then moves on to a next sequence asking B whether he remembers what the chef

270
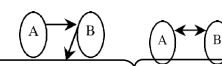
of that boyscout camp served them. The first request for recognition is answered at line 6 and the second one at line 9. The two participants sustain mutual gaze throughout the silences and during all of these turns. Then at line 10, A assesses the fact that B remembers what they were eating at the camp where they met and this allows A to move on with the planned course of action, which is to provide B with some additional advice on a recipe. The advice at line 11 is accepted at line 12 and A adds a strong positive assessment after the acceptance (it is very good). After accepting the advice, A looks away, while B keeps looking at A until his strong assessment is complete. At this point, he then also looks away. After some silence (line 13), A resumes the previous course of action (i.e., reading the instructions on the recipe in front of him) after having completed the current one (i.e., providing advice about how to improve a recipe).

While we may be tempted to consider this entire stretch of talk a series of pre-expansions followed by the base sequence at lines 10-11, I would rather suggest that this is a sequence of sequences that accomplishes one side course of action. This side course of action is certainly relevant for the larger activity of A's instructing B, but that activity reaches completion at the end of lines 11-12. If we consider the general organization of action of this example we can represent it in this way: participants are involved in the activity of giving and receiving instructions for preparing specific recipes. Within this larger framework, each recipe can be taken as a step in the activity and can potentially be treated as a single course of action, or an episode of the larger course of action. In particular, A moves out of the details of one recipe and provides general advice on how to prepare specific ingredients. As such, there are multiple sequences of talk that accomplish very specific actions; actions that are functional to the accomplishment of all of the above levels.
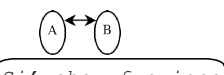
Taking a closer look at this example in terms of action, we can see that A is providing some general advice to B, but in order to do this he has to make sure that B is capable of recognizing the quality and taste of what he is suggesting. There is no clear logical way of establishing that a question about where the two participants first met could project that onion with tomatoes is tasty. However, by proceeding step by step we can see how A displays the link between these two items and their relevance for B. By doing this, he constructs one course of action performed through multiple sequences. Thus, this example shows that if participants can recognize a sequence initiating action as a possible first move in a larger course of action, they can sustain their orientations toward the other participant until that course of action is complete, independently of the ending and starting of other sequences along the way.
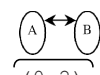
Example 4.12 is similar to 4.11 and shows an even more elaborate course of action performed through multiple sequences, which ends with gaze withdrawal by both participants. Before the beginning of this extract, A has mentioned that he is planning to bring his scooter (il Ciao) from his hometown to the university town where he now lives. B had previously mentioned possible problems regarding newer gasoline not being appropriate for the old scooter. A rejected this suggestion, justifying the scooter's safety by citing his grandfather (i.e., if his grandfather gave it to him then it must be working). They change topic for a while until B returns to it at line 29.

**(4.12) 2PLUNCH1- Il Ciao 13:53**
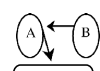


```
29 B:   Il   Ciao funziona?
        The Ciao work.3s
        The Ciao works?
```



```
30        (0.5)
```



```
31 A:   Si' che  funziona
        Yes that works
        Yes it works
```



```
32        (0.2)
```



```
33 B:   Tuo nonno        lo usa¿
        Your grandfather it use.3s
        Your grandfather uses it¿
```



```
34        (0.2)
```

35 A:   *Sempre usato*
       Always used
       **Always used**

36 B:   *Non l' ho    mai vis[to   tuo  nonno   girare] col Ciao=*
       Not it have.2s ever seen   your grandpa ride    with Ciao
       **I have never se[en your granfather riding] with the Ciao=**

37 A:                   *[Chiaro    che non  fa    i duecento]*
                         Clear       that not do.3s the two-hundred
                   **[Clearly it doesn't get to two hundred]**

38 A:   *=Non l'  hai    mai  visto?*
       Not him have.2s ever seen
       **=You have never seen him?**

39   (0.5)

40 B:   *No hha*
       No hha
       **No hha**

41 A:   *Ci    credo     che  non l'  hai     mai visto. (.)*
       Cl    believe.1s that not him have.2s ever seen
       **Of   course you have never seen him. (.)**
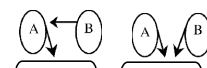
42    *Perche' non lo usa.=*
      Because not it   uses
      **Because he doesn't use it.=**
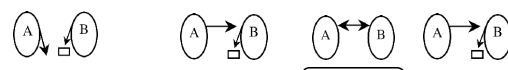

43    *=Ah e(h)cco*
      Oh t(h)here it is
      **=Oh the(h)re it is**


44    (1.3)


45 B: *La  tua  nonna   gi(h)rare col C(h)iao*
      The  your grandma ride       with Ciao
      **Your grandmother ri(h)ding with the C(h)iao**


46    *non l' ho     mai vi(h)sta*
      not her have.1s ever seen
      **I have never s(h)een her ((either))**


47    (2.0)   (3.0)   ((A shakes his head))
      **(5.0)**


47 B: *Io il casco te  lo presto (0.8) e   poi::*
      I the helmet you it lend.1s      and then
      **I will lend you the helmet (0.8) and the::n**


48    *prima o poi te  lo compri*
      first or then you it buy.2s
      **sooner or later you buy it ((for yourself))**


274

At line 29, B again challenges A's claim that the scooter is working. This is meant as a tease, because previously B had criticized A for indicating a desire to bring the scooter to town and then coming up with multiple excuses for not doing so later on. In this context, asking whether the scooter is working is meant as a tease, to pre-empt a possible future excuse for not bringing the scooter to town: it is not working. The answer at line 31 displays his stance toward the tease and constitutes a po-faced response (Drew, 1987) to the tease. That is, he not only produces an affirmative answer but he also adds "che funziona" (it works) indicating that this is something already discussed. At this point both participants keep looking at each other and neither of them backs down. B escalates the challenge/tease by asking whether A's grandfather uses the scooter (and therefore doubting the fact that the scooter actually works). Again, A answers positively (sustaining the challenge), but B continues doubting that his grandfather does use the scooter (line 36). Lines 38 and 40 are an inserted repair sequence produced with sustained mutual gaze, and at lines 41-42, A responds to B's challenge/tease confirming that his grandfather indeed does not use it anymore (notice that A did say that his grandfather has always used the scooter, not that he is using it at the moment).

The end of line 42 marks a point of possible sequence completion but the participants sustain mutual gaze. B produces a sequence closing third that teases A by claiming that he confirmed what she already knew: that the scooter does not actually work. However, she soon realizes that the only thing ruled out is that A's grandfather is using the scooter, but not that someone else is using it and so she adds a teasing next turn where she says that she has not seen A's grandmother using the scooter either. B keeps looking at him while uttering line 45-46. However, A has already re-oriented his gaze away from her and reacts to the tease by shaking his head but not responding verbally (which could lead the course of action to be further expanded). After 2 seconds of silence B looks away and begins eating (line 47). Then she starts a new sequence about her willingness to lend him her helmet. At this point, the course of action about the scooter and the fact that it does not work is complete and this new talk about the helmet, while being topically related, constitutes a new course of action.

In this example, just as in example 4.11, we have seen that participants sustain their gaze toward each other across multiple sequences and yet they look away at a certain point, when the larger course of action is possibly complete (in this case, challenging that the scooter works).

Finally, example 4.13 provides further evidence of the relationship between gaze withdrawal from the other participant and course of action, rather than sequence, completion.

275

In this example, B is visiting A, and, at the very beginning of their interaction, B appears to seek an invitation from A to go out with her and her friends. Our focus is the gaze behavior at each sequence possible completion: lines 2, 4, 8, 10, 13, 14, 17, 18.

**(4.13) 2GGOSS-stasera 00:15**

```
01 B:   .hh A(h)l l o r a stas(h)era cos' e' che  fate
             So              tonight     what is that do.2s
        .hh S(h)o t(h)onight what are you doing


02 A:   Eh     andiamo a  Villa Chiara=
        Eh     go.1p    to Villa Chiara
        Eh     we go to Villa Chiara=


03 B:   =Ma   a   che   ora  vi  incontrate
         But at which hour you meet.2p
        =But at what time do you meet


04 A:   Vado  alle nove e   mezza dalla Gloria.
        Go.1s at   nine and half  to    Gloria
        I go to Gloria's (house) at nine thirty.


05      ((B annuisce))
        ((B nods))


06   ((some turns with side sequence omitted))
```

276

07 B:        [*Nove  e   mezza ma] andate  subito      a  Villa*
             Nine  and half  but  go.2p  immediately to Villa
             **[Nine thirty but] do you go immediately to Villa**

08           *Chiara alle  nove e    mezza?=*
             Chiara at    nine and half
             **Chiara at nine thirty?=**

09 A:        *=Con- le   dieci con  gli altri.*
              With the ten    with the others
             **=With- at ten with the others.**

10           (0.4)

11 A:        *Ci  incontriamo.*
             Cl. meet.1p
             **We meet.**

12           (0.6)

13 A:        *Ci [v i e n i?*
             Cl. come.2s
             **Will [you come? ((to Villa Chiara))**

14 B:   [Io    esco        alle nove_ (0.5) ah io pero' se arrivo
        I      get.1s out at   nine        oh I   but  if arrive.1s
        **[I      get out (of work) at nine_ (0.5) oh but if I come**

15      arrivo    a mezzanotte eh
        arrive.1s at midnight eh
        **I arrive at midnight eh**

16      (0.6) ((A displays shocked facial expression see Fig. 4.2))



Figure 4.2. A's facial expression during line 16.

17 A:   Alle nove esci?
        At   nine get.2s out
        **At nine you get out (of work)?**

18 B:   E[h
        Eh
        **Y[es ((confirming something already said))**

19 A:   [Cazzo ((Poi A annuisce, guardando in basso))
        Dick
        **[Shit ((and then A nods, looking down))**

```
20     (0.6)        (1.2)
         (1.8)



21 A:  Domani  sera   che  fai?
       Tomorrow evening what do.2s
       Tomorrow evening what do you do?



22      (0.5)
```
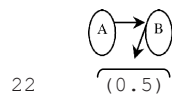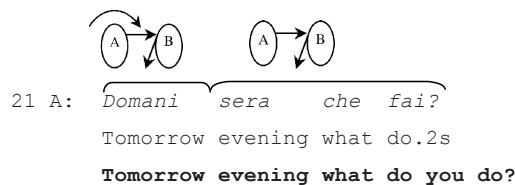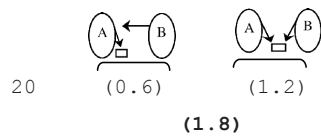
At line 1, the recipient starts looking at the speaker before the speaker turns toward her and the participants engage in mutual gaze. At the completion of line 2, both participants continue looking at each other and the same happens at the end of lines 4 and 8 and 10. These all constitute possible completions of a sequence for the three adjacency pair sequences that begin with the questions at line 1, 3 and 6-7. Each of the 3 questions can be analyzed as a request for information by B about A's plans for the night, probably meant to get an invitation by A. During the silence at line 11, A, the person who has been asked about her plans, continues looking at B and at line 12 resolves to invite B to go out with her and her friends. B's SPP at lines 13-14 is an indirect answer to A's invitation at line 12. B is providing her schedule indicating when she might be joining them, and again, both at completion of the first TCU of line 13 (I get out at nine) and at the end of line 14, the participants sustain mutual gaze. They sustain mutual gaze during the following silence and into the next adjacency pair sequence (a repair sequence; lines 16-18). A, the person who produced the actual invitation, looks away while producing the negative assessment at line 18, which displays the unlikelihood that A and B will meet that night. B continues looking at A briefly into the silence at line 19 and then looks away as well. At this point, they withdraw their gaze for 1.2 seconds until A resumes asking about B's plans for the following evening. Although topically related to the previous course of action and clearly occasioned by the failed invitation, A's turn at line 20 starts a different episode of the course of action as it is asking about B's plans for the following day. Furthermore, as observable in both participants'

279

behavior, A's turn at line 20 was not projectable at line 18 and during the silence at line 19 both participants act as if the course of action is complete.

In this example, then, we again see that participants sustain mutual gaze across multiple sequences that constitute a larger course of action and withdraw from looking at the other participant only when the course of action appears to be complete (at line 19), when at least one participant has produced a bid for closure (in this case, when A withdrew her gaze while producing line 18). Together with the two other examples in this section, these three examples show that participants orient toward the completion of sequences that are part of a larger course of action in a similar way as they do toward the completion of an inserted repair sequence in a base sequence. This confirms that participants' gaze orients toward the accomplishment of the projectable course of action more than toward the completion of a single sequence.

## 4.5.2 Is the Course of Action Still Accomplishable?

The second piece of evidence that participants' gaze behavior displays an orientation toward the completion of a course action consists of showing how participants orient toward the failure of a projectable course of action, independent of whether the sequence is actually complete. Examples 4.14 and 4.15 are instances of this kind of orientation.

In example 4.14, B is visiting A to chat and eat dessert. The excerpt starts with the participants discussing whether they should finish the cake on the table in front of them. This leads to a brief discussion of what they each had for dinner. Our focus is the course of action that starts at line 6.

**(4.14) 2GC-pizza 17:10**

```
01 A:  Ah dici   ce la magnamo tutta=
       Oh say.2s cl it eat.1p  all
       Oh you say we will eat it all=

02 B:  =No oh che tutta L(h)uisa. Con calma [che  io sono obe:sa
        No oh what all  Luisa     With calm  that I  am   obese
       =No oh what all L(h)uisa. Calm down [that I am obe:se
```

03 A:
    [Anche perche' io
    *Also   because I*
    **[Also because I**

04
*mi sono mangiata una pizza intera a  cena.*
me has eaten    a   pizza whole at dinner
**ate a whole pizza for dinner**.

05
    (1.0)    (0.2)
    **(1.2)**

06 B:
*Ma   dove  hai    mangiato? Qui?*
But where have.2s eaten     Here
**But where did you eat? Here?**

07
    (0.4)

08 A:
*Si'.*
Yes
**Yes.**

09 B:
*E   hai    mangiato da sola?*
And have.2s eaten     alone
**And did you eat alone?**

10 A:
*No.  Con  la  mamma.*
No.  With the mum
**No.  With mum.**

281

```
          A      B
         ( )    ( )
          ↓ ↘ ↙
           □
11 B: ⌐ Ah::::
      Oh

      Oh::::


          A      B
         ( )    ( )
          ↓ ↘ ↙
           □
12      (0.2)


                   A      B
                  ( )    ( )
                   ↓ ↘ ↙
                    □
13 B: Perche' non me lo hai    detto. Potevo   venire.
      Because not me it have.2s said   Could.1s come

      Why didn't you tell me. I could have come.
```

At lines 3-4, A announces that she ate a whole pizza for dinner, affiliating with B's stance against eating the whole cake. A is trying to cut the cake and as soon as she looks up at B (the end of silence at line 5), B asks a question (line 6). This FPP is a wh-question that gets turned into a yes-no question by adding a candidate answer. After a short silence, A responds with a type-conforming answer, at which point the sequence could be over. But the participants are still looking at each other. And what follows is a new sequence that starts with another yes-no question (line 9); however, it is also understandable as part of the same course of action. This time, A looks at the cake before answering at line 10, while B looks down at the cake at the beginning of the second TCU of that turn ("with mum"). Thus, the first aversion of gaze occurs as soon as B hears the first TCU of that turn ("no"), but the course of action is not over despite her averted gaze. At this point, B produces a minimal post-expansion that claims information receipt (Heritage, 1984), while both participants are looking down (see § 4.7.1). Finally, B makes the interactional plan of the two previous questions explicit: that is, she complains about not having been invited for dinner (line 13). Indeed, B has only been invited for dessert whereas it is quite common and ordinary to be invited for both dinner and dessert. In this sense, only being invited for dessert is a noticeable, and perhaps accountable, action.

As we have seen in example 4.11, it is not completely clear what kind of action B is projecting. That is, at the end of the question at line 6, A does not know that B is going to complain about the fact that she was not invited for dinner, though it is possible to recognize the link between the two events. At lines 6 and 9, B checks whether A could have actually invited her for dinner. While A's answer to the question at line 6 allows B to continue this

course of action, the first TCU at line 10 ("no") blocks it. By saying that she did not have dinner alone, A blocks the possible complaint by B. At the same time, the fact that she had dinner with her mum, whom B knows quite well, would make it less problematic in terms of making explicit her disappointment for not having been invited for the pizza. The production of the complaint at line 13 could be a *post hoc* adjustment to the new information ("with mum"), rather than a failure of gaze withdrawal as a way of orienting toward closure of the course of action. By looking away after the "no" at line 10, B treats that response as blocking and the course of action as potentially over. A subsequently lowers her gaze, but by adding the TCU about her mum, she leaves open the possibility of further talk by B, which indeed occurs. The main point here is that the participants sustain mutual gaze across two different sequences of talk that are both part of the same projected course of action. A could continue if the answers to B's question were always preferred. However, this is not the case after the second question and we can see a clear orientation to this by the person who initiated the course of action, who withdraws her gaze.

Example 4.15 shows something similar, although there is no sustained gaze across multiple sequences of talk. In the following excerpt, B is visiting A to fix his computer. At line 10, B mentions what he is going to do. Our focus is the quip at line 17, where A plays with the double meaning of the word "importa": to import and to care.
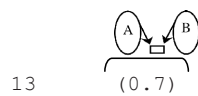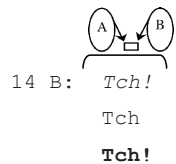
**(4.15) 2PCOMP-battuta 06:32**

```
                       ⎛A⎞↘↙⎛B⎞
                        └□┘
10 B:  ⌜↑Ora (1.2) io dovrei cercare di- file importa.⌝
        Now       I should  try     to  file import.
       ↑Now (1.2) I should try to- file import.
```

```
                   ⎛A⎞↘↙⎛B⎞
                     └□┘
11      ⌜(1.4)⌝
```
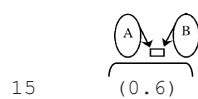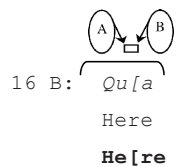
```
                 ⎛A⎞↘↙⎛B⎞
                   └□┘
12 B:  ⌜°uhh
         uhh
       °uhh
```

13　　　(0.7)



14 B:　*Tch!*
　　　Tch
　　　**Tch!**
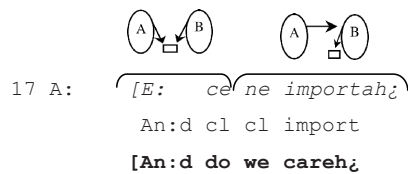

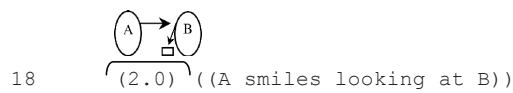
15　　　(0.6)



16 B:　*Qu[a*
　　　Here
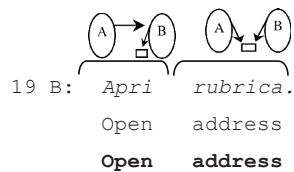　　　**He[re**



17 A:　*[E:　ce ne importah¿*
　　　An:d cl cl import
　　　**[An:d do we careh¿**
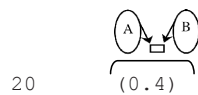


18　　　(2.0) ((A smiles looking at B))



19 B:　*Apri　rubrica.*
　　　Open　address book
　　　**Open　address book.**



20　　　(0.4)

At line 17, A produces a quip while looking toward B, who continues looking at the computer screen. Line 17 ("and do we care?") is hearable as either asserting a position

(Schegloff, 1984) or a quip, so either agreement or laughter are appropriate responses to it. And yet, neither occurs. After 2 seconds of silence, during which A sustains his gaze toward B, B indicates his engagement with fixing the computer problems, thereby indicating that he is not going to react to A's "and do we care?" (line 17). Interestingly, A looks away and re-orients toward the screen after the first word of line 19 ("open").

To briefly summarize, both examples 4.14 and 4.15 have shown that as soon as one participant blocks a course of action, the eyes of the other participant move away from them and, as a result, the blocked participant re-engages in competing activities. This displays a clear sensitivity toward the accomplishment of a course of action rather than simply the completeness of a sequence.

### 4.5.3 Looking Away Before Starting a New Course of Action

The third and final piece of evidence of the importance of courses of action and action completion for participants gaze behavior can be observed in example 4.16 (and later in example 4.25). If the claim is that participants withdraw their gaze to display an orientation toward the completeness of a course of action rather than a sequence, then we should see that before new courses of action participants look away from each other. This, however, need not be performed by both participants, as they could have different understandings of the actual completion of the course of action. Nonetheless, if the function of the withdrawal is systematic, then even if there are unilateral departures and one participant starts a new course of action while interrupting another one, this person should look away before actually starting the new course of action, therefore marking the action-in-progress as over and allowing for the possibility that something new can be produced. This is illustrated in example 4.16.

In this excerpt, the participants have been talking about possible summer plans, and B has complained about the fact that she will be alone for a few days near the end of August because her boyfriend is going away with friends. While producing a turn that suggests A is treating the course of action as complete ("yes, I understood"; line 3), at line 7 A looks away immediately before starting a completely different topic and course of action. Our focus is what happens at line 7.
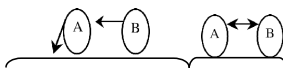
**(4.16) 2GGOSS-insomma 05:25**

01 B:  *volevo  far qualcosa  tipo mm non so      anche a  m- m-*
       want.1s do   something like mm not know.1s also  to m- m-
       **I wanted to do something like mm I don't know also to m- m-**
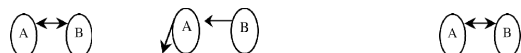
02     *me va     ben[issimo magari Rave]nna delle robe   cosi'.*
       me go.3s very well   maybe  Ravenna  some   things so
       **for me it is tot[ally fine maybe Ravenna or things like that.**

03 A:              *[Si' ho      capito] ((annuendo))*
                    Yes have.1s understood
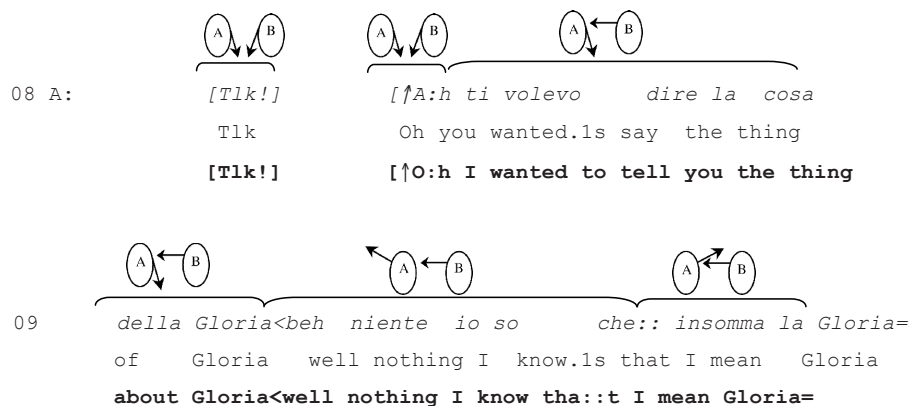                   **[Yes I understood ] ((nodding))**

04         (0.3)

05 B:  *Pero' °insomma°.*
       But    so
       **Well °so°.**

06     (0.2)

07 B:  *No guard[a mi] e' disp[iaciuto*
       No look    me  is sad
       **No loo[k I]  was   sa[d**

08 A:        *[Tlk!]*            *[↑A:h ti volevo    dire la  cosa*
             Tlk                 Oh  you wanted.1s say  the thing
             **[Tlk!]**          **[↑O:h I wanted to tell you the thing**

09          *della Gloria<beh  niente  io so      che:: insomma la Gloria=*
            of    Gloria  well nothing I  know.1s that  I mean   Gloria
            **about Gloria<well nothing I know tha::t I mean Gloria=**

At line 3, A acknowledges B's previous turn and explicitly states that she understands what B means by it. Nonetheless she continues looking at B, who produces a *postmortem* (Schegloff, 2007b: 142) turn at line 5. At line 7, B produces another turn, apparently pursuing further affiliation with her unfortunate circumstances, but at this point A has already looked away. In overlap with B's affiliation-seeking turn, she starts a completely new course of action. She projects this through prosodic, lexical and visual means: first, with a pitch reset at the beginning of this new TCU (Couper-Kuhlen 2004); second, by using a verbal preface that marks this new talk as distinct from what came before ("I wanted to tell you about"); and third, by looking away immediately before starting this new TCU. What follows is a long discussion about the attitude of a friend of theirs, a clearly different topic than the one of their previous conversation. This example shows that the normative association between looking away when the course of action can be considered complete or blocked can also be found in the behavior of a single participant, who performs it without necessarily waiting for the other's alignment.

Examples 4.11 to 4.16 both contribute evidence that participants orient toward the completion of courses of action more than toward the completion of a sequence of talk, although most of the time the completion of one constitutes the completion of the other. The next section provides quantitative evidence that further supports the claims presented in this chapter.

## 4.6 A Systematic Practice

The previous sections provided qualitative evidence for two claims about the organization of gaze behavior:

1) If a possible sequence completion is treated by participants as the actual completion of the course of action, they will not be looking at each other after the completion of the sequence;

2) If participants sustain their gaze toward the other once they reach possible sequence completion, this engenders sequence expansion.

This section provides further evidence of the regularity of the patterns exemplified thus far. The following results are based on having coded gaze behavior and whether a sequence/course of action was possibly complete in 10 minutes of Italian dyadic interactions for each of the 10 different interactions[11] (total of 100 minutes). The data were also coded in terms of whether the sequence/course of action was actually expanded or not. Participant gaze behavior was coded by considering where the participants were looking 0.2 seconds after the completion of the turn that constitutes possible completion of a sequence. Time was not measured in terms of speed of current speech (that is to say, using a participant perspective) but rather using software (ELAN, see chapter 1) that measures the time that passes between the completion of a turn and the beginning of another. For this reason, I added a 0.2 window to the actual completion of the turn, which approximately corresponds to what conversation analysts consider the amount of time that normally elapses during turn transitions and corresponds to the average turn latency described in Stivers et al. (2009).[12]

The criteria for coding sequences were based on the categories outlined by Schegloff in *Sequence Organization in Interaction* (2007b). As mentioned in § 4.4, I considered sequence expansion not only in terms of new turns produced by other participants, but also additional TCUs produced by the speaker who has brought the sequence to possible completion. Both of these situations were considered sequence expansion. Moreover, if the turn following a possible completion of a sequence was considered an expansion of the

---

[11] The interactional data coded for this chapter correspond to that used in chapters 2 and 3 and further described in Appendix A.

[12] See de Ruiter, Mitterer and Enfield (2006) and Stivers et al. (2009) for work on turn taking that uses "atomic clock" time rather than the more local participant's speaking speed. But see Jefferson (1988) for an argument against using "clock measured time".

current course of action, then this was coded as expansion. If two sequences were part of what Schegloff (2007b: 195-216) calls a "sequence of sequences", the second one would *not* be coded as "expansion" of the first one, since the course of action would only be considered complete at the completion of the second sequence. This follows from the claim outlined in § 4.5 that participants orient toward the possible completion of the course of action started with the sequence initiating action and not simply to the completion of each step, unless, of course, they can be considered different episodes of a course of action and each episode can be considered complete by itself. For example, if A asks B how difficult each of her next three exams will be, the course of action cannot be complete once B has provided an answer concerning one exam. It is only after B has provided information about all three exams that the course of action can be considered possibly complete. On the other hand, if A asks B what she did on Friday night, and, after the answer is provided A then asks about Saturday night, the completion of the sequence concerning Friday night does constitute the possible completion of that initial course of action. The subsequent question about Saturday night was not clearly foreseeable earlier. In other words, possible completion of a sequence is coded as such only when further talk by another participant is not projectable by the design of the sequence initiating action alone.

These coding measures yielded 469 instances of possible completions of sequences and simultaneously of courses of actions. Table 4.2 illustrates the frequency of sequence expansion following sequence possible completions.

*Table 4.2* Distribution of sequence expansions following possible sequence completion

| Sequence Expanded | Instances |
|---|---|
| Yes | 270 (57.6% ) |
| No | 199 (42.4%) |
| **Total** | 469 (100%) |

The information in this Table shows that just over half of possible completions of sequences are followed by sequential expansion. Thus, even though participants might orient toward specific interactional moments as making closure relevant next, it does not mean that the course of action is actually treated by both participants as complete, and further sequentially appropriate expansion might ensue.

*Table 4.3* Participant expanding the sequence when sequence gets expanded

| Who expands | Instances |
|---|---|
| Current Speaker | 89 (33%) |
| Other Participant | 181 (67%) |
| **Total** | 270 (100%) |

Table 4.3 shows that in two thirds of the sequences that get expanded, it is the person who was the recipient of the turn that constituted possible completion of the sequence who actually expands the sequence/course of action, and not the current speaker. Yet notice that in one third of cases, the speaker who brought a sequence to possible completion adds at least another TCU.

The first factor to consider in terms of what could affect sequence closure is the *sequential structure* of the actions immediately preceding possible completion of the sequence (see Schegloff, 2007b). Indeed, if we consider more precisely what kind of sequential structure had just reached possible completion, Table 4.4 shows that the data can be divided into three main categories.

*Table. 4.4.* Types of sequential structure culminating in a point of possible completion of the sequence/course of action.

| Sequential Structure | Acronym |
|---|---|
| Base Sequence | BS[13] |
| Post-Expansion of the Base Sequence | Post Ex. BS |
| Sequence Closing Third | SCT |

A base sequence (BS) is usually made of two adjacently paired turns produced by two different speakers. It constitutes the core of the course of action, and any pre-expansion or post-expansion has to do with completing the base sequence. If we think about an invitation, for example, we know that its delivery makes structurally and interactionally relevant the occurrence of an acceptance or rejection by the other participant. Now, the actual delivery of an invitation might be preceded by pre-sequences aimed at checking whether the person who is going to be invited might actually accept the invitation. Moreover, inserted sequences can

---

[13] BS is the structure that precedes possible completion of a sequence only if there are no post-expansions of it. So, BS and Post Ex. BS are mutually exclusive for this coding.

occur between the invitation and the occurrence of the acceptance, for example, to make sure that both participants understand the terms of the invitation. However, once an acceptance or rejection is produced, the base sequence reaches a point of possible completion and further expansions, by means of post-expansions of the base sequence (Post Ex. BS), are thought to occur in three circumstances:

1) when the SPP of the base sequence is interactionally dispreferred;
2) when there is some disagreement between the participants;
3) if something about the SPP is unclear  and requires further clarification.

However, both a base sequence and its post-expansions, which usually take the sequential form of adjacently pair turns, can also be followed by SCT turns. These SCT turns take the form of a single turn aimed at closing the sequence rather than expanding it. This does not mean, however, that the entire course of action necessarily achieves closure once a SCT turn is produced. Table 4.5 shows which sequential structures precede the possible completions of a sequence and Table 4.6 and Figure 4.3 show how often each one gets expanded.

*Table 4.5* Types of sequential structures that precede possible completion of a sequence.

| Preceding Sequential Structure | Instances |
|---|---|
| BS | 265 (56.5%) |
| Post Exp. BS | 141 (30.1%) |
| SCT | 63 (13.4%) |
| **Total** | 469 (100%) |

*Table 4.6* Distribution of sequence expansion after specific sequential structures

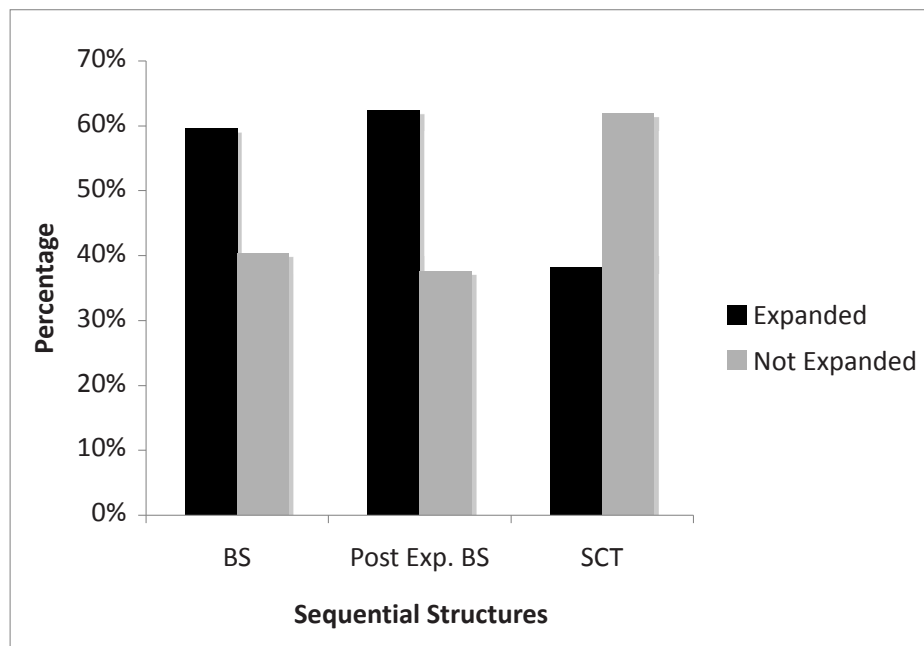| Preceding Sequential Structure | Expanded | Not Expanded | Total |
|---|---|---|---|
| BS | 158 (59.6%) | 107 (40.4%) | 265 (100%) |
| Post Exp. BS | 88 (62.4%) | 53 (37.6%) | 141 (100%) |
| SCT | 24 (38.1%) | 39 (61.9%) | 63 (100%) |

*Figure 4.3* Distribution of sequence expansion vs. non-expansion following specific sequential structures

Table 4.5 shows that just over half of the sequence possible completions in the data set occur after a BS and only 13.4% occur after a SCT turn. However, Table 4.6 and Figure 4.3 show that distinguishing between these different sequential structures might be relevant in terms of what is likely to follow next. Indeed, while sequence expansion is more likely than no expansion after a base sequence, it is much less likely after a SCT. Indeed, a SCT is significantly more likely to lead to sequence closure than a BS or a PEBS ($\chi^2$ (2)= 11.59; p<.01). These sequential structures occur while people are co-present and face-to-face, and although SCTs may make sequence closure more relevant, different gaze behaviors can still be implemented on top of these sequential structures. Previous research on telephone conversations suggests that SCTs are especially designed to make sequence closure relevant next by functioning as a bid for closure (see Schegloff, 2007b: 118-148). Hence, the sequential structure that precedes sequence completion is an additional variable that might affect the occurrence of sequence expansion.

A second variable is the *preference* of the turn that precedes possible completion of the sequence. One of the reasons a sequence may get expanded is that the last turn was

interactionally dispreferred[14] and engendered either an account by the current speaker (e.g., accounting for why s/he cannot accept an invitation to a party) or the production of further talk by the other participant (e.g., seeking accounts, negotiating disagreements). Table 4.7 shows how often the turn that preceded sequence possible completion is analyzed as a preferred one and Table 4.8 and Figure 4.4 show how often this is correlated with sequence expansion.

*Table 4.7* Preference of turn preceding sequence possible completion

| Preference of turn | Instances |
|---|---|
| Preferred | 297 (63.3%) |
| Dispreferred | 172 (36.7%) |
| **Total** | 469 (100%) |

*Table 4.8* Sequence expansion by preference of the turn preceding possible sequence completion

| Preference of turn | Sequence expanded | Sequence not expanded | Total |
|---|---|---|---|
| Preferred | 145 (48.8%) | 152 (51.2%) | 297 (100%) |
| Dispreferred | 125 (72.7%) | 47 (27.3%) | 172 (100%) |

---

[14] See chapter 3 for a more detailed explanation of preference and dispreference in terms of social action.
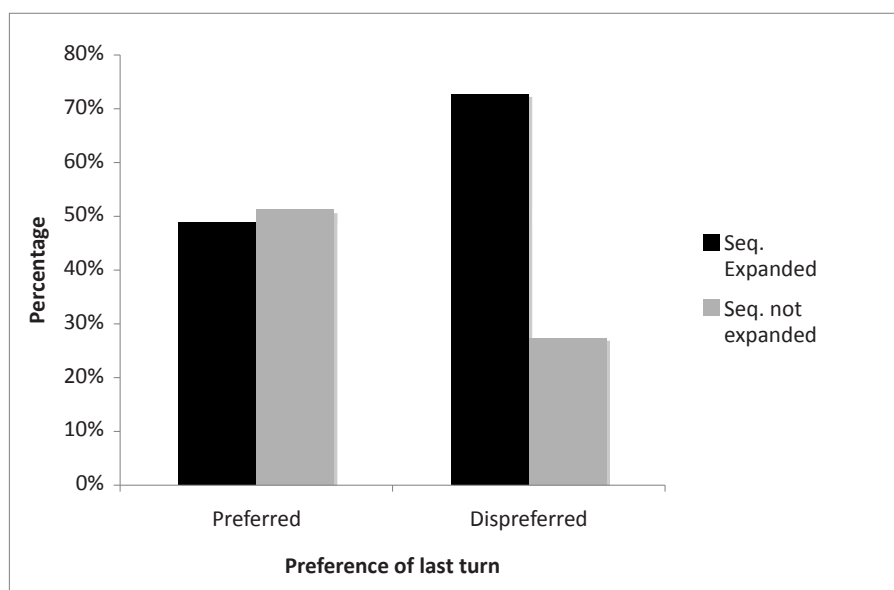
*Figure 4.4* Sequence expansion by preference of the turn preceding possible sequence completion

Table 4.7 shows that approximately two thirds of the turns leading to possible sequence completion are interactionally preferred responses to the preceding actions. These numbers reveal either a tendency by recipients of FPPs to "play along" and cooperate interactionally and produce actions that are structurally preferred, or the ability of speakers of FPPs to avoid producing FPPs that would yield a dispreferred response.

On the other hand, Table 4.8 and Figure 4.4 show that the preference of the turn preceding sequence completion affects the likelihood that any participant will expand the sequence. While expansion follows a dispreferred turn in 73% of the cases, it appears that a preferred turn does not necessarily lead to sequence closure, as this occurs only half of the time. A dispreferred turn, however, is significantly more likely to lead to sequence expansion than a preferred one ($\chi^2(1)= 25.37$; $p<.0001$).

Sections 4.4 and 4.5 documented the relevance of gaze behavior to sequence expansion. This means that *gaze* should be considered another variable affecting sequence expansion. Table 4.9 shows that in slightly more than half of possible completions of a sequence, at least one participant was still looking at the other.

*Table 4.9* Distribution of sequence possible completions with respect to gaze behavior

| Coding | Instances |
|---|---|
| Both participant gaze away from each other | 196 (41.8%) |
| At least one participant gazes at other | 273 (58.2%) |
| **Total** | 469 (100%) |

Considering the claims outlined in § 4.4 and 4.5, we should expect that if participants are not looking at each other at possible sequence completion then a sequence expansion is less likely to occur, while if they keep looking, then the sequence is more likely to be expanded. Tables 4.10 and 4.11 show that this is indeed the case. Figure 4.5 represents the two distributions graphically. In 83% of the cases in which both participants are looking away, the sequence is not expanded, and, on the other hand, in 87% of the cases in which at least one participant looks at the other, the sequence gets expanded.

*Table 4.10* Distribution of sequence expansions in relation to gaze down by both participants

| Coding | Instances |
|---|---|
| Gaze down → Sequence ends | 163 (83.2%) |
| Gaze down → Sequence continues | 33 (16.8%) |
| **Total** | 196 (100%) |

*Table 4.11* Distribution of sequence expansions in relation to gaze up by at least one participant

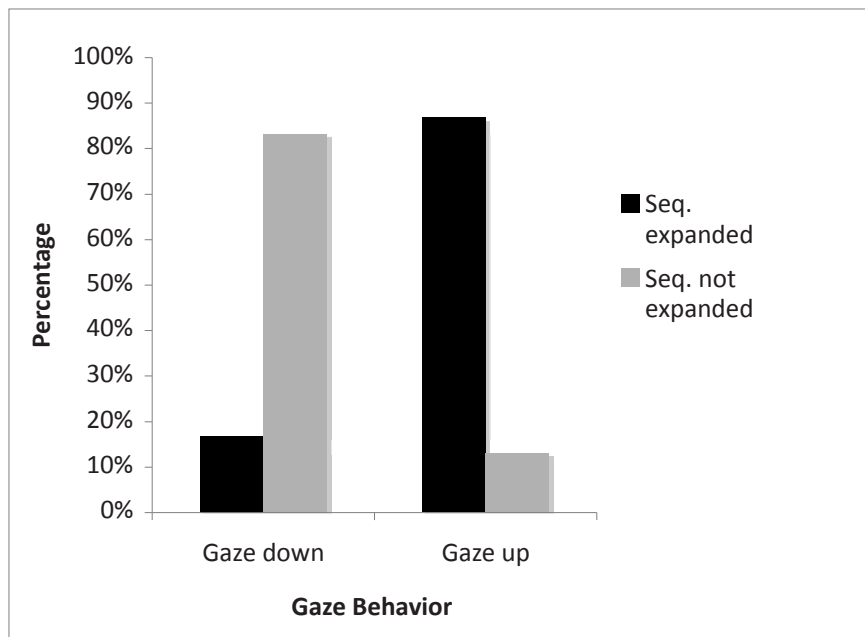| Coding | Instances |
|---|---|
| Gaze up → Sequence ends | 36 (13.2%) |
| Gaze up → Sequence continues | 237 (86.8%) |
| **Total** | 273 (100%) |

Figure 4.5 Distribution of sequence expansion in relation to participant's gaze behavior

To investigate whether participants' gaze predicted sequence expansion independently of the sequential position and preference of the turn, a logistic regression model was fitted to the data using the standard error corrected for the clustering of sequences in the 10 interactions. The results of the logistic regression are shown in Table 4.12 as an odds ratio with the 95% confidence interval.

*Table 4.12* Results of logistic regression predicting sequence expansion in relation to gaze toward addressee, sequential placement of the turn and preference of the turn.[15]

| **Variables** | **Odds Ratio** | **95% Confidence Interval** |
|---|---|---|
| Gaze to other participant | 31.02*** | 16.47, 58.41 |
| Sequential position | 0.39*** | 0.22, 0.69 |
| Preference of turn | 0.45** | 0.26, 0.78 |

Table 4.12 shows that if at least one participant keeps looking at the other at sequence possible completion, then the odds that the sequence will be expanded increase 31 times, in contrast with a situation in which both participants are looking away. The odds increase even

---

[15] *** denotes p<.001; ** denotes p<.01

if we take into account the sequential position and the preference of the turn, though much less than if we simply consider their gaze behavior. The results were adjusted for interactions so that one particular interaction would not bias the general result.

These results provide further quantitative evidence of the association between gaze behavior and sequence expansion: when participants look away, they treat the sequence as complete, and if they keep looking at each other, this routinely engenders sequence expansion. However, as outlined in Table 4.1 in § 4.4, there are not simply two possible gaze configurations when two participants reach sequence completion but in fact four. The following tables show a more fine-grained representation of the data using the four categories outlined in Table 4.1 in § 4.4 (participants not looking at each other, both participants looking at each other, only the speaker looking at the recipient or only the recipient looking at the speaker). Table 4.13 illustrates how many instances of each gaze configuration occurred in the 100 minutes of data analyzed. Table 4.14 shows how often each gaze configuration at possible sequence completion was followed by a sequence expansion. Figure 4.6 provides a graphical representation of the distribution presented in Table 4.14.

*Table 4.13* Distribution of the four gaze configurations at possible sequence completion points

| Gaze configuration | Instances |
|---|---|
| 1 | 196 (41.8%) |
| 2 | 121 (25.8%) |
| 3 | 78 (16.6%) |
| 4 | 74 (15.8%) |
| **Total** | 469 (100%) |

The information presented in this table demonstrates that both participants sustain their gaze toward each other's face in less than half of the cases in which at least one person keeps looking at the other (121/273 = 44.3%). Given that the participants did not always engage in mutual gaze (see Table 4.15 below), the distribution of gaze configurations 3 and 4 does not necessarily mean that in half the cases one participant treated the sequence as possibly complete while the other did not. On the other hand, the number of instances of gaze

configurations 3 and 4 indicates that these two configurations actually occur regularly and are not just atypical cases of the gaze configuration 2.

*Table 4.14* Distribution of gaze configurations followed by sequence expansion

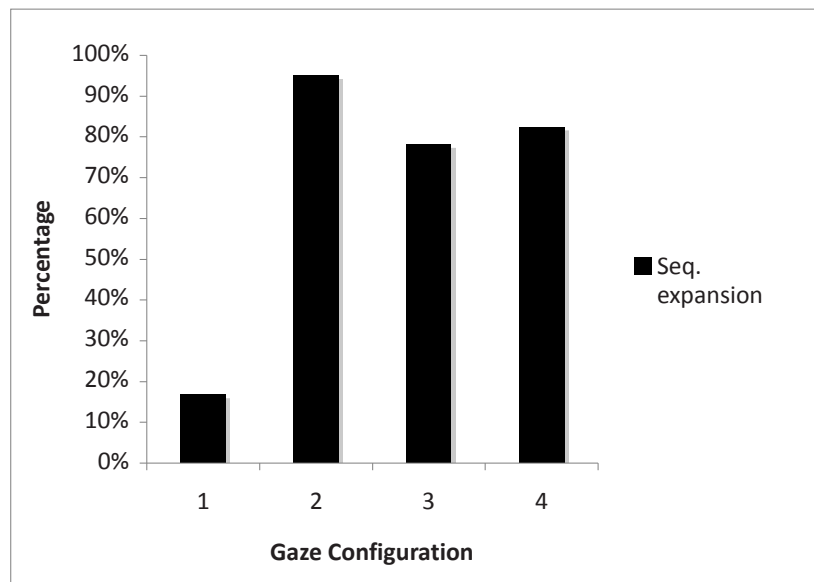| Gaze configuration | Instances of Sequence expansion |
|---|---|
| 1 | 33/196 (16.8%) |
| 2 | 115/121 (95.0%) |
| 3 | 61/78 (78.2%) |
| 4 | 61/74 (82.4%) |



*Figure 4.6* Distribution of gaze configurations followed by sequence expansion

The data shown in Table 4.14 (and Figure 4.6) indicates that although sequences tend to be expanded rarely after gaze configuration 1, they get regularly expanded after configurations 2, 3 and 4. However, the likelihood of a sequence expansion following configurations 3 and 4 is significantly less than after gaze configuration 2 ($\chi^2$ (2)= 13.87; p<.001). This is a reasonable finding given that the occurrence of gaze configuration 2 indicates that both participants are treating the current course of action as incomplete, while configurations 3 and 4 mainly occur when only one participant treats the course of action as incomplete. The above results show that configurations 2, 3 and 4 are alike in that in all

298

cases, sequence expansion is more likely than sequence closure; however, the probability of a sequence expansion depends on which specific gaze configuration occurs at sequence possible completion.

Now that we have a clearer picture of the strong association between the occurrence of specific gaze configurations and the expansion of courses of action, a closer look at the data can provide us with a more detailed understanding of the mechanisms at play, which lead to these correlations. First of all, both participants do not always look at each other during the turn(s) that lead(s) to possible sequence completion. Table 4.15 shows that although both participants look at each other in two-thirds of the sequential structure preceding a possible completion of a sequence, in 21% of these cases only one participant looks toward the other, while 12% occur without any participant gaze toward the other.

*Table 4.15* Occurrence of gaze toward other participant before sequence completion

| Gaze behavior | Instances |
|---|---|
| No gaze at all | 56 (11.9%) |
| Only one participant looks | 97 (20.7%) |
| Both participant look at each other | 316 (67.4%) |
| Total | 469 (100%) |

The high number of sequential structures produced without any gaze (i.e., almost 12%) can be explained mainly in two ways:

1) both participants are looking toward something else, usually because of competing activities such as eating, drinking, or reading;

2) many of the sequential structures in which no gaze occurs are actually sequence closing third turns.

Table 4.16 and Figure 4.5 show that 80% of the time, both participants look at each other during post-expansions of base sequences (Post Ex. BS), while they do so only 50% of the time during sequence closing thirds (SCT). They also show that approximately one quarter of base sequences and sequence closing third turns are produced with only one participant looking at the other.

*Table 4.16* Gaze behavior during sequential structures preceding sequence possible completion

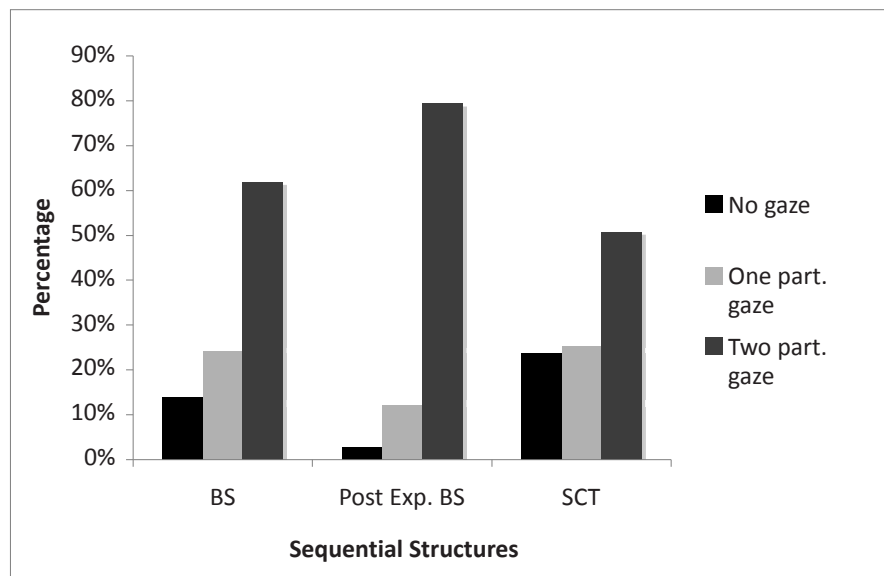| Gaze | BS | Post Exp. BS | SCT |
|---|---|---|---|
| No gaze at all | 37 (14.0%) | 4 (2.8%) | 15 (23.8%) |
| Only one participant looks | 64 (24.2%) | 17 (12.1%) | 16 (25.4%) |
| Both participants look at each other | 164 (61.9%) | 112 (79.4%) | 32 (50.8%) |
| **Total** | 265 (100%) | 141 (100%) | 63 (100%) |



*Figure 4.7* Gaze behavior during sequential structures preceding sequence possible completion

The data presented in Table 4.16 and Figure 4.7 suggest an interesting association between the co-occurrence of specific gaze behaviors (e.g., the small number of post-expansions produced with no gaze compared to the relatively large number of sequence closing thirds produced with no gaze) and the tendency of these structures to be followed by sequence expansion or not, as described in Table 4.6. One possible explanation for this association is that participants provide multiple simultaneous cues to the other participants in terms of their interpretation of the gist of the ongoing course of action and whether it might be brought to

300

closure or not. Alternatively, the association between sequential structures (and conversational units) and sequence expansion might be actually driven by the gaze behavior implemented during them. The first explanation seems more probable than the second, mainly because of evidence from telephone conversations. However, further work comparing such conversations with a larger corpus of video-recorded, co-present interaction is necessary in order to establish this conclusively.

Finally, while describing what was happening in terms of gaze in the examples in § 4.4 and 4.5, it was reported that the first person to withdraw gaze when both participants end up in configuration 1 (i.e., both participants look away at sequence completion) was the person who had initiated the course of action. If we focus on the 85 sequence possible completions in which configuration 1 is achieved, after both participants had looked toward the other, an interesting finding emerges. Table 4.17 shows the distribution of sequence non-expansion in relation to who looks away first.

*Table 4.17* Distribution of sequence expansion by which participant first looks away preceding gaze configuration 1 ( )

| First who looks away | Sequences expanded | Sequence not expanded |
|---|---|---|
| Initiator of course of action | 5.9% (3/51) | 94.1% (48/51) |
| Other participant | 23.5% (8/34) | 76.5% (26/34) |

Table 4.17 indicates that when the first person who looks away is the person who initiated the course of action, sequence closure is significantly more likely (Fisher Exact test $p < .05$) than when the first person looking away is the recipient of the turn initiating the course of action.

This section has shown how multiple factors affect the occurrence of sequence expansion (e.g., type of sequential structure and preference of what preceded the sequence possible completion). If both people sustain gaze toward each other at possible completion of a course of action, then the course of action is usually expanded, while if both participants look away, the course of action is usually brought to closure. Moreover, if only one participant keeps looking at the other at possible completion of the course of action, the course of action is likely to be expanded, though this is significantly less so than if both participants were looking at each other in the same sequential environment. Finally, if the first person who looks away is the initiator of the course of action, when both participants have been looking at each other, the course of action is more likely to be brought to closure at

the first possible completion of the sequence than if the first person looking away is the recipient of the initiating action.

In what follows, cases that do not exhibit this pattern will be considered and an interactional account for their occurrence provided.

## 4.7 "Deviant Cases"

Although § 4.6 documented the regularity of the patterns outlined in § 4.4, an argument that these patterns are normative requires not only that they usually occur but that participants are oriented to their non-occurrence. Thus, in this section the 5% of cases in which both participants keep looking and the sequence does not get expanded are examined.

### 4.7.1 Gaze Up → No Sequence Expansion

Previously, the claim has been made that if both participants sustain their gaze toward each other, the sequence is treated as incomplete and should be expanded. However, the quantitative distribution of cases shows that in 6 cases of mutual gaze at possible completion of the sequence, the sequence does not get expanded.[16] A closer look at these cases shows that sequence expansion does not occur for the following reason: in the context of a challenge or reiterated disagreement, avoiding expansion averts conflict and resists aligning with the other participant.
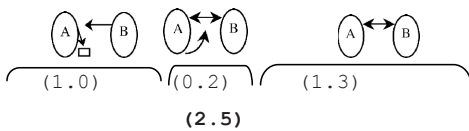
Example 4.17 represents a case in point. In this example, a complaint does not get an immediate response and the response is pursued through gaze (see chapter 3). However, a response is performed gesturally, although it is not clear whether it is appropriately responding to the complaint or doing teasing. The person who initiated the complaint is the first person who looks away, although she does briefly look up again to monitor what the other participant is doing. The participant who produces the SPP then looks away, and, at this point, also the initiator of the sequence withdraws her gaze and re-engages the competing activity of eating. At the beginning of this excerpt, A complains about the fact that there was no surprise inside an Easter chocolate rabbit she recently bought (line 2). Our focus is the gaze behavior during line 5.

---

[16] The following analysis will not present all 6 deviant cases but the argument put forward nevertheless holds for them all.

**(4.17) 2PLUNCH1-sorpresa 9:12**

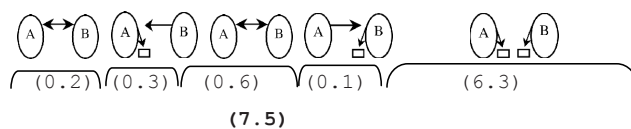01      (2.5)

02 A: *Io ci son rimasta male che non c' era  la sorpresa.*
      I  cl. am left    bad that not cl. was the surprise
      **I was disappointed because there was not the surprise.**
   **((Inside the chocolate rabbit))**

03        (1.0)   (0.2)   (1.3)
               **(2.5)**

04 ((B makes gesture + facial expression like "what a pity", Fig. 4.8))



Figure 4.8. Frame representing line 4

05     (0.2) (0.3)  (0.6)  (0.1)   (6.3)
               **(7.5)**

06 A: *Beh insomma Angela ha vinto l'  Erasmus ad Heidelberg (0.2)*
      Well in sum Angela has won   the Erasmus at Heidelberg
      **Well Angela won the Erasmus for Heidelberg (0.2)**

A pursues a response to the complaint during the silence at line 3 by looking toward the addressee who withholds responding to the gaze pursuit for 1.3 seconds, until he produces a facial expression and a mildly empathetic gesture that can be glossed as "what a pity" (it appears ironic). Immediately afterwards, A looks down briefly but B keeps looking. A looks

up again but this time B looks down. A looks down as well and they re-engage eating for some seconds. Then, at line 6, A starts a completely new course of action telling B some news about a mutual friend.

The fact that A looks up again and sustains brief mutual gaze after the responsive gesture suggests she either treats the gesture as insufficient *as* a response or she is checking whether B was actually teasing her; however, by the time B looks down, she aligns with A by also looking down and re-engaging with eating. On the other hand, B had just been accused by A of having eaten the entire chocolate rabbit and A has admitted his fault. Thus, he is clearly withholding response also after the engagement of mutual gaze at line 3. After the accusation, A seeks an affiliative response while B may continue looking to understand what A is trying to do and whether she is possibly launching a telling.[17] This means that at possible sequence completion both participants could be orienting toward the possibility of sequence expansion for very different reasons. And, besides these reasons, A looks away first but then also looks back to monitor whether B has actually aligned with her gaze withdrawal (A is the person who initiated the sequence).

Participants may also resist providing a response or a sequence expansion and evade the situation by unilaterally starting a new course of action, at a point in which further sequential talk was expected instead. In example 4.18, A has just reported to B that some of their colleagues did not attend one of their colleague's husband's funeral. After admitting that she did not go either, B also admits that she has not even called this colleague (presumably to give her condolences). In the telling that follows (lines 4-8), A reports that someone else did not go to the funeral and did not call their colleague and yet the colleague who lost her husband decided to call the one who had not attended the funeral after receiving a text message. Our target here is what happens during the silence at line 12.

**(4.18) 2PERLINE-chiamata 14:12**

```
01     (1.4)

02 B:  Io non l'ho       chiamata
       I  not her have.1s called
       I did not call her
```

---

[17] Notice that some of the features of line 2 correspond to the ones discussed in chapter 2 as being usually associated with beginnings of extended tellings.
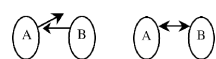
```
03      (0.4)

04 A:   .hhh Io so      che  Marco le ha mandato un messaggio
        .hhh I   know.1s that Marco her has sent   a   message
        .hhh I know that Marco sent her a sms

05      dicendo=scusandosi del fatto che non: non l'aveva
        saying  apologizing of fact  that not not her had
        saying=apologizing for the fact that not he had not

06      chiamata e   che  non si sentiva di chiamarla   e   lei
        called   and that not cl felt     of calling her and she
        called her and that he had not felt like calling her and she
```
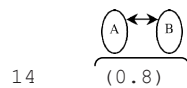


```
07      l'ha chiamato  e   ha  detto °ma  no insomma (.)
        him has called and has said  but no I mean
        called him and she said °but no I mean (.)
```
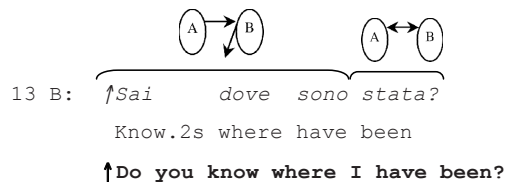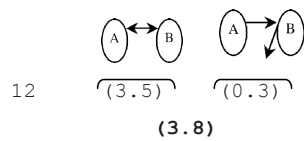


```
08      si sono parlati°
        cl are  talked
        they talked°
```



```
09      (0.7)
```



```
10 B:   L'ha     chiamato lei,
        Him has called    her
        She called him,
```



```
11 A:   Si'
        Yes
        Yes
```

12   (3.5)   (0.3)

       **(3.8)**

13 B:  *↑Sai     dove   sono  stata?*
       Know.2s  where  have  been

       **↑Do you know where I have been?**

14   (0.8)

15 B:  *Cs- uh::::: sabato   sera (0.5) sono andata a  casa della*
         uh     Saturday evening     have gone   to home of
       **Cs- uh::::: Saturday evening(0.5) I went to the house of**

After A's report, B asks A for confirmation about her understanding that the widow actually called the person who did not attend the funeral of her husband and this is confirmed by A (line 11). At this point, the request for confirmation is complete and so is A's telling. Given that B has not gone to the funeral and has not called her colleague, A's telling is an attempt at consoling B for not having done what is expected with regards to their mutual colleague's situation. Moreover, once A confirms that their colleague called someone who did exactly what B has done, it would seem possible for B to acknowledge A's action in some way (e.g., state that she, too, could send a text message to their colleague). This means that this point represents possible sequence completion (the end of line 11). However, a closer look at the interactional contingencies of the course of action they are engaged in suggests that further sequential talk by B might occur at this point. Both participants keep looking at each other for 3.5 seconds (line 12), until B looks down and starts a complete new course of action at line 13 (where she projects the delivery of a telling that starts at line 15). By starting a new course of action, B does not publicly acknowledge A's telling as a candidate solution to B's problem, and at the same time, B does not state that she will do anything to fix the situation she was complaining about. By starting a new sequence, she avoids accepting or rejecting A's indirect advice.

The fact that B looks away 0.3 seconds before starting a new course of action also becomes a public display to A that she will not continue on the previous course of action. Indeed, it is interesting that before moving to a new course of action, and after having sustained mutual gaze, the speaker of the new sequence first withdraws from mutual gaze and only then starts the new turn of talk. While it could be claimed that there are cognitive reasons for looking away before producing a new turn of talk (see, e.g., Kendon, 1967 who claims that speakers look away before speaking to help planning their turn), it is significant that this looking away does not occur before *every* turn of talk, but rather before starting a new course of action once another one is still potentially open. Thus, in example 4.18 we have seen that in a situation in which some talk by one participant is potentially relevant—though not conditionally relevant—next, that participant can resist providing such talk, and actually evade the situation by starting a completely new course of action.

This section has shown how the few cases in which participants sustain mutual gaze and the sequence does not get expanded are explainable by providing a closer look at the interactional contingencies relevant in the local context in which they occur. Such "deviant" cases all involve one participant resisting aligning with the other and they also tend to occur in the context of (in)direct conflict or sanctioning. Such situations mainly demonstrate two things. First, that participants treat the sustaining of mutual gaze in this context as unusual or problematic, and, consequently, either repeatedly monitor what the other participant is doing with her/his eyes (see example 4.17), or unilaterally depart from the mutual gaze by looking away and starting a complete new course of action (example 4.18). In both examples, participants disengage from mutual gaze before the beginning of the next turn of talk, thereby displaying a clear orientation toward the normativity of gazing away to mark sequence completion. While they initially orient to possible sequence expansion at a TRP by sustaining their mutual gaze, during this TRP one participant initiates closing by looking away. This move toward closure can be acknowledged and "mirrored" by the other participant (4.17) or can be ignored (4.18) so that what follows is a unilateral closing and the beginning of a new course of action. Second, both examples confirm that transition relevance can be extended through time (metaphorically, it is a place and not a point). As a consequence, these examples suggest that limiting analysis to numbers (as is necessary for quantitative coding), which treats the TRP in terms of a very limited amount of time around turn completion, may mask our understanding that apparently 'deviant' cases are not actually deviant but in fact display the same orientation toward the norm as the other examples. If the norm is to expand the sequence when participants keep looking at each other at possible completion of a course of

action, then not expanding indicates some form of resistance. Returning to the cases that did not fit the proposed pattern sheds additional light on the nuances of how gaze behavior is related to sequence organization and the organization of action in interaction.

## *4.8 Discussion*

This chapter outlines specific patterns of gaze behavior occurring when participants approach possible completion of a course of action, and shows their interactional uses. It is an empirical and methodological issue to show that participants orient toward a gesture, posture, facial expression or gaze behavior as a practice for doing something and not toward something else in the conversation. It is also an issue to show that there is a possible causal link between the occurrence of visible behavior and the reactions that participants produce. As Schegloff (2004: 22) states:

> *Efforts to ground analytic characterizations of conduct in interaction in the demonstrable orientations of the participants are subject to challenge if the participants do not overtly announce such orientations. […] How in that case can a claim be warranted? Or how can such a linkage be disputed if there are theoretical grounds for expecting it? In general, how do we know anything about the orientations of persons to their environment, setting, "umwelt," "context," etc., or specific aspects of it, without their articulate expression? Can we know such things, and, if so, how?*

This work has provided support for specific characterizations of visible behavior as effective interactional moves. Such characterizations are possible by building up collections of examples that take into account different variables and by choosing aspects of behavior that are visible, ostensive and can be enacted by other participants as well.

The general findings presented in this chapter are the following:

When a sequence *can* end,

1. if both look away, it usually *will* end.

2. otherwise it will usually be expanded.

Qualitative and quantitative evidence for these claims has been provided and all the cases that do not seem to fit with the predicted pattern have been accounted for. Some questions remain in terms of the import of these findings, and, in what follows, I outline some of them.

First of all, these gaze practices confirm something conversation analysts have long argued about the relationship between sequences and actions (e.g., it is not always the case that a course of action is accomplished via a single sequence of talk; see, for example, Schegloff, 2007b). In addition, however, this chapter shows how participants orient toward *action* completion rather than only *sequence* completion. As shown in § 4.5, gaze is sensitive to participants' understanding of where they are in a course of action. If gaze is sustained by participants once they get to possible completion of a sequence, this displays an orientation toward that sequence as one step within a course of action. A course of action that can be projected by what gets said and done in every single sequence. Gaze withdrawal, on the other hand, deals with action closure. It signals closure even when the action cannot be accomplished because of specific contingencies. So, for example, a glance at example 4.14 might suggest that what precedes a complaint are just two different sequences that request specific information. But if we look at the gaze behavior of the participants, we can see that they orient toward each sequence as one step in a larger course of action or interactional project. In differentiating between "sequences" and "interactional project", or more specifically between "sequences" and "courses of action", this chapter follows a distinction previously suggested by Clark (1996) and Schegloff (2007b) and currently further emphasized by Levinson (in press). In particular, this chapter provides further empirical evidence for the necessity of this distinction.

This chapter has also addressed Schegloff's work on sequence organization and his recent distinction between the unit "sequence" and courses of action performed through multiple sequences. The analysis put forward here suggests through their gaze behavior, participants orient toward completion of courses of action independently of whether those actions are accomplished through a single sequence or multiple ones. It therefore confirms the relevance of showing the close relationship between these extended courses of action and single sequences.

There are a number of reasons for distinguishing single sequences of talk from "sequences of sequences". Apart from noting their labels and structural similarities, it is also important to see what participants make of them. As far as gaze behavior is concerned, this chapter shows that participants more than orienting toward these two structures, orient toward the accomplishment of courses of action. A completion of a course of action naturally

constitutes also the completion of a sequence of talk or of a sequence of sequences. This is because sequences refer to a structure related to the relationship between turns produced by different participants, while course of action refers to what participants are trying to do via those turns, in other words their illocutionary force (Austin, 1962). This work also further contributes to research on sequence organization by showing an orientation by both participants toward the adjacency pair as not only a possible unit for building action, but the most minimal sequence (Schegloff & Sacks, 1973). Considering only the range of actions presented in this chapter (and not for example sequences of talk that might occur in a classroom interaction), a sequence with a third position turn is treated as an expanded sequence. This does not mean, however, that there are no actions that by default require more than two turns (e.g., summons-answer-reason for the summons). It simply shows that there are many sequences for which two turns are necessary and sufficient.

A more general point throughout this dissertation concerns what gaze can tell us about how participants process and analyze what is happening on-line and where they are in the development of a course of action. This is crucial for performing appropriate next actions and avoiding accountability for misunderstanding of which action was expected from us. We show others how we understand what has just happened in terms of being the possible end of the course of action or not. And, gaze can be an additional window into how others understand where they are in the ongoing sequence. Coordination of closure is not additive. It is mutual. An individual can show an orientation toward the norm (as I have shown throughout this chapter), but cannot usually accomplish efficient closure unless the other aligns with it at a certain point. One can try to initiate closure, but this does not mean that s/he can easily accomplish it (see example 4.17).

These last considerations resonate with what Schegloff and Sacks (1973: 290) wrote about closing a conversation:

> *The 'closing problem' is […] a problem for conversationalists; […] closings are to be seen as achievements, as solutions to certain problems of conversational organization.*

Why is closure achieved when both participants coordinate their withdrawal of gaze before or at possible completion? Schegloff and Sacks (1973: 297) explained the necessity of two utterances produced by different speakers to do openings or closings of conversations:

*By an adjacently positioned second, a speaker can show that he understood what a prior aimed at, and that he is willing to go along with that. Also, by virtue of the occurrence of an adjacently produced second, the doer of a first can see that what he intended was indeed understood, and that it was or was not accepted.*

The adjacent positioning of gaze withdrawal is relevant for coordinating sequence closure. If gaze toward the other participant at possible completion of a sequence indicates that a possible expansion could occur, then no-gaze by both participants in the same sequential position indicates an implicit agreement toward non-expansion.

But who should initiate closure and what are the consequences? As we have seen, it is rare that one interactant lowers her/his gaze then looks up briefly to check whether the other is looking down too. Accordingly, there must be some kind of preference for the initiation of closure by one of the participants in a way that allows the other one to recognize the action and possibly (dis)align with it. Throughout this chapter it was noted that often the person who first withdraws her/his gaze is the person who initiated the entire course of action and the other usually follows. Table 4.17 provided quantitative evidence that the participant who looks away first affects the likelihood of the occurrence of sequence expansion. There are obvious reasons for this such as the usual order of gaze withdrawal: if a participant is trying to accomplish some action, once that action is accomplished, s/he can display her/his satisfaction by closing it down. As far as that individual is concerned, nothing further is necessary. However, it can happen that the person who initiated the course of action looks away but at that point the course of action is potentially changing its trajectory (see, for example, 4.17-4.18). Or, it is possible that the person who looks away first is not the person who started the ongoing course of action, but is nevertheless the one who wants to end it, possibly to initiate a new one. These possibilities show that every completion is, indeed, an interactional achievement as it deals with the contingencies of what can happen during the development of a course of action. Moreover, this shows an asymmetry of agency in action sequences: the one who instigates the action sequence is the 'owner' of it, meaning that he/she has a greater right to determine its closure or expansion (Enfield, 2011).

I have already stressed that this gaze behavior has the double feature of being context free and context sensitive at the same time. All four of the gaze configurations can occur and the implications remain the same, independent of the conversational topic and of the

participants.[18] This makes it "context-free". On the other hand, as shown in § 4.7, local contingencies can account for what appears to be, at least, slight deviation from the usual pattern. This makes it "context-sensitive".

How can people project whether a turn is a first pair part or the beginning of a sequence of sequences? That is, how do people know that something is a pre-sequence rather than a base sequence? The fact that participants sustain mutual gaze at completion of sequences, which are part of larger courses of action, confirms that people can systematically distinguish a pre-sequence from a base sequence. Our capacity to project what will come next in a conversation and display our understanding on-line is a crucial feature that allows our interactions to proceed so rapidly and effectively, a matter taken up in the conclusions.

A final, more general question to address here concerns whether gaze withdrawal is doing closing. Is it proposing closure? Is it displaying participants' engagement in the conversation? Or is it just an emergent structure of the concurrent occurrence of closure of the sequence? In other words, is gaze withdrawal a signal people use or is it something that co-occurs with sequence closure, but does not play any specific role in expanding or closing the sequence? By looking away and not looking back at the other participant, one displays a stance toward and an understanding of the course of action that can be aligned with or not by the other participant. One participant's gaze withdrawal does not make anything conditionally relevant, but it does make a bid for closing. If the gaze withdrawal is aligned by the other participant, then closure can occur and usually does occur. If not aligned, usually some sequence expansion will occur. The claim here is that gaze negotiates sequence closure in cases where closure is in principle problematic; while in more routine cases, it may merely signal understanding that the sequence is complete. In other words, it can be both a signal and a symptom.

---

[18] It is an open empirical question whether the meaning of gaze behavior at possible sequence completion changes in specific institutional settings. It is, on the other hand, very likely affected by culture.

# 5 Conclusions

*In all viable systems, there must be an area where the individual is free to make choices so as to manipulate the system to his advantage.*
Leach (1962: 133) On Certain Unconsidered Aspects of Double Descent Systems

## *5.1 The Findings*

This dissertation has examined how humans use their eyes and their bodies during face-to-face interactions. Previous research on eye gaze in face-to-face interaction claimed that gaze behavior was affected by variables exogenous to the interaction (e.g., gender, age, social status; see, e.g. Ellsworth & Ludwig, 1972; Kleinke, 1986 for reviews), that it displays the degree of engagement of participants in the conversation (e.g., Argyle & Cook, 1976; Goffman, 1963; Goodwin, 1981), that it has a regulatory role in turn taking (e.g., Duncan, 1975; Kendon, 1967) and, finally, that it has a role in implementing social actions, depending on its context and delivery (e.g., Kidwell, 2005, 2009; Sidnell, 2006). This dissertation provides several key findings that necessitate a rethinking of some of these claims.

Chapter 2 shows that when a person is the recipient of a multi-unit telling, they will tend to direct their gaze to the speaker during the first turn of that telling. But when a person is the recipient of the first turn of an adjacency pair, they do not show a particular tendency in gaze behaviour. As such gaze is used differently depending on the social actions and activities in which interactants are involved. Interactants are shown to have different norms for gazing at their co-interactants depending on whether they are involved in extended multi-unit turns (extended tellings) or turn-by-turn talk (adjacency-pair-based sequences). In the former context, listeners should gaze at speakers within the completion of the first TCU of the projected telling. By contrast, in turn-by-turn talk such as question-answer sequences, listener gaze is not treated as normatively required during the first TCU of the sequence initiating action. This poses a puzzle for how, within the first TCU, listeners ascertain whether or not they are hearing the first TCU of an extended telling or a single TCU announcement, for instance. The answer lies, at least in part, in the fact that listeners rely on "fast and frugal heuristics" (Gigerenzer & Goldstein, 1996), based on the semantic domains entailed in the utterance, to recognize the course of action the speaker is initiating. For

example, while TCUs that initiate tellings often have a first person (singular or plural) subject, contain time references to the past or the future, indicate epistemic access to the events reported, and/or mention a third person not previously discussed, first TCUs that initiate adjacency pairs often have a second person (singular or plural) subject and contain deictics and modal verbs. The different categories of words that participants choose to design their sequence initiating actions provide cues to the recipients about whether that turn is the beginning of an extended telling or the first pair part of an adjacency-pair-based sequence.[1] The findings of this chapter provide a refining of Goffman's (1981a) analysis of the different types of hearers that may occur in a social situation. Goffman differentiated between "official" and "unofficial" participants. Within "official" participants, he distinguished "addressed" and "unaddressed" recipients and among the "unofficial" he differentiated "eavesdroppers" from "overhearers", "bystanders" and "audiences". Chapter 2 shows that within the category of "addressed recipients", we can further distinguish a recipient of an extended telling from a recipient of the first pair part of an adjacency pair sequence. This added specification refines our understanding of participation in a social situation and recognizes the behavior that an individual is expected to implement for acting as an attentive and "addressed" recipient. These findings, then, indicate the need for a recontextualization of many of the prior claims about gaze, in particular the ones developed in experimental settings, since those studies did not usually take into account the content of the conversation and the social actions performed through each turn at talk. Moreover, the findings reported in chapter 2 suggest that prior studies suffered from the methodological problem of not controlling for conversational activity type, to be discussed further in § 5.2.

Chapter 3 describes the way gaze can be used to mobilize recipient response in conversation. Whereas previous work focused on vocal resources and how they might be implemented to pursue recipient response, the findings reported in this chapter reveal that gaze alone is sufficient to mobilize response. And, it shows that this can be done not only at a transition relevance place but also in the context of a story telling, after a laughable or as a try-marker. These findings suggest that gaze is a resource for doing far more in interaction than was previously shown. Specifically, chapter 3 shows how, even in the context of silence, speaker gaze can mobilize recipient response and this finding adds to prior literature on pursuits and on linguistic features that might be deployed for mobilizing responses, such as syntax or prosody.

---

[1] See Levinson (in press) for an overview of action formation and action ascription in social interaction.

314

Finally, Chapter 4 discusses how speakers navigate the closure of sequences. The analysis shows that participants orient to the relevance of gaze withdrawal at sequence completion when this coincides with the completion of a course of action. This contrasts with previous studies that argued that gaze was primarily related to turn taking. The data reported in this chapter show that participants' gaze is sensitive to where participants are in a course of action. When participants can and want to close a sequence, they will tend withdraw their gaze. In contexts where gaze is not withdrawn, sequences are expanded until they can be closed in the absence of interactants' gaze.

Tables 5.1 and 5.2 summarize how the chapters of this dissertation have explored both the speaker's and the recipient's gaze behavior at the beginning, during and at completion of the interactional environments discussed in chapter 2, namely, Extended-Telling Sequences (ETS) and Adjacency-Pair-Based Sequences (APBS).

*Table 5.1* Speaker and recipient gaze behavior documented for ETS.

|  | **First TCU** | **During telling** | **Possible completion** |
|---|---|---|---|
| **Speaker** | (Minimally Ch. 2) | (Minimally Ch. 3) | (Minimally Ch. 4) |
| **Recipient** | Ch. 2 | Ch. 2 | Ch. 4 |

*Table 5.2* Speaker and recipient gaze behavior documented for APBS.

|  | **First TCU**[2] | **Between FPP and SPP** | **Possible completion** |
|---|---|---|---|
| **Speaker** | Ch. 3 | Ch. 3 | Ch. 4 |
| **Recipient** | Ch. 2, Ch. 3 | Ch. 3 | Ch. 4 |

Tables 5.1 and 5.2 indicate that this dissertation has documented both participants' gaze behavior in both sequential environments. This is the first work to extensively document the systematicity of gaze behavior at this level of granularity, and the first to show such a strong correlation between gaze behavior and the overall interactional project, the participants' goals and the actions participants accomplish. The analyses here presented suggest that the *principal* units around which gaze behavior in face-to-face conversation is organized are not just turns at talk, but rather the achievement of course of action and the organization of the

---

[2] See also Rossano (2010) and Rossano et al. (2009) for work on gaze during the first TCU of APBSs, with a focus on questions.

sequential structures that were used in order to do so. Some specific gaze behaviors may be related to other cognitive or biological factors, rather than a particular sequential environment (e.g., closing one's eyes during a word search to facilitate a lexical search, and looking at objects before, and at times during, reaching for them, respectively). While it might be tempting to interpret the examples just mentioned as deviations from the interactional order outlined in this dissertation, I rather see them as additional factors affecting the orderliness of gaze behavior. They could also be interpreted as part of the "natural meaning" of gaze: It remains to be established what takes precedence between the biological, cognitive/processing and interactional factors when their outcome would convey conflicting signals (for example if one would like to grab a glass but also continue the sequential development of a course of action, withdrawing gaze from the addressee at possible sequence completion to look toward the glass might be problematic and be misinterpreted). Note also that the regulatory functions of gaze behavior reported in this dissertation are not independent of what is done and said during any interaction and the courses of actions accomplished within that interaction, and so they do not apply to every single turn at talk.

## 5.2 Methodological Contribution

Harvey Sacks once remarked that "the important theories in the social sciences have tended to view a society as a piece of machinery with relatively few orderly products, where, then, much of what else takes place is more or less random" (1984: 21). He was, with Garfinkel (1967), one of the first to suggest that orderliness of social conduct in social interaction should be an assumption, rather than an exception, otherwise social life and social understanding for practical purposes would be impossible. This dissertation extends Sacks' claim by showing that not only *talk*-in-interaction, but also *gaze behavior* in talk-in-interaction is orderly and not random. To discover the orderliness of gaze behavior, I focused on a series of practices implemented through gaze. Following Heritage (2011: 213), I use the term "practice" in the following way:

> A 'practice' is any feature of the design of a turn in a sequence that (i) has a distinctive character, (ii) has specific locations within a turn or sequence, and (iii) is distinctive in its consequences for the nature or the meaning of the action that the turn implements.

For example, one of the practices described in this dissertation is pursuing responses through gaze. The practice consists of a speaker looking toward the recipient at the TRP following a sequence-initiating action that has not yet been responded to. The "distinctive consequences" are that recipients often interpret this gaze behavior as a pursuit and upon its occurrence provide a response. Part of the "distinctive character" of the social action implemented through this gaze behavior is that a response is pursued without officially treating the lack of response as evidence of forthcoming dispreferred response or trouble in hearing or understanding (an argument developed in chapter 3).

To identify and analyze some of these different practices implemented through gaze in social interaction, I relied on a systematic investigative method that developed through the following steps:

*Step 1*: A detailed analysis of a corpus of interactional fragments led to the development of a characterization of a practice, of the variables relevant for the occurrence and effectiveness of the practice and of the interactional environment in which the practice would usually occur (e.g., recipients looking toward speakers during the first TCU of an ETS or an APBS).

*Step 2*: Using a different portion of the dataset, the data were then coded for multiple factors to establish the systematicity of the practice and test its effectiveness statistically within a specific interactional domain (e.g., by selecting all first TCUs of ETS and APBS in the dataset and verifying how often recipients would look toward speakers if they were not already looking).

*Step 3*: Deviant cases were analyzed to assess the accuracy of the initial characterization of the practice (e.g., by looking at what happened in the instances in which a recipient did not look at the speaker within the first TCU of an ETS).

The method of analysis here reported can be summarized as follows:

*Single instances → model of practice → test reliability of practice → account for deviant cases*

To analyze the corpus of interactional fragments (step 1) from a qualitative perspective, I adopted the methods of conversation analysis (CA). Pomerantz (1990: 231) outlines three types of claims analysts often provide:

a. A characterization of the social action investigated (e.g., offer, request, invitation);

b. Analyses of methods [i.e. practices] used by the interactants for accomplishing that action;

c. A detailed characterization of the sequential features and interactional consequences of those method(s) [practices].

After developing a characterization of the possible practice from a corpus of possible instances (step 1 of my methods above) as Pomerantz proposes, in step 2 I did not look for more instances of the same practice (e.g., I did not look for all the times in which a recipient might look at a speaker during a first TCU of an ETS or an APBS). Rather, I looked for the slot relevant for the occurrence of that practice and investigated the gaze behavior that occurred in that slot, coding what happened, interactionally, after the implementation of the practice or the occurrence of any alternative gaze behavior in that "slot". By doing so, I have attempted to test the accuracy of my characterization of the practice and its function by attempting to falsify it (Popper, 1959 [1934]).

In step 2 of each chapter this dissertation combines qualitative observations developed through CA with quantitative methods in order to document the robustness and reliability of the practices investigated across the data. The use of quantitative methods has been debated within CA. On one hand, there are strong proponents of the inclusion of frequencies and statistics in analytic reports, especially in the context of applied research (e.g., Robinson, 2007; Heritage, 1999). On the other hand, there are those who consistently warned about losing track of the specificity of each instantiation of every interactional practice when they are grouped for the purposes of abstract quantification (e.g., Schegloff, 1993, 1997). Quantitative methods nevertheless continue to be adopted by some CA scholars. Sometimes the frequencies of the distribution of some interactional phenomenon are provided (Egbert, 1996; Ford & Thompson, 1996; Jefferson, 2004b), while other times, correlations between interactional phenomena are described (e.g., social action performed through talk and occurrence of applause in political speeches, see Heritage & Greatbach, 1986). Perhaps most often, quantitative approaches are applied in the context of institutional settings, in order to assess the association between an interactional phenomenon and some exogenous variable, be it time (e.g., in Clayman, Heritage, Elliott, & McDonald, 2007), race or class (e.g., in Stivers

& Majid, 2007) or prescribing outcomes or provider perception (e.g., in Mangione-Smith, Stivers, Elliott, McDonald, & Heritage, 2003). Most recently, some conversation analysts have used quantitative methods for analyzing ordinary conversation from a cross-cultural perspective (e.g., Fox et al., 2009; Rossano et al., 2009; Stivers et al., 2009). By combining qualitative and quantitative methods, the findings significantly enhance our understanding of the functions of gaze in social interaction by showing the practices described to be recurrent, systematic and reliable in their production. I have addressed the problems of quantitative work by a) first establishing the validity of the practices I am counting, and b) not allowing the 'deviant' cases to be dismissed as mere noise.

This dissertation also shows the advantages of relying on large corpora in that it reveals the robustness of specific claims about interactional practices. One of the goals of building a corpus of instances of a practice is to assess the systematicity of that practice and what factors might play a role in its deployment and effectiveness (see, e.g., Schegloff, 1996a, 1997). Single-case analyses can help us appreciate the complexity of a practice or raise the possibility that a practice might deal with a specific interactional issue (e.g., Goodwin, 1979). However, if one describes an interactional phenomenon as a practice but only looks at a single example then s/he is simply hypothesizing that this behavior is not idiosyncratic but is, in fact, a practice. A more systematic investigation (including an analysis of a large corpus of data and examples that range in terms of type of interaction and the participants involved) is necessary to confirm that hypothesis and identify its normative implications. Notice here that the contrast is not between doing single-case analyses and developing large corpora, as both options are important and central components to the enterprise of understanding social behavior. The contrast rather lies in the goals of the analytic enterprise. Indeed, the issue often lies in the attempt of making general claims about the normative dimensions or systematic implementation of some behavior relying on data from a single interaction or a single participant in that interaction. This is problematic for two reasons:

1) Idiosyncratic behaviors exist and, even in a normative system, sometimes participants "let [them] pass" (Garfinkel, 1967) or avoid displaying an orientation to a violation (e.g., Robinson, 2007) either because they are familiar with the idiosyncratic behavior of that individual or they want to avoid altering the progressivity of the conversation (Schegloff, 2007b);

319

2)      Confirmation biases exist and we tend to notice the existence of infrastructures and practices when they fail or break down. So, without attempting to falsify a claim, the risk is to present something that is actually not normative nor typical as if it were normative or typical.

An inductive approach to the data investigated in this dissertation might have led to presenting 3 examples of recipients not looking at the speaker at the beginning of an ETS (see chapter 2) as evidence of a practice aimed at avoiding looking at speakers at beginning of tellings, or 6 examples of sequence closure after sustaining mutual gaze at possible sequence completion (see chapter 4) as evidence of a practice aimed at suggesting sequence closure via sustaining mutual gaze at possible completion of a sequence. In both cases, such a report would not account for whether such behaviors are representative or rather atypical. Typicality can be even more difficult to assess since certain behaviors can occur in a specific environment in a significant number of cases and yet not be the 'typical' or 'normal' thing to do. For example, Rossano et al. (2009) showed that in approximately 75% of questions in 3 unrelated languages (out of 300 questions per language), speakers look toward their addressee. This means that for each language, the authors of this study could have presented up to 75 instances (the remaining 25% of the cases) in which the speaker *did not* look toward the addressee during questions and claimed the behavior to be "typical" for the speakers of those languages. In doing so, they would have represented a behavior that occurs only approximately one fourth of the time as representative of the default strategy or most common gaze behavior for naturally occurring questions in each of the three languages. One of the tasks for the analyst remains accounting for the occurrence of cases either considered deviant, or that do not follow the usual deployment or effectiveness of the practice identified. For all these reasons, it seems clear that combining qualitative and quantitative methods can only add strength to our claims, rather than weaken them. According to Max Weber (1978 [1922]: 12):

> *A correct causal interpretation of typical action means that the process which is claimed to be typical is shown to be both adequately grasped on the level of meaning and at the same time the interpretation is to some degree causally adequate. [...] even the most perfect adequacy on the level of meaning has causal significance from a sociological point of view only insofar as there is some kind of proof for the existence of a probability that action in fact normally takes the course which has been held to be meaningful. For this there must be some*

> *degree of determinable frequency of approximation to an average or a pure type.*

A closer look at the data presented in Table 5.3 can further illustrate the importance of combining the strengths of both qualitative and quantitative approaches to social interaction. In the previous chapters, statistics concerning participants' gaze behavior during 10 minutes selected from each interaction were reported (for more details about each interaction, see Appendix A). Table 5.3 reports the total percent and amount of time each participant in each interaction looks toward the other participant's face and the amount of time the two participants engaged in mutual gaze.

*Table 5.3* Summary of frequency and duration of participants' gaze toward each other and of mutual gaze.

| Name of Extract | Time Gaze A→B | Time Gaze B→A | Time Mutual Gaze |
|---|---|---|---|
| 2PERLINE | 02:36 (26%) | 02:35 (26%) | 01:29 (15%) |
| 2GC | 05:30 (55%) | 04:56 (49%) | 03:18 (33%) |
| 2GSOFA | 07:44 (77%) | 08:06 (81%) | 06:17 (63%) |
| 2GGOSS | 06:55 (69%) | 08:59 (90%) | 06:08 (61%) |
| 2GSTUDYING | 06:07 (61%) | 05:22 (54%) | 03:20 (33%) |
| 2PCOMP | 00:54 (9%) | 02:31 (25%) | 00:35 (6%) |
| 2PLUNCH | 03:32 (35%) | 04:04 (40%) | 02:00 (20%) |
| 2PPLAN | 04:06 (41%) | 03:35 (36%) | 01:49 (18%) |
| 2GCOLL | 04:53 (49%) | 03:01 (30%) | 02:10 (22%) |
| 2PEXAM | 03:56 (39%) | 01:37 (16%) | 01:14 (12%) |

There are mainly three ways of interpreting the data reported in Table 5.3, but what I wish to highlight is how uninformative and arguably unproductive it is to simply report how long each participant looked toward the other during a particular interaction. The first way would consist in explaining the numbers as idiosyncrasies of the participants, related to individual differences (e.g., participant B in 2PCOMP, a young man, looks at participant A 9% of the time, while in 2GGOSS participant B, a young woman, looks at participant A 90% of the time). Another way of interpreting those numbers would assume that they are actually orderly, but would use macro-categories such as gender or type of relationship between the participants to account for them. A, third, alternative, interpretation, grounded in the

participants' behaviors in the interactions, is to examine the activities the participants are engaged in during the interactions under investigation. For example, participant B in 2PCOMP spends the majority of his time looking at a computer screen and a picture album, while participant B in 2GGOSS is mainly engaged in listening to participant A's extended tellings. Keeping in mind these different activities, if we then consider the examples of these participants' gaze behavior presented throughout this dissertation, such behavior does *not* in fact reflect any individual idiosyncrasies. Moreover, regarding the relationship between these two participants and their co-interlocutors, there is no doubt that they are all very close friends, but the 2PCOMP interaction is full of quips, jokes and teases to a much greater extent than the 2GGOSS one. So according to the relationship variable, there should not be much of a difference, but there is difference in terms of the activities that are performed throughout the interaction. The fact that the amount of mutual gaze in interaction does *not* necessarily indicate the degree of intimacy of the participants, contrary to what has been previously claimed in the literature (see, e.g., Thayer & Schiff, 1974), is confirmed by comparing any of the interactions in Table 5.3. For example, comparing the 2GC interaction (33% of time spent in mutual gaze) and 2PERLINE (15%), there is no difference in gender in terms of participants in the interactions, as all participants are women. In terms of the intimacy of the relationship, one might expect that the ones who are closer and better friends would look at each other more than the others. However, the two friends in 2GC had a problematic interaction in that they disagreed on multiple occasions during that conversation and as a consequence they stopped seeing each other after that interaction, while the participants in 2PERLINE have remained very good friends. On the other hand, the activities the participants are involved in, the type of sequences they produce and the amount of silence that occurs vary significantly. In 2GC, the participants are engaged in catching up over dessert and they produce multiple extended tellings. In 2PERLINE, the participants are also catching up, but while one is instructing the other on how to make flowers with beads. A large part of their conversation develops in turn-by-turn talk and they produce only few extended tellings. Also, to work on the manual task both of them must pay attention to the materials they are manipulating and as we know, this requires visual focus on the items that need to be picked up and moved. Like the previous two interactions compared, the participants in 2GC and in 2PERLINE do not show any noticeable deviation in terms of their gaze behavior as described in the terms of this dissertation. In other words, what participants do with their eyes in any interaction, regardless of the variation in terms of exogenous factors such as gender or degree of intimacy, can be explained across individuals mainly by taking

into account the activities they are engaged in, the actions they perform through their talk and the sequential environment in which their gaze behavior occurs. Thus, it is crucial to analyze and report exactly what the participants are doing during an interaction, both with their bodies and with their talk, if we want to understand any individual's gaze behavior. So, for example, rather than simply assuming that men look at each other less than women (see, e.g., Ellsworth & Ludwig, 1972), we might look at the nature of the interaction and the kinds of conversational activities men engage in as compared to women. Individual differences may of course exist, as may gender differences, but what has previously been reported as striking individual differences, gender effects, or clear indicators of the participants' relationship, might better be explained and accounted for by focusing on the structure of participants' talk in relation to the tasks and competing activities, they are engaged in. Thus, this dissertation not only provides evidence of the significant relationship between gaze, talk, action and activities but also highlights the importance of re-evaluating previous claims in light of a more nuanced understanding of this relationship in talk-in-interaction.

Along these lines, another methodological contribution of this dissertation concerns the focus on the relationship between what people are doing through their gaze behavior and their talk. Often analysts tend to focus only on talk or only on visible behavior and fail to appreciate the intricate connection between the two. In recent CA work, some authors have resisted using similar categories to describe the organization and function of talk and visible behavior in social interaction, for example, by suggesting that "there is no reliable empirical evidence for treating physically realized actions as being in principle organized as adjacency pairs" (Schegloff, 2007b: 11). For others, "[…] nonverbal conduct is subordinate to the verbal conduct with which it is intermeshed; it's probably true to say that none of the practices, devices or patterns identified in CA research are shaped or altered in any significant ways by accompanying nonverbal conduct" (Drew, 2005: 78). Yet, a number of scholars of social interaction have repeatedly indicated and demonstrated the importance of integrating visible and verbal behavior in the analysis of what participants are doing during an interaction. Some have done this, approaching the issue from a multimodal perspective and with a special interest in the organization of participation in social activities, like Charles Goodwin (e.g., 1981, 1986c, 2000a, 2006) and Lorenza Mondada (e.g., 2002, 2006, 2007), others by approaching it from a more linguistic perspective, like Nick Enfield (e.g., 2001, 2005, 2009a), or from a gestural and communicative perspective, like Jürgen Streeck (e.g.,

1993, 1994, 2003, 2009) and finally by combining their contributions in edited volumes (e.g., Streeck, Goodwin, & LeBaron, 2011).

This dissertation has attempted to follow the path indicated by these scholars in terms of integrating visible and verbal behavior in the analysis of social interaction. Indeed, one of the central aims of this work has been to show how an analysis of talk and an analysis of gaze behavior can be combined, and how the analysis of one informs the analysis of the other. In addition to that, this dissertation has also shown some ways in which gaze behavior does "shape" or "alter" the accompanying verbal conduct. The practice of pursuing a response through gaze, or the practices concerning gaze behavior at possible sequence completion, for example, add to our basic knowledge of sequence organization. An understanding of action in interaction, of sequence organization and of the relevance of responses to sequence initiating actions is needed to capture what is going on in face-to-face interactions. Moreover, an understanding of the organization of gaze cannot and should not be developed by looking at eye movements in isolation. In this respect, while the goal of documenting as many interactional practices as possible is crucial for those interested in the organization of social interaction, it appears that a broader perspective on the role and organization of visible behavior and gaze in social interaction is necessary. With few exceptions, in particular, the work of Charles Goodwin (e.g., 1979; 1980; 1981) and Mardi Kidwell (e.g., 2005, 2009), most recent work on gaze in interaction has not approached it as a system in its own right, for example, by describing its 'grammar' in relation to other interactional systems. Indeed, most studies have either investigated gaze as just one component of a more complex range of visible resources (and therefore, as not deserving its own attention), or more disconcertingly, have reduced the interactional environment to some set of dichotomous exogenous variables (e.g., gender, power asymmetries, some aspect of the relationship between participants, etc.) and correlated them with particular gaze behaviors. By contrast, this dissertation shows the power of gaze in social interaction, accepting its membership in a vast array of visible resources that participants can rely on during social interactions, but also emphasizing the existence of a system, which can be described and appreciated once the interactional environment is investigated at a more fine-grained level of complexity.

324

## 5.3 Implications for a Model of Gaze Behavior in Social Interaction

One of the goals of this investigation has been to provide more precise descriptions of the mechanisms at play in participants' gaze behavior so that, for example, robots or avatars may be programmed to use their eyes in a more human-like fashion during interaction with human beings. The gaze practices reported in this dissertation are better understood in terms of their semiotics, rather than in terms of norms and rules. Here is how these practices work: if I do X, I mean to convey Y, if I do not do X, I mean to convey Z (see quote from Sacks, 1972 above). What this is meant to capture is that 'doing X' has one meaning and 'not doing X' has a different meaning, and the relevance and accountability of each behavior can be assessed taking into account the sequential environment of the talk. Practices are not usually accountable per se, but the actions they implement might have normative implications or their absence might be considered accountable. Consider, for example, the practice of using gaze to pressure for a response, presented in chapter 3. It is not the case that the participant who produced a FPP *must* look toward the recipient or is expected to do so. That is, the occurrence of that practice (i.e. using gaze to pressure for response) is not normative and other similar practices might be implemented instead (see Stivers & Rossano, 2010 on different ways of mobilizing response). However, that same participant might look toward the recipient in order to increase the likelihood of obtaining a response. By doing so, s/he might appear to put more pressure on the recipient, but there may be occasions where that might not the preferred thing to do. When speaker gaze toward the recipient occurs in specific sequential environments, however, it does increase pressure for a response, and, as such, it might have important interactional consequences. This means, for example, that while one does not have to use gaze to pursue a response, once gaze has been used to pursue it, a response might be normatively expected. The occurrence of a pursuit is what creates a normative expectation for a response. It is important to keep in mind that this is potentially different for other gaze practices, such as the one presented in chapter 2 (recipient gaze toward the speaker during the first TCU of an ETS). In that case, there may be a normative expectation that recipiency will be displayed through recipient gaze toward the speaker, rather than in some other way. Here, the occurrence of that specific practice for displaying recipiency, rather than alternative verbal one, might be normatively expected in that sequential position.

It is important to remember, here, that the effectiveness of a practice or the occurrence of behavioral regularity (e.g. the fact that most questions are produced with speaker gaze toward the recipient) should not, per se, be taken as evidence of the existence of a norm about it (see Bicchieri, 2006 for the difference between conventions and social norms). As Weber (1978 [1922]: 30) reminds us:

> *Many of the especially notable uniformities in the course of social action are not determined by orientation to any sort of norm which is held to be valid, nor do they rest on custom, but entirely on the fact that the corresponding type of social action is in the nature of the case best adapted to the normal interests of the actors as they themselves are aware of them.*

So, in using gaze to pursue a response, one might say that its deployment would depend on the interactional situation and whether the participants have an interest in pursuing a response without modifying the terms of the sequence initiating action. Alternatively, participants might interpret the delay in responding as due to a problem in hearing (e.g., they could repeat the sequence initiating turn) or prefiguring a dispreferred response (e.g., the speaker could produce a modified version of the sequence initiating turn where s/he might either back down or change the polarity of the initiating turn to allow for the production of a rejection as preferred response).[3]

Another important aspect to consider for the investigation of gaze in interaction concerns the semiotic resources that eye movements can use to accomplish specific communicative or regulatory functions. After all, the view of gaze as simply "on" or "off" (i.e., that participants either look at the other participant or not) is symptomatic of an inadequate understanding of how gaze works. Participants in interaction do not simply have the option of either looking at the other participants' faces or not looking at them. As the symbols used to represent gaze behavior in this dissertation show (see Appendix B), there are clearly more than two options. For example, looking at an object relevant for a competing activity (e.g., eating), is a different type of engagement than looking at a bare wall, just as looking middle distance, or looking upward (e.g., when somebody is thinking) is different from closing both eyes. In any of these cases, the orientation of the eyes with respect to the orientation of the head might affect the interpretation of a specific gaze behavior as well.

---

[3] For example, after a lack of response following "would you like to go to the cinema tonight?", the speaker might pursue a response with "or would you prefer another day?". A "no" in response to the first question can become a "yes" to the second one. As such, a dispreferred response to the first question can become a preferred response to the second question, produced as a pursuit.

Clearly, looks that are not directed toward somebody's face are not all equal in terms of their relevance and meaning for a recipient. Each look will be interpreted in terms of its relevance for the ongoing interaction and some looks will be more noticeable than others.

Aside from the complex ways of doing "not looking", two additional aspects of gaze contribute to their semiotic power: *timing* and *duration*. What this dissertation has shown is that, in order to understand the function of any particular gaze behavior, it is not enough to look at the composition of participants' gaze but it is in fact crucial to also examine its position, with respect to the participants' talk. Thus, it is not the case that, for example, *any* gaze withdrawal will work as a bid for closure nor that *any* speaker look toward the recipient will mobilize response, but that gaze can be used in those ways in some specific sequential environments. Similarly, the duration of a specific gaze behavior has important implications for its interpretation. For example, closing one's eyes for a second during a conversation might be understood as a way of displaying thinking; however, closing one's eyes for a longer duration (e.g., 5 or 6 seconds) might invite an alternative interpretation (e.g., the person may be not feeling well or has stopped listening). The same applies to participants' looks toward objects. Glancing at a glass before reaching for it might project that the object will be picked up, but a look toward that same glass for 5 or 6 seconds might suggest something else (e.g., recipient disengagement). In fact, there is a "natural duration"[4] of looks toward people, objects and empty spaces that relates to what else is going on in the conversation. Gaze behavior that extends beyond that natural duration, will be taken to mean something else. Accordingly, the particular meaning of a gaze behavior will also be affected by how long it lasts. It will also be affected by its placement in a course of action. Therefore, it is not enough to take into account whether a participant is looking at another one. Rather, it is crucial to map participants' gaze behaviors onto the social actions the participants are involved in. And, only at that point, can the specific meaning of that gaze behavior emerge.

A final core interest of this dissertation has been to try to specify gaze behavior with sufficient precision that one could, in principle, program an avatar to perform adequate gaze behavior—in practice, this can only be done within very limited domains at the present time (e.g. in the context of ETS, or during the TRP or at possible completion of a sequence), for reasons to be spelled out in what follows. There are (at least) four main constraints that could affect, for example, the occurrence and duration of mutual gaze, but would be difficult for

---

[4] See Morris (1957) on "typical intensity" of a signal from an ethological perspective.

machines to measure and compute. Yet, they play a clear role in the organization of gaze behavior in social interaction. They include:

1) A BIOLOGICAL CONSTRAINT: Mutual gaze can create arousal and be interpreted as flirting or threatening. The amount of arousal acceptable between individuals may vary and be affected by multiple variables.

2) A SOCIO-CULTURAL CONSTRAINT: Looking at the other participant may be used to display recipiency, commitment to the course of action and sustained attention. Moreover, gaze behavior might be an important semiotic resource for managing social relations.[5] Yet this might vary from activity to activity and from culture to culture.

3) AN INFORMATION PROCESSING CONSTRAINT: Looking at each other could be related to minimizing problems in hearing, speaking and understanding and to facilitate the information exchange (e.g., looking at lip movements to better understand speech, or looking at faces to read facial expressions associated with the content of the talk). Yet there might be individual differences in terms of people capacities to process information with and without the visual modality, in particular in individuals with cognitive impairments. Moreover, information processing during a conversation might at times require to look at the gestures of the other participant, rather than their face.

4) AN ATTENTION CONSTRAINT: There is a strong association between focus of attention and where people direct their eyes, so looking away from a speaker can suggest shift in a focus of attention or momentary lack of attention.

Clearly, there is a relationship between these constraints and they might have a cascading effect on one another. The possibility of exceeding arousal (biological constraint) might lead people to look away from time to time. However, at least in some cultures, looking is a way of displaying sustained engagement (social constraint), so looking away should be limited to things that are not *cognitively demanding*. For example, looking at

---

[5] As Enfield (2009b: 60) suggests "human pragmatics is about using semiotic resources to try to meet goals in the realm of social relationships". As such, gaze might be one of these resources and yet the ways in which it might help managing social relationships might be culturally specific.

objects in the environment (e.g., pens, glasses, food) would be interactionally acceptable as long as such objects are not too cognitively demanding. However, engagement with these objects might be related to *competing activities*, and focusing on them as such, might mean diminished attention (attention constraint) to the ongoing interaction. It might also mean potentially reduced capacity in fully processing the information obtainable by looking at the other participant's face (information processing constraint). At the same time, it is important to note that people might look at each other to perceive emotions and/or take a particular stance toward something rather than to parse words or better process information. Indeed, it is possible that the processing of emotions plays a major role in the development of the social brain and in the processing of the face (e.g., Adolphs, 2009; Grossmann & Johnson, 2007; Grossmann, Johnson, Farroni, & Csibra, 2007; Haxby, Hoffman, & Gobbini, 2000; Spezio, Huang, Castelli, & Adolphs, 2007). Finally, when one thinks of information processing during a conversation, the main focus is usually on information provided through lip movements or facial expressions. Yet we know that much information is provided through speech-accompanying gestures and their uptake might at times require fixating on the gestures themselves. In general, we know that interactants fixate on very few gestures, but also that the likelihood of an addressee fixating a gesture increases under the following three circumstances:

1. When speakers first look at their own gestures (Gullberg & Holmqvist, 1999, 2006; Gullberg & Kita, 2009);

2. When a gesture is produced in the periphery of gesture space in front of the speaker's body (Gullberg & Holmqvist, 1999, 2006);[6]

3. When a gestural movement is suspended momentarily in mid-air and goes into a hold (Gullberg & Holmqvist, 1999, 2006; Gullberg & Kita, 2009).

These variables, then, have to be compared with the actual effect on uptake of the information conveyed by the gestures. Indeed, while addressees' fixation of gestures produced by speakers is affected by the above variable, it appears that the only one that actually affect uptake of gestural information is speakers previous fixation of their own

---

[6] Though see Gullberg & Kita, 2009 for evidence that suggests otherwise. Their criticism is that the claim at point 2 has been made in situations in which the variables 1-3 were often conflated.

gestures. So direct fixation of a gesture does not necessarily entails better uptake of the information that it is conveying.

To summarize, an optimal account of gaze behavior should not only account for when people look at each other and when they do not, but also for what they are doing when they are (not) looking at each other and why they might do that. And, an optimal description of the gaze machinery needs to take into account not only what people usually do, but the biological, cognitive and social constraints under which they do it. Thus, I have suggested an approach that requires a basic understanding of four different levels of organization (see § 1.1):

1. Individual

2. Micro-social (Dialogical/Conversational)

3. Macro-social/Cultural

4. Species-specific

In order to be able to completely understand what humans do when they interact with each other, these four different levels of organization are a reminder that, in any interaction:

1. We have individual cognition and individual memories of our previous interactions.

2. We use the interactional machinery we believe we share with other interactants and we recipient-design our communications and actions, because of what we know about others.

3. Our actions are informed by what we deem socially and culturally appropriate.

4. We are biological organisms, with specific limits and capacities and our bodies have evolved for specific adaptive purposes.

The findings reported in this dissertation, the methods deployed to investigate gaze behavior and, more generally, the interest in a broader approach to gaze in interaction will hopefully constitute a contribution to current research on the organization of social interaction and the practices deployed to sustain it.

# Appendix A: Information about the Data

*Interactions used for qualitative and quantitative purposes*

1) **2GC**

**PARTICIPANTS:** 2

**GENDER:** both females

**AGE:** In their early 20s

**RELATIONSHIP:** Friends

**LOCATION:** A's house (kitchen)

**SEATING ARRANGEMENTS:** in front of each other, sitting at a dinner table


2) **2GCCOLL**

**PARTICIPANTS:** 2

**GENDER:** both females

**AGE:** In their early 20s

**RELATIONSHIP:** housemates and friends

**LOCATION:** in the kitchen of the house they share

**SEATING ARRANGEMENTS:** in front of each other, sitting at the dinner table


3) **2GGOSS**

**PARTICIPANTS:** 2

**GENDER:** both females

**AGE:** In their early 20s

**RELATIONSHIP:** Friends

**LOCATION:** A's house (kitchen)

**SEATING ARRANGEMENTS:** 90 degrees, sitting at table

4) **2GSOFA**

**PARTICIPANTS:** 2

**GENDER:** both females

**AGE:** In their early 20s

**RELATIONSHIP:** Friends

**LOCATION:** a friend's living room.

**SEATING ARRANGEMENTS:** 90 degrees, sitting on two couches with a coffee table near both of them


5) **2GSTUDYING**

**PARTICIPANTS:** 2

**GENDER:** both females

**AGE:** In their early 20s

**RELATIONSHIP:** Friends

**LOCATION:** A's bedroom in A's house

**SEATING ARRANGEMENTS:** A is sitting on a bed facing B, who is sitting at a 90 degree corner at a desk, but often turns chair to face A.


6) **2PCOMP**

**PARTICIPANTS:** 2

**GENDER:** both males

**AGE:** In their early 20s

**RELATIONSHIP:** Friends

**LOCATION:** A's bedroom in A's house

**SEATING ARRANGEMENTS:** side-by-side at a desk, both facing a computer


7) **2PERLINE**

**PARTICIPANTS:** 2

**GENDER:** both females

**AGE:** A is in her 30s, B is 50

**RELATIONSHIP:** friends and colleagues


334

**LOCATION:** in B's house (kitchen)

**SEATING ARRANGEMENTS:** in front of each other, sitting at the dinner table


8) **2PEXAM**

**PARTICIPANTS:** 2

**GENDER:** both males

**AGE:** In their early 20s

**RELATIONSHIP:** Friends

**LOCATION:** B's bedroom in B's house

**SEATING ARRANGEMENTS:** A is sitting on a chair, while B is sitting at a 90 degree angle, in front of a desk,  looking at a computer, but often turns chair to face A.


9) **2PPLAN**

**PARTICIPANTS:** 2

**GENDER:** both males

**AGE:** In their early 20s

**RELATIONSHIP:** Friends

**LOCATION:** A's living room in A's house

**SEATING ARRANGEMENTS:** in front of each other, on 2 couches


10) **2PLUNCH**

**PARTICIPANTS:** 2

**GENDER:** A is male, B is female

**AGE:** In their early 20s

**RELATIONSHIP:** Couple

**LOCATION:** B's house (kitchen)

**SEATING ARRANGEMENTS:** 90 degrees, sitting at table

*Interactions used only for qualitative purposes*


11) **2PRON1**

**PARTICIPANTS:** 2

**GENDER:** both males

**AGE:** In their early 20s

**RELATIONSHIP:** Housemates

**LOCATION:** in the kitchen of the apartment they are sharing

**SEATING ARRANGEMENTS:** in front of each other at the dinner table


12) **3PDRIVING**

**PARTICIPANTS:** 3

**GENDER:** all males

**AGE:** In their early 20s

**RELATIONSHIP:** Friends

**LOCATION:** in A's car

**SEATING ARRANGEMENTS:** A is driving, C is in the passenger seat and B is sitting the back seats, leaning forward in between A and C


13) **3PMARIA**

**PARTICIPANTS:** 3

**GENDER:** two females and a male

**AGE:** In their early 20s

**RELATIONSHIP:** Housemates

**LOCATION:** in the kitchen of the apartment they are sharing

**SEATING ARRANGEMENTS:** two women in front of each other at the dinner table, man at the head of the table at a 90 degree angle from both of them

## Appendix B: Transcription Conventions

### *Transcription Conventions for the Talk*

The symbols used in the transcriptions in this article follow the notation system developed by Gail Jefferson (see, for example, Jefferson 2004a) for conversation analytic research.

1. Temporal and sequential relationships

| | |
|---|---|
| [ | A left bracket indicates the onset of overlapping speech |
| ] | A right bracket indicates the point at which overlapping utterances end |
| = | An equals sign indicates contiguous speech |
| (0.5) | Silences are indicated as pauses in tenths of a second |
| (.) | A period in parentheses indicates a micro-pause (less than two tenths of a second) |

2. Aspects of speech delivery

| | |
|---|---|
| . | A period indicates a falling intonation contour |
| , | A comma indicates continuing intonation |
| ? | A question mark indicates rising intonation contour |
| ¿ | An inverted question mark indicates a rise stronger than the comma but weaker than the question mark |
| _ | An underscore indicates flat intonation contour |
| : | Colons indicate lengthening of preceding sound (the more colons, the longer the lengthening) |
| - | A hyphen indicates an abrupt cutoff sound |
| yes | Underlining indicates emphatic stress |
| YES | Upper case indicates noticeably increased amplitude or pitch reset |
| °yes° | The degree sign indicates noticeably decreased amplitude in speech |
| >yes< | Indicates talk that is noticeably faster than surrounding talk |
| <yes> | Indicate talk that is noticeably slower than surrounding talk |

| hh | The letter 'h' indicates audible aspirations (the more hs the longer the breath) |
|---|---|
| .hh | A period preceding the letter 'h' indicates audible inhalations (the more hs the longer the breath) |
| y(h)es | h within parentheses within a word indicates aspiration, possibly laughter |

3. Other notational devices

| (guess) | words within single parentheses indicate likely hearing of that word |
|---|---|
| ((coughs)) | information in double parentheses indicate additional details |
| (     ) | empty parentheses indicate hearable yet indecipherable talk |

### *Multi-linear transcription conventions (see Sidnell, 2009: xvii)*

The conversations reported in this dissertation are in Italian. To facilitate comprehension, I used three-line transcripts where the first line is a broad phonetic representation of the talk in the original language and follows the notations reported above. The second line is a morpheme-by-morpheme gloss using a combination of word-by-word translation and abbreviations to indicate particles and other functional items that do not admit of a direct translation into English. The third line presents an idiomatic English gloss, meant to capture, as close as possible, the contextual meaning of the utterance.

## *Symbols for Gaze Orientations*

$\overset{\text{A}}{\bigcirc}\longleftrightarrow\overset{\text{B}}{\bigcirc}$ = Mutual gaze

$\overset{\text{A}}{\bigcirc}\quad\overset{\text{B}}{\bigcirc}$ = A looks away and B looks away

$\overset{\text{A}}{\bigcirc}\quad\overset{\text{B}}{\bigcirc}$ = A looks down oriented towards B. B looks away

$\overset{\text{A}}{\bigcirc}\quad\overset{\text{B}}{\bigcirc}$ = A looks away. B looks down oriented towards A

$\overset{\text{A}}{\bigcirc}\overset{\text{B}}{\bigcirc}$ = A and B are looking down in front of them

$\overset{\text{A}}{\bigcirc}\longrightarrow\overset{\text{B}}{\bigcirc}$ = A looks at B. B looks down

$\overset{\text{A}}{\bigcirc}\longrightarrow\overset{\text{B}}{\bigcirc}$ = A looks at B. B looks away

$\overset{\text{A}}{\bigcirc}\longleftarrow\overset{\text{B}}{\bigcirc}$ = B looks at A. A looks down

$\overset{\text{A}}{\bigcirc}\longleftarrow\overset{\text{B}}{\bigcirc}$ = B looks at A. A looks away

$\overset{\text{A}}{\bigcirc}\quad\overset{\text{B}}{\bigcirc}$ = A looks down. B eyes closed

$\overset{\text{A}}{\bigcirc}\quad\overset{\text{B}}{\bigcirc}$ = A eyes closed. B looks down
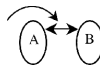
$\overset{\text{A}}{\bigcirc}\quad\overset{\text{B}}{\bigcirc}$ = A eyes closed. B eyes closed

$\overset{\text{A}}{\bigcirc}\longrightarrow\overset{\text{B}}{\bigcirc}$ = A looks at B. B eyes closed

$\overset{\text{A}}{\bigcirc}\longleftarrow\overset{\text{B}}{\bigcirc}$ = A eyes closed. B looks at A

$\overset{\text{A}}{\bigcirc}\quad\overset{\text{B}}{\bigcirc}$ = A away. B eyes closed

 = A turns towards B who is looking down

 = A turns towards B who is already looking at A

 = A raises gaze towards B who is looking down
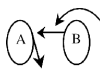
 = A raises gaze towards B who is already looking at A

 = A looks away B looks mid distance up left
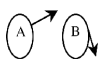
 = A looks away B looks mid distance up right
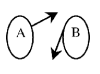
 = A looks down B looks mid distance up left

 = A looks down B looks mid distance up right

 = B turns towards A who is looking down

 = A raises gaze towards B who is looking down

 =A looks mid distance up left. B is looking away

 = A looks mid distance up left. B is looking down

 = A looks mid distance up right. B is looking away

 = A looks mid distance up right. B is looking down

 = A turns towards B who is looking away

340

 = B turns towards A who is already looking at B

 = B raises gaze towards A who is looking away

 = B turns towards A who is looking away

 = B raises gaze towards A who is already looking at B

 = B raises gaze towards A who is looking down

 = A looking at B. B looks mid distance up right

 = A mid distance right. B eyes closed

 = A mid distance right. B looking at A

 = A mid distance left. B looking at A

 = A looking at B. B mid distance right

 = A looking at B. B mid distance left

 = A looking at object. B looking away

 = A looking at object. B looking down

 = A looking at object. B eyes closed

 = A looking at object. B looking at A

 = A looking at object. B raises gaze towards A

 = A looking at object. B mid distance left

 = A looking at object. B mid distance right

 = A looking at object. B turns gaze towards A

 = A looking at one object. B looking at a different object

 = A looking away. B looking at object

 = A looking down. B looking at object

 = A eyes closed. B looking at object

 = A looking at B. B looking at object

 = A raises gaze towards B. B looking at object

 = A mid distance left. B looking at object

 = A mid distance right. B looking at object

 = A turns gaze towards B. B looking at object

 = A and B looking at the same object

# References

Adolphs, R. (2009). The social brain: neural basis of social knowledge. *Annual Review of Psychology, 60*, 693-716.

Alexander, R. D. (1987). *The biology of moral systems*. New York, NY: Aldine de Gruyter.

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Science, 4*, 267-278.

Allman, J. (1977). Evolution of the visual system in the early primates. *Progress in Psychobiology and Physiological Psychology, 7*, 1-53.

Argyle, M., & Cook, M. (1976). *Gaze and Mutual Gaze*. Cambridge: Cambridge University Press.

Argyle, M., & Dean, J. (1965). Eye-Contact, Distance and Affiliation. *Sociometry, 28*, 289-304.

Argyle, M., & Graham, J. A. (1976). The Central Europe experiment: Looking at persons and looking at things. *Journal of Environmental Psychology and Nonverbal Behavior, 1*(1), 6-16.

Aslin, R. N. (2007). What's in a look? *Developmental Science, 10*(1), 48-53.

Atkinson, J. M., & Heritage, J. (Eds.). (1984). *Structures of Social Action: Studies in Conversation Analysis*. Cambridge: Cambridge University Press.

Auer, P. (1984). Referential Problems in Conversation. *Journal of Pragmatics, 8*(5-6), 627-648.

Auer, P. (2005). Projection in interaction and projection in grammar. *Text, 25*(1), 7-36.

Austin, J. L. (1962). *How to Do Things with Words: Second Edition*. Oxford: Clarendon.

Bard, K. A., Myowa-Yamakoshi, M., Tomonaga, M., Tanaka, M., Costall, A., & Matsuzawa, T. (2005). Group differences in the mutual gaze of chimpanzees (Pan Troglodytes). *Developmental Psychology, 41*(4), 616_624.

Baron-Cohen, S. (1995). *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge MA.: The MIT Press.

Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology letters, 2*, 412-414.

Bavelas, J. B., Coates, L., & Johnson, T. (2000). Listeners as co-narrators. *Journal of Personality and Social Psychology, 79*, 941-952.

Bavelas, J. B., Coates, L., & Johnson, T. (2002). Listener Responses as a Collaborative Process: The Role of Gaze. *Journal of Communication, 52*, 566-580.

Beach, W. A. (1995). Preserving and constraining options: "okays" and 'official' priorities in medical interviews. In B. Morris & R. Chenail (Eds.), *Talk of the Clinic*. Hillsdale NJ: Erlbaum.

Beattie, G. W. (1978). Floor Apportionment and Gaze in Conversational Dyads. *British Journal of Social and Clinical Psychology, 17*, 7-16.

Beattie, G. W. (1979). Planning Units in Spontaneous Speech: Some Evidence from Hesitations in Speech and Speaker Gaze Direction in Conversation. *Linguistics, 17*, 61-78.

Bente, G., Donaghy, W. C., & Suwelack, D. (1998). Sex differences in body movement and visual attention: an integrated analysis of movement and gaze in mixed sex dyas. *Journal of Nonverbal Behavior, 22*(1), 31-58.

Bicchieri, C. (2006). *The Grammar of Society: the Nature and Dynamics of Social Norms*. Cambridge: Cambridge University Press.

Birdwhistell, R. L. (1970). *Kinesics and Context: Essays on Body Motion Communication*. New York: Ballentine Books.

Boll, H. (1994 [1963]). *The Clown*: Penguin Classics.

Bowmaker, J. K. (1991). The evolution of vertebrate visual pigments and its implications for colour deficiencies. In J. R. Cronly-Dillon & R. L. Gregory (Eds.), *Vision and visual dysfunction* (Vol. 2, pp. 63-81). London: Macmillan.

Brazelton, T. B., Koslowski, B., & Main, M. (1974). The origins of reciprocity: The early mother-intant interaction. In M. Lewis & L. A. Rosenblum (Eds.), *The effect of the infant on its caregivers* (pp. 137-154). New York: Wiley.

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science, 8*(6), 535-543.

Brothers, L. (1990). The social brain: a project for integrating primate behavior and neuropsychology in a new domain. *Concepts in Neuroscience, 1*, 27-51.

Brown, P., & Levinson, S. C. (1987). *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.

Bruce, V., Green, P. R., & Georgeson, M. A. (2003). *Visual Perception: physiology, psychology and ecology*. New York: Psychology Press.

Bruner, J. (1983). *Child's talk*. New York: Norton.

Butterworth, G., & Cochran, E. (1980). Towards a Mechanism of Joint Visual Attention in Human Infancy. *International Journal of Behavioral Development, 3*, 253-272.

Byrne, R. (2006). Parsing Behavior: A Mundane Origin for an Extraordinary Ability? In N. J. Enfield & S. C. Levinson (Eds.), *Roots of Human Sociality* (pp. 478-505). New York: Berg.

Calder, A. J., Beaver, J. D., Winston, J. S., Dolan, R. J., Jenkins, R., Eger, E., et al. (2007). Separate Coding of Different Gaze Direction in the Superior Temporal Sulcus and Inferior Parietal Lobule. *Current Biology, 17*, 20-25.

Carpenter, M., & Tomasello, M. (2000). Joint attention, cultural learning, and language acquisition: Implications for children with autism. In A. M. Wetherby & B. M. Prizant (Eds.), *Communication and language issues in autism and pervasive developmental disorder: A transactional developmental perspective* (pp. 31-54). Baltimore, MD: Brookes.

Cartmill, M. (1974). Rethinking primate origins. *Science, 184*, 436-443.

Cartmill, M. (1992). New Views on Primate Origins. *Evolutionary Anthropology, 1*(3), 105-111.

Charman, T. (2004). Why is joint attention a pivotal skill in autism? *Philosophical Transactions of the Royal Society of London, 358*, 315-324.

Clark, H. (1996). *Using Language*. Cambridge: Cambridge University Press.

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

Clayman, S. E., Heritage, J., Elliott, M. N., & McDonald, L. (2007). When Does the Watchdog Bark? Conditions of Aggressive Questioning in Presidential News Conferences. *American Sociological Review, 72*, 23-41.

Cordell, D. M., & McGahan, J. R. (2004). Mutual Gaze Duration as a Function of Length of Conversation in Male-Female Dyads. *Psychological Reports, 94*, 109-114.

Couper-Kuhlen , E. (2004). Prosody and sequence organization in English conversation: The case of new beginnings. In E. Couper-Kuhlen  & C. Ford (Eds.), *Sound Patterns in Interaction* (pp. 335-376). Amsterdam: John Benjamins.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*(4), 148-153.

Curl, T., & Drew, P. (2008). Contingency and action: A comparison of two forms of requesting. *Research on Language and Social Interaction, 41*(2), 129-153.

D'Entremont, B., Hains, S. M. J., & Muir, D. W. (1997). A demonstration of gaze following in 3- to 6- month-olds. *Infant Behavior and Development, 20*(4), 569-572.

Darwin, C. (1859). *On the origins of species by means of natural selection*. London: John Murray.

Darwin, C. (1872). *The expression of the emotions in man and animals*. London: John Murray.

Davidson, J. (1984). Subsequent Versions of Invitations, Offers, Requests, and Proposals Dealing with Potential or Actual Rejection. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action* (pp. 102-128). Cambridge: Cambridge University Press.

Davidson, J. A. (1990). Modifications of Invitations, Offers and Rejections. In G. Psathas (Ed.), *Interaction Competence* (pp. 149-180). Washington: International Institute for Ethnomethodology and Conversation Analysis and University Press of America.

Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., et al. (2004). Early social attention impairments in autism: social orienting, joint attention and attention to distress. *Developmental Psychology, 40*(2), 271-283.

De Ruiter, J. P. (2005). The role of eye-gaze in visual dialogue tasks. *AMLaP*.

De Ruiter, J. P., Mitterer, H., & Enfield, N. J. (2006). Projecting the End of a Speaker's Turn: a Cognitive Cornerstone of Conversation. *Language, 82*(3), 504-524.

de Waal, F. (2001). *The ape and the sushi master: Cultural reflections by a primatologist*. London: Penguin Press.

Drew, P. (1984). Speakers' Reportings in Invitation Sequences. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action* (pp. 152-164). Cambridge: Cambridge University Press.

Drew, P. (1987). Po-Faced Receipts of Teases. *Linguistics, 25*, 219-253.

Drew, P. (1995). Interaction Sequences and Anticipatory Interactive Planning. In E. N. Goody (Ed.), *Social Intelligence and Interaction: Expressions and Implications of the Social Bias in Human Intelligence* (pp. 111-138). Cambrdige: Cambridge University Press.

Drew, P. (1997). 'Open' class repair initiators in response to sequential sources of trouble in conversation. *Journal of Pragmatics, 28*, 69-101.

Drew, P. (1998). Complaints about transgressions and misconduct. *Research on Language and Social Interaction, 31*(3-4), 295-325.

Drew, P. (2005). Conversation Analysis. In K. L. Fitch & R. E. Sanders (Eds.), *Handbook of Language and Social Interaction* (pp. 71-102). Mahwah, NJ: Erlbaum.

Drew, P., & Holt, E. (1988). Complainable Matters: The Use of Idiomatic Expressions in Making Complaints. *Social Problems, 35(4)*, 398-417.

Drew, P., & Holt, E. (1998). Figures of Speech: Figurative Expressions and the Management of Topic Transition in Conversation. *Language in Society, 27*(4), 495-523.

Drew, P., & Walker, T. (2009). Going too far: Complaining, escalating and disaffiliation. *Journal of Pragmatics, 41*(12), 2400-2414.

Drew, P., & Walker, T. (2010). Requesting assistance in calls to the police. In M. Coulthard & A. Johnson (Eds.), *Handbook of Forensic Linguistics* (pp. 95-110): Routledge.

Dunbar, R. (1998). The social brain hypothesis. *Evolutionary Anthropology, 6*, 178-190.

Duncan, S., Jr. (1975). Interaction units during speaking turns in dyadic, face-to-face conversations. *Organization of Behaviour in Face-to-face Interaction*, 199-212.

Duncan, S., Jr., & Fiske, D. W. (1977). *Face-to-Face Interaction: Research, Methods, and Theory*. New York: Wiley.

Duncan, S., Jr., & Niederehe, G. (1974). On Signalling that it's your turn to speak. *Journal of Experimental Social Psychology, 10*, 234-247.

Edwards, D. (2000). Extreme case formulations: Softeners, investment, and doing nonliteral. *Research on Language and Social Interaction, 33*(4), 347-373.

Egbert, M. (1996). Context sensitivity in conversation analysis: eye gaze and the German repair initiator "bitte". *Language in Society, 25*(4), 587-612.

Ehardt, C. L., & Blount, B. G. (1984). Mother-infant visual interaction in Japanese macaques. *Developmental Psychobiology, 17*, 391-405.

Eibl-Eibesfeldt, I. (1989). *Human Ethology*. New York: Aldine de Gruyter.

Ekman, P. (1992). Facial expressions of emotions: New findings, new questions. *Psychological Science, 3*(1), 34-38.

Ekman, P. (1993). Facial expression and emotion. *American Psychologist, 48*(4), 376-379.

Ekman, P., & Friesen, W. V. (1974). Nonverbal Leakage and Clues to Deception. In S. Weitz (Ed.), *Nonverbal Communication* (pp. 269-290). New York: Oxford University Press.

Ekman, P., & Friesen, W. V. (1978). *The facial action coding system: A technique for the measurement of facial movement*. Palo Alto: The Consulting Psychologists Press.

Ekman, P., & Oster, H. (1979). Facial expressions of emotion. *Annual Review of Psychology, 30*, 527-554.

Ekström, M. (2011). *Coordinating talk and gaze in political media interviews*. Paper presented at the ICA Conference.

Ellsworth, P. C., & Ludwig, L. M. (1972). Visual Behavior in Social Interaction. *The Journal of Communication, 22*(4), 375-403.

Emery, N. J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews, 24*, 581-604.

Enfield, N. J. (2001). 'Lip-pointing': A discussion of form and function with reference to data from Laos. *Gesture, 1*(2), 185-212.

Enfield, N. J. (2005). The body as a cognitive artifact in kinship representations: Hand gesture diagrams by speakers of Lao. *Current Anthropology, 46*(1), 1-26.

Enfield, N. J. (2009a). *The anatomy of meaning: Speech, gesture, and composite utterances*. Cambridge: Cambridge University Press.

Enfield, N. J. (2009b). Relationship thinking and human pragmatics. *Journal of Pragmatics, 41*, 60-78.

Enfield, N. J. (2011). Sources of asymmetry in human interaction: Enchrony, status, knowledge and agency. In T. Stivers, L. Mondada, & J. Steensig (Eds.), *The morality of knowledge in conversation* (pp. 285-312). Cambridge: Cambridge University Press.

Enfield, N. J., & Levinson, S. C. (Eds.). (2006a). *Roots of human sociality: culture, cognition and interaction*. Oxford: Berg.

Enfield, N. J., & Levinson, S. C. (2006b). Introduction: Human sociality as a new interdisciplinary field. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of human sociality: Culture, cognition and interaction* (pp. 1-35). Oxford: Berg.

Erickson, F. (1979). Talking Down: Some Cultural Sources of Miscommunication in Interracial Interviews. In A. Wolfgang (Ed.), *Nonverbal Communication: Applications and Cultural Implications* (pp. 99-126). New York: Academic Press.

Exline, R. V. (1963). Explorations in the Process of Person Perception: Visual Interaction in Relation to Competition, Sex, and Need for Affiliation. *Journal of Personality, 31*, 1-20.

Fantz, R. L. (1963). Pattern Vision in Newborn Infants. *Science, 140*, 296-297.

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Science (PNAS), 99*, 9602-9605.

Farroni, T., Menon, E., & Johnson, M. H. (2006). Factors influencing newborns' preference for faces with eye contact. *Journal of Experimental Child Psychology, 95*, 298-308.

Farroni, T., Pividori, D., Simion, F., Massaccesi, S., & Johnson, M. H. (2004). Eye gaze cueing of attention in newborns. *Infancy, 5*, 39-60.

Fernald, R. D. (2000). Evolution of eyes. *Current Opinion in Neurobiology, 10*(4), 444-450.

Fernald, R. D. (2006). Casting a genetic light on the evolution of eyes. *Science, 313*, 1914-1918.

Field, T. M. (1979). Visual and cardiac responses to animate and inanimate faces by young term and preterm infants. *Child Development, 50*, 188-194.

Field, T. M. (1981). Infant gaze aversion and heart rate during face-to-face interactions. *Infant Behaviour & Development, 4*, 307-315.

Ford, C. E. (2004). Contingency and units in interaction. *Discourse Studies, 6*(1), 27-52.

Ford, C. E., Fox, B. A., & Thompson, S. A. (2002). Constituency and the Grammar of Turn Increments. In C. E. Ford, B. A. Fox & S. A. Thompson (Eds.), *The Language of Turn and Sequence* (pp. 14-38). Oxford: Oxford University Press.

Ford, C. E., & Thompson, S. A. (1996). Interactional Units in Conversation: Syntactic, Intonational, and Pragmatic Resources for the Management of Turns. In E. Ochs, E. A. Schegloff & S. A. Thompson (Eds.), *Interaction in Grammar* (pp. 134-184). Cambridge: Cambridge University Press.

Foucault, M. (1979). *Discipline and Punish: The Birth of the Prison*. New York: Random House.

Fox, B., Wouk, F., Hayashi, M., Sorjonen, M., Laakso, M.-J., Fincke, S., et al. (2009). A cross-linguistic investigation of the site of initiation in same-turn self-repair. In J. Sidnell (Ed.), *Conversation Analysis: Comparative Perspectives* (pp. 60-103). Cambridge: Cambridge University Press.

Frith, U., & Frith, C. (2001). The biological basis of social interaction. *Current Directions in Psychological Science, 10*(5), 151-155.

Gardner, R. (2001). *When Listeners Talk*. Amsterdam/Philadelphia: John Benjamins.

Garfinkel, H. (1967). *Studies in Ethnomethodology*. Englewood Cliffs, N.J.: Prentice-Hall.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

Gibson, J. J., & Pick, A. D. (1963). Perception of Another Person's Looking Behavior. *American Journal of Psychology, 76*(3), 386-394.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the Fast and Frugal Way: Models of Bounded Rationality. *Psychological Review, 103*(4), 650-669.

Goffman, E. (1955). On face work. *Psychiatry, 18*, 213-231.

Goffman, E. (1959). *The Presentation of Self in Everyday Life*. Garden City NY: Doubleday.

Goffman, E. (1963). Behavior in Public Places: Notes on the Social Organization of Gatherings.

Goffman, E. (1964). The Neglected Situation. In The Ethnography of Communication. John J. Gumperz and Dell Hymes, eds. *American Anthropologist, 66, 6, pt. II*, 133-136.

Goffman, E. (1967). *Interaction Ritual: Essays in Face to Face Behavior*. Garden City, New York: Doubleday.

Goffman, E. (1971). *Relations in Public: Microstudies of the Public Order*. New York: Harper and Row.

Goffman, E. (1981a). Footing. *Forms of Talk*, 124-159.

Goffman, E. (1981b). *Forms of Talk*. Philadelphia: University of Pennsylvania Press.

Goffman, E. (1981c). Response cries. *Forms of Talk*, 78-123.

Golato, A. (2005). *Compliments and compliment responses: Grammatical structure and sequential organization*. Amsterdam: John Benjamins.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of Ecological Rationality: The Recognition Heuristic. *Psychological Review, 109*(1), 75-90.

Goodenough, J., McGuire, B., & Wallace, R. (1993). *Perspectives on Animal Behavior*. New York: Wiley.

Goodwin, C. (1979). The Interactive Construction of a Sentence in Natural Conversation. In G. Psathas (Ed.), *Everyday Language: Studies in Ethnomethodology* (pp. 97-121). New York: Irvington Publishers.

Goodwin, C. (1980). Restarts, Pauses, and the Achievement of Mutual Gaze at Turn-Beginning. *Sociological Inquiry, 50*(3-4), 272-302.

Goodwin, C. (1981). *Conversational Organization: Interaction Between Speakers and Hearers*. New York: Academic Press.

350

Goodwin, C. (1984). Notes on Story Structure and the Organization of Participation. *Structures of Social Action*, 225-246.

Goodwin, C. (1986a). Audience Diversity, Participation and Interpretation. *Text, 6(3)*, 283-316.

Goodwin, C. (1986b). Between and Within: Alternative Treatments of Continuers and Assessments. *Human Studies, 9*, 205-217.

Goodwin, C. (1986c). Gesture as a Resource for the Organization of Mutual Orientation. *Semiotica, 62(1/2)*, 29-49.

Goodwin, C. (1987). Forgetfulness as an Interactive Resource. *Social Psychology Quarterly, 50, No.2*, 115-130.

Goodwin, C. (2000a). Action and Embodiment Within Situated Human Interaction. *Journal of Pragmatics, 32*(1489-1522).

Goodwin, C. (2000b). Practices of Seeing, Visual Analysis: An Ethnomethodological Approach. In T. van Leeuwen & C. Jewitt (Eds.), *Handbook of Visual Analysis* (pp. 157-182). London: Sage.

Goodwin, C. (2006). Human Sociality as Mutual Orientation in a Rich Interactive Environment: Multimodal Utterances and Pointing in Aphasia. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of human sociality: culture, cognition and interaction* (pp. 97-125). New York: Berg.

Goodwin, C., & Goodwin, M. H. (1987). Concurrent Operations on Talk: Notes on the Interactive Organization of Assessments. *IPrA Papers in Pragmatics, 1*(1), 1-52.

Goodwin, C., & Goodwin, M. H. (1992). Assessments and the Construction of Context. In A. Duranti & C. Goodwin (Eds.), *Rethinking Context: Language as an Interactive Phenomenon* (pp. 147-190). Cambridge: Cambridge University Press.

Goodwin, M. H. (1980). Processes of Mutual Monitoring Implicated in the Production of Description Sequences. *Sociological Inquiry, 50*, 303-317.

Goodwin, M. H. (1982). 'Instigating': Storytelling as a Social Process. *American Ethnologist, 9*, 799-819.

Goodwin, M. H., & Goodwin, C. (1986). Gesture and Coparticipation in the Activity of Searching for a Word. *Semiotica, 62(1/2)*, 51-75.

Gordon, R. A., & Donzis, P. B. (1985). Refractive Development of the Human Eye. *Archives of Ophthalmology, 103*(6), 785-789.

Gould, S. J. (1997). The exaptive excellence of spandrels as a term and prototype. *Proceedings of the National Academy of Sciences (PNAS), 94*, 10750-10755.

Gould, S. J., & Lewontin, R. C. (1979). The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proceedings of the Royal Society of London Series B-Biological Sciences, 205*, 581-598.

Gregory, R. L. (1972). *Eye and brain (2nd Ed)*. London: Weidenfeld and Nicolson.

Gregory, R. L. (1980). Perceptions as hypotheses. *Philosophical Transactions of the Royal Society of London, Series B, 290*, 181-197.

Grice, H. P. (1975). Logic and Conversation. In P. Cole & N. L. Morgan (Eds.), *Syntax and Semantics, Vol. 3, Speech Acts* (pp. 41-58). New York: Academic Press.

Griffin, Z. M. (2004). Why look? Reasons for eye movements related to language production. In J. M. Henderson & F. Ferreira (Eds.), *The integration of language, vision and action: Eye movements and the visual world*. New York: Psychology Press.

Grossmann, T., & Farroni, T. (2009). Decoding social signals in the infant brain: A look at eye gaze perception. In M. de Haan & M. Gunnar (Eds.), *Handbook of Developmental Social Neuroscience* (pp. 87-106): Guilford Publications.

Grossmann, T., & Johnson, M. H. (2007). The development of the social brain in human infancy. *European Journal of Neuroscience, 25*, 909-919.

Grossmann, T., Johnson, M. H., Farroni, T., & Csibra, G. (2007). Social perception in the infant brain: gamma oscillatory activity in response to eye gaze. *SCAN, 2*, 284-291.

Gullberg, M., & Holmqvist, K. (1999). Keeping an eye on gestures: Visual perception of gestures in face-to-face communication. *Pragmatics & Cognition, 7*, 35-63.

Gullberg, M., & Holmqvist, K. (2006). What speakers do and what listeners look at. Visual attention to gestures in human interaction live and on video. *Pragmatics and Cognition, 14*(1), 53-82.

Gullberg, M., & Kita, S. (2009). Attention to speech-accompanying gestures: Eye movements and information uptake. *Journal of Nonverbal Behavior, 33*(4), 251-277.

Haddington, P. (2006). The organization of gaze and assessments as resources for stance taking. *Text & Talk, 26*(3), 281-328.

Halliday, M. A. K. (1967). *Intonation and Grammar in British English*. The Hague: Mouton.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Science, 4*, 223-233.

Hayhoe, M. (2000). Vision using routines: a functional account of vision. *Visual Cognition, 7*, 43-64.

Hayhoe, M., & Ballard, D. (2005). Eye Movements in Natural Behavior. *Trends in Cognitive Science, 9*(4), 188-194.

Heath, C. (1984). Talk and Recipiency: Sequential Organization in Speech and Body Movement. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action* (pp. 247-265). Cambridge: Cambridge University Press.

Heath, C. (1986). *Body Movement and Speech in Medical Interaction*. Cambridge: Cambridge University Press.

Heritage, J. (1984). A Change-of-State Token and Aspects of Its Sequential Placement. *Structures of Social Action*, 299-345.

Heritage, J. (1999). CA at century's end: practices of talk-in-interaction, their distributions and their outcomes. *Research on Language and Social Interaction, 32*, 69-76.

Heritage, J. (2007). Intersubjectivity and progressivity in person (and place) reference. In N. J. Enfield & T. Stivers (Eds.), *Person Reference in Interaction: Linguistic, Cultural and Social Perspectives* (pp. 255-280). Cambridge: Cambridge University Press.

Heritage, J. (2011). Conversation analysis: Practices and methods. In D. Silverman (Ed.), *Qualitative Sociology* (3rd ed., pp. 208-230). London: Sage.

Heritage, J., & Greatbatch, D. (1986). Generating Applause: A Study of Rhetoric and Response at Party Political Conferences. *American Journal of Sociology, 92(1)*, 110-157.

Heritage, J., & Maynard, D. W. (2006a). Introduction: Analyzing interaction between doctors and patients in primary care encounters. *Communication in Medical Care: Interaction between primary care physicians and patients*.

Heritage, J., & Maynard, D. W. (Eds.). (2006b). *Communication in Medical Care: Interaction between primary care physicians and patiens*. Cambridge: Cambrdige University Press.

Heritage, J., & Raymond, G. (2005). The terms of agreement: Indexing epistemic authority and subordination in talk-in-interaction. *Social Psychology Quarterly, 68*(1), 15-38.

Heritage, J., & Roth, A. (1995). Grammar and Institution: Questions and Questioning in the Broadcast News Interview. *Research on Language and Social Interaction, 28*(1), 1-60.

Heritage, J., & Sorjonen, M.-L. (1994). Constituting and Maintaining Activities Across Sequences: And-Prefacing as a Feature of Question Design. *Language in Society, 23*, 1-29.

Hietanen, J. K., Leppanen, J. M., Peltola, M. J., Linna-Aho, K., & Ruuhiala, H. J. (2008). Seeing direct and averted gaze activates the approach-avoidance motivational brain systems. *Neuropsychologia, 46*(9), 2423-2430.

Howland, H. C., Merola, S., & Basarab, J. R. (2004). The allometry and scaling of the size of vertebrate eyes. *Vision Research, 44*(17), 2043-2065.

Hutchinson, J. M. C., & Gigerenzer, G. (2005). Simple heuristics and rules of thumb: Where psychologists and behavioural biologists mght meet. *Behavioural Processes, 69*, 97-124.

Jacobs, G. H. (1993). The distribution and nature of colour vision in mammals. *Biological Reviews, 68*(3), 413-471.

Jacobs, G. H. (1995). Variations in primate colour vision: mechanisms and utility. *Evolutionary Anthropology, 3*, 196-205.

Jeannerod, M. (1986). The formation of finger grip during prehension. A cortically mediated visuomotor pattern. *Behavioural Brain Research, 19*, 99-116.

Jefferson, G. (1973). A Case of Precision Timing in Ordinary Conversation: Overlapped Tag-Positioned Address Terms in Closing Sequences. *Semiotica, 9*, 47-96.

Jefferson, G. (1978). Sequential Aspects of Storytelling in Conversation. In J. Schenkein (Ed.), *Studies in the Organization of Coversational Interaction* (pp. 219-248). New York: Academic Press.

Jefferson, G. (1983). Notes on some orderlinesses of overlap onset. In V.D'Urso & P.Leonardi (Eds.), *Discourse Analysis and Natural Rhetoric* (pp. 11-38). Padua: Cleup Editore.

Jefferson, G. (1986). Notes on `Latency' in Overlap Onset. *Human Studies, 9*, 153-183.

Jefferson, G. (1988). Notes on a possible metric which provides for a 'standard maximum' silence of approximately one second in conversation. In D. R. a. P. Bull (Ed.), *Conversation; An interdisciplinary perspective*. Clevedon, UK: Multilingual Matters.

Jefferson, G. (1989). Preliminary Notes on a Possible Metric which Provides for a 'Standard Maximum' Silences of Approximately One Second in Conversation. In D. Roger & P. Bull (Eds.), *Conversation: An Interdiscipoinary Perspective* (pp. 166-196). Clevedon: Multilingual Matters.

Jefferson, G. (2004a). Glossary of transcript symbols with an introduction. In G. Lerner (Ed.), *Conversation Analysis: Studies from the First Generation* (pp. 13-31). Amsterdam: Benjamins.

Jefferson, G. (2004b). A note on laughter in 'male-female' interaction. *Discourse Studies, 6*(1), 117-133.

Jefferson, G. (2004c). A sketch of some orderly aspects of overlap in natural conversation. In G. Lerner (Ed.), *Conversation Analysis: Studies from the First Generation* (pp. 43-59). Amsterdam: Benjamins.

Johansson, R. S., Westling, G., Backstrom, A., & Flanagan, J. R. (2001). Eye-Hand Coordination in Object Manipulation. *The Journal of Neuroscience, 21*(17), 6917-6932.

Johnson, M. H., Griffin, R., Csibra, G., Halit, H., Farroni, T., De Haan, M., et al. (2005). The emergence of the social brain network: Evidence from typical and atypical development. *Developmental Psychopathology, 17*(3), 599-619.

Kampe, K. K. W., Frith, C. D., & Frith, U. (2003). "Hey John": Signals conveying communicative intention toward the self activate brain regions associated with "mentalizing," regardless of modality. *The Journal of Neuroscience, 23*(12), 5258-5263.

Kaplan, G., & Rogers, L. J. (2002). Patterns of Gazing in Orangutans (Pongo pygmaeus). *International Journal of Primatology, 23*(3), 501-526.

Kendon, A. (1967). Some Functions of Gaze-Direction in Social Interaction. *Acta Psychologica, 26*, 22-63.

Kendon, A. (1973). The Role of Visible Behavior in the Organization of Social Interaction. *Social Communication and Movement: Studies of Interaction and Expression in Man and Chimpanzee*, 29-74.

Kendon, A. (1977). Spatial Organization in Social Encounters: The F-formation System. In A. Kendon (Ed.), *Studies in the Behavior of Social Interaction* (pp. 179-208). Lisse, Holland: Peter DeRidder Press.

Kendon, A. (1978). Looking in Conversation and the Regulation of Turns at Talk: A Comment on the Papers of G. Beattie and D.R. Rutter et al. *British Journal of Social and Clinical Psychology, 17*, 23-24.

Kendon, A. (1990). Conducting Interaction: Patterns of Behavior in Focused Encounters.

Kendon, A. (1995). Gestures as Illocutionary and Discourse Structure Markers in Southern Italian Conversation. *Journal of Pragmatics, 23*, 247-279.

Kendon, A. (2002). Some uses of the head shake. *Gesture, 2*(2), 147-182.

Kidwell, M. (2005). Gaze as Social Control: How Very Young Children Differentiate "The Look" From a "Mere Look" by Their Adult Caregivers. *Research on Language and Social Interaction, 38*(4), 417-449.

Kidwell, M. (2006). 'Calm down!': the role of gaze in the interactional management of hysteria by the police. *Discourse Studies, 8*(6), 745-770.

Kidwell, M. (2009). Gaze Shift as an Interactional Resource for Very Young Children. *Discourse Processes, 46*, 145-160.

Kleinke, C. L. (1986). Gaze and Eye Contact: A Research Review. *Psychological Bulletin, 100*(1), 78-100.

Kobayashi, H. (1997). Unique morphology of the human eye. *Nature, 387*, 767-768.

Kobayashi, H. (2001). Unique morphology of the human eye and its adaptive meaning: comparative studies on external morphology of the primate eye. *Journal of Human Evolution, 40*, 419-435.

Kobayashi, H., & Hashiya, K. (2011). The gaze that groomsL contribution of social factors to the evolution of primate eye morphology. *Evolution and Human Behavior, 32*(3), 157-165.

Koenig, C. J. (2007). Question-initiated oblique sequences: A recipience's practice in multi-unit turn environments. Unpublished manuscript.

Labov, W., & Fanshel, D. (1977). *Therapeutic Discourse: Psychotherapy as Conversation*. New York: Academic Press.

LaFrance, M. (1974). Nonverbal cues to conversational turn-taking between black speakers. *Personality and Social Psychology Bulletin, 1*, 240-242.

LaFrance, M., & Mayo, C. (1976). Racial Differences in Gaze Behavior During Conversation: Two Systematic Observational Studies. *Journal of Personality and Social Psychology, 33*, 547-552.

Land, M. F. (2006). Eye movements and the control of actions in everyday life. *Progress in Retinal and Eye Research, 25*, 296-324.

Land, M. F., & Fernald, R. D. (1992). The evolution of eyes. *Annual Review of Neuroscience, 15*, 1-29.

Land, M. F., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception, 28*, 1311-1328.

Land, M. F., & Nilsson, D. E. (2002). *Animal Eyes*. Oxford: Oxford University Press.

Lavelli, M., & Fogel, A. (2005). Developmental changes in the relationship between the infant's attention and emotion during early face-to-face communication: The 2-month transition. *Developmental Psychology, 41*(1), 265-280.

Le Gros Clark, W. E. (1959). *History of the Primates: An Introduction to the Study of Fossil Man*. Chicago: Univeristy of Chicago Press.

Leach, E. (1972). The Influence of Cultural Context on Non-Verbal Communication in Man. In R. Hinde (Ed.), *Non-Verbal Communication* (pp. 315-347). Cambridge: Cambridge University Press.

Lee, S.-H. (2008). Extended requesting: Interaction and collaboration in the production and specification of requests. *Journal of Pragmatics, 41*(6), 1248-1271.

Lerner, G. H. (1991). On the Syntax of Sentences in Progress. *Language in Society, 20*, 441-458.

Lerner, G. H. (1992). Assisted Storytelling: Deploying Shared Knowledge as a Practical Matter. *Qualitative Sociology, 15*(3), 247-271.

Lerner, G. H. (1996). On the 'semi-permeable' character of grammatical units in conversation: conditional enmtry into the turn-space of another speaker. In E. Ochs, E. A. Schegloff & S. Thompson (Eds.), *Interaction and Grammar* (pp. 238-276). Cambridge: Cambridge University Press.

Lerner, G. H. (2003). Selecting next speaker: The context sensitive operation of a context-free organization. *Language in Society, 32*, 177-201.

Lerner, G. H. (2004). On the place of linguistic resources in the organization of talk-in-interaction: grammar as action in prompting a speaker to elaborate. *Research on Language and Social Interaction, 37*(2), 151-184.

Leslie, A. M. (1987). Pretense and Representation: The Origins of "Theory of Mind". *Psychological Review, 94*(4), 412-426.

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

Levinson, S. C. (2005). Living with Manny's dangerous idea. *Discours Studies, 7*(4-5), 431-453.

Levinson, S. C. (2006). On the human "interaction engine". In N. J. Enfield & S. C. Levinson (Eds.), *Roots of human sociality: Culture, cognition and interaction* (pp. 39-69). Oxford: Berg.

Levinson, S. C. (in press). Action Formation and Ascription. In J. Sidnell & T. Stivers (Eds.), *Handbook of conversation analysis*. Malden, MA: Wiley-Blackwell.

Liversedge, S. P., Gilchrist, I. D., & Everling, S. (2011). *Oxford handbook of eye movements*. Oxford: Oxford University Press.

Local, J., & Walker, G. (2004). Abrupt-joins as a resource for the production of multi-unit, multi-action turns. *Journal of Pragmatics, 36*, 1375-1403.

Lorenz, K. Z. (1966). Evolution of ritualization in the biological and cultural spheres. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 251*, 273-284.

Loveland, K. A., & Landry, S. H. (1986). Joint attention and language in autism and developmental language delay. *Journal of Autism and Developmental Disorders, 16*(3), 335-349.

Mangione-Smith, R., Stivers, T., Elliott, M. N., McDonald, L., & Heritage, J. (2003). The Relationship Between Online Commentary Use and Prevention of Inappropriate Antibiotic Prescribing by Pediatricians. *Social Science and Medicine*.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: W.H. Freeman & Co.

Mazeland, H. (2007). Parenthetical Sequences. *Journal of Pragmatics, 39*(10), 1816-1869.

Milinski, M., Semmann, D., & Krambeck, H.-J. (2002). Reputation helps solve the 'tragedy of the commons'. *Nature, 415*, 424-426.

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*. Oxford: Oxford University Press.

Milner, A. D., & Goodale, M. A. (2008). Two visual systems re-viewed. *Neuropsychologia, 46*(3), 774-785.

Mondada, L. (2002). Describing surgical gestures: the view from researcher's and surgeon's video recordings. *Proceedings of the Gesture Conference, Austin TX*.

Mondada, L. (2006). Participants' online analysis and multimodal practices: projecting the end of the turn and the closing of the sequence. *Discourse Studies, 8*(1), 117-129.

Mondada, L. (2007). Multimodal resources for turn-taking: pointing and the emergence of possible next speakers. *Discourse Studies, 9*(2), 194-225.

Monzoni, C. M. (2008). Introducing direct complaints through questions: the interactional achievement of 'pre-sequences'? *Discourse Studies, 10*(1), 73-87.

Morris, D. (1957). "Typical Intensity" and its relation to the problem of ritualisation. *Behaviour, 11*(1), 1-12.

Morris, D. (1985). *Body Watching*. Oxford: Equinox Ltd.

Mundy, P. (1995). Joint attention and social-emotional approach behavior in children with autism. *Development and Psychopathology, 7*, 63-82.

Mundy, P., & Newell, L. (2007). Attention, Joint Attention and Social Cognition. *Current Directions in Psychological Science, 16*, 269-274.

Myowa-Yamakoshi, M., Tomonaga, M., Tanaka, M., & Matsuzawa, T. (2003). Preference for human direct gaze in infant chimpanzees (Pan Troglodytes). *Cognition, 89*, 53-64.

Nichols, K. A., & Champness, B. G. (1971). Eye Gaze and the GSR. *Journal of Experimental Social Psychology, 7*, 623-626.

Nielsen, G. (1962). *Studies in Self-Confrontation*. Copenhagen: Munksgaard.

Nilsson, D. E., & Pelger, S. (1994). A pessimistic estimate of the time required for an eye to evolve. *Proceedings of the Royal Society of London Series B-Biological Sciences, 256*(1345), 53-58.

Norrick, N. R. (2010). Laughter before the punch line during the performance of narrative jokes in conversation. *Text & Talk, 31*(1), 75-95.

Parker, A. R. (2003). *In the blink of an eye*. New York, NY: Basic Books.

Pelphrey, K. A., Morris, J. P., & McCarthy, G. (2004). Grasping the intentions of others: The perceived intentionality of an action influences activity in the superior temporal sulcus during social perception. *Journal of Cognitive Neuroscience, 16*, 1706-1716.

Pomerantz, A. (1978). Compliment Reponses: Notes on the Co-Operation of Multiple Constraints. In J. Schenkein (Ed.), *Studies in the Organization of Conversational Interaction* (pp. 79-112). New York: Academic Press.

Pomerantz, A. (1984a). Agreeing and Disagreeing with Assessments: Some Features of Preferred/Dispreferred Turn Shapes. *Structures of Social Action: Studies in Conversation Analysis*, 57-101.

Pomerantz, A. (1984b). Pursuing a Response. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action* (pp. 152-164). Cambridge: Cambridge University Press.

Pomerantz, A. (1986). Extreme Case Formulations: A Way of Legitimizing Claims. *Human Studies, 9*, 219-229.

Pomerantz, A. M. (1990). Conversation analytic claims. *Communication Monographs, 57*, 231-235.

Popper, K. (1959 [1934]). *The Logic of Scientific Discovery*. London: Hutchinson.

Raymond, G. (2003). Grammar and social organization: Yes/No interrogatives and the structure of responding. *American Sociological Review, 68*, 939-967.

Rayner, K. (1998). Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin, 124*(3), 372-422.

Redican, W. K. (1975). Facial expressions in Nonhuman Primates. In L. A. Rosenblum (Ed.), *Primate Behavior. Developments in Field and Laboratory Research* (Vol. 4, pp. 103-194). New York: Academic Press.

Regan, B. C., Julliot, C., Simmen, B., Vienot, F., Charles-Dominique, P., & Mollon, J. D. (2001). Fruits, foliage and the evolution of primate colour vision. *Philosophical Transactions of the Royal Society of London- Series B Biology, 356*, 229-283.

Robinson, J. D. (1998). Getting Down to Business: Talk, Gaze, and Body Orientation During Openings of Doctor-Patient Consultations. *Human Communication Research, 25*(1), 97-123.

Robinson, J. D. (2007). The role of numbers and statistics within Conversation Analysis. *Communication Methods and Measures, 1*(1), 65-75.

Ross, C. F., & Kirk, E. C. (2007). Evolution of eye size and shape in primates. *Journal of Human Evolution, 52*, 294-313.

Rossano, F. (2010). Questioning and responding in Italian. *Journal of Pragmatics, 42*(10), 2756-2771.

Rossano, F., Brown, P., & Levinson, S. C. (2009). Gaze, Questioning and Culture. *Conversation Analysis: Comparative Perspectives*, 187-249.

Rutter, D. R. (1984). *Looking and Seeing: The Role of Visual Communication in Social Interaction*. Chichester: Wiley.

Ruusuvuori, J., & Peräkylä, A. (2009). Facial and verbal expressions in assessing stories and topics. *Research on Language and Social Interaction, 42*(4), 377-394.

Sacks, H. (1972). On the Analyzability of Stories by Children. In J. J. Gumperz & D. Hymes (Eds.), *Directions in Sociolinguistics: The Ethnography of Communication* (pp. 325-345). New York: Holt, Rinehart and Winston.

Sacks, H. (1974). An Analysis of the Course of a Joke's Telling in Conversation. *Explorations in the Ethnography of Speaking*, 337-353.

Sacks, H. (1984a). Notes on Methodology. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action* (pp. 21-27). Cambridge: Cambridge University Press.

Sacks, H. (1984b). On Doing 'Being Ordinary'. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action* (pp. 413-429). Cambridge: Cambridge University Press.

Sacks, H. (1986). Some Considerations of a Story Told in Ordinary Conversation. *Poetics, 15*, 127-138.

Sacks, H. (1987). On the Preferences for Agreement and Contiguity in Sequences in Conversation. *Talk and Social Organisation*, 54-69.

Sacks, H. (1992 [1964-72]). *Lectures on Conversation (2 Vols.)*. Oxford: Basil Blackwell.

Sacks, H., & Schegloff, E. A. (1979). Two Preferences in the Organization of Reference to Persons and Their Interaction. In G. Psathas (Ed.), *Everyday Language: Studies in Ethnomethodology* (pp. 15-21). New York: Irvington Publishers.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language, 50*, 696-735.

Salvini-Plawen, L. V., & Mayr, E. (1977). On the evolution of photoreceptors and eyes. *Evolutionary Biology, 10*, 207-263.

Sapir, E. (1963 [1927]). The Unconscious Patterning of Behavior in Society. In D. G. Mandelbaum (Ed.), *Selected Writings of Edward Sapir in Language, Culture and Society* (pp. 544-559). Berkeley: University of California Press.

Sartre, J.-P. (1956 [1943]). *Being and Nothingness*. New York: Philosophical Library.

Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature, 253*, 265-266.

Scheflen, A. E. (1964). The Significance of Posture in Communication Systems. *Psychiatry*, 316-331.

Scheflen, A. E. (1975a). Micro-territories in human interaction. In A. Kendon, R. M. Harris & M. R. Key (Eds.), *The Organization of Behavior in Face-to-Face Interaction* (pp. 159-173). The Hague: Mouton Publishers.

Scheflen, A. E. (1975b). Models of epistemologies in the study of interaction. In A. Kendon, R. M. Harris & M. R. Key (Eds.), *Organization of Behavior in Face-to-Face Interaction* (pp. 63-91). The Hague: Mouton Publishers.

Schegloff, E. A. (1968). Sequencing in Conversational Openings. *American Anthropologist, 70*(6), 1075-1095.

Schegloff, E. A. (1979). The Relevance of Repair for Syntax-for-Conversation. *Syntax and Semantics 12: Discourse and Syntax*, 261-288.

Schegloff, E. A. (1980). Preliminaries to Preliminaries: 'Can I Ask You a Question'. *Sociological Inquiry, 50*, 104-152.

Schegloff, E. A. (1982). Discourse as an Interactional Achievement: Some Uses of 'Uh huh' and Other Things that Come Between Sentences. *Georgetown University Roundtable on Languages and Linguistics*, 71-93.

Schegloff, E. A. (1984). On Some Questions and Ambiguities in Conversation. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action* (pp. 28-52). Cambridge: Cambridge University Press.

Schegloff, E. A. (1993). Reflections on Quantification in the Study of Conversation. *Research on Language and Social Interaction, 26*, 99-128.

Schegloff, E. A. (1995). Discourse as an Interactional Achievement III: The Omnirelevance of Action. *Research on Language and Social Interaction*((Special issue of *Co-Construction* edited by Sally Jacoby and Elinor Ochs)), 185-212.

Schegloff, E. A. (1996a). Confirming Allusions: Toward an Empirical Account of Action. *American Journal of Sociology, 104*, 161-216.

Schegloff, E. A. (1996b). Some practices for referring to persons in talk-in interaction: a partial sketch of a systematics. In B. Fox (Ed.), *Studies in Anaphora* (pp. 437-485). Amsterdam: John Benjamins.

Schegloff, E. A. (1996c). Turn Organization: One Intersection of Grammar and Interaction. *Interaction and Grammar*, 52-133.

Schegloff, E. A. (1997). Practices and Actions: Boundary Cases of Other-initiated Repair. *Discoure Processes, 23*(3), 499-545.

Schegloff, E. A. (1998). Body Torque. *Social Research, 65*(3), 535-596.

Schegloff, E. A. (2000a). Overlapping Talk and the Organization of Turn-Taking for Conversation. *Language in Society, 29*, 1-63.

Schegloff, E. A. (2000b). When 'Others' Initiate Repair. *Applied Linguistics, 21*(2), 205-243.

Schegloff, E. A. (2001). Getting Serious: Joke -> Serious 'No'. *Journal of Pragmatics, 33*(12), 1947-1955.

Schegloff, E. A. (2001). *Increments: Where they are and what they do.* Paper presented at the Linguistic Institute, Santa Barbara, CA.

Schegloff, E. A. (2004). *Whistling in the Dark: Notes from the Other Side of Liminality.* Paper presented at the Twelfth Annual symposium about Language and Society, Austin.

Schegloff, E. A. (2006). Interaction: the infrastructure for social institutions, the natural ecological niche for language, and the arena in which culture is enacted. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of Human Sociality: Culture, Cognition and Interaction* (pp. 70-96). Oxford: Berger.

Schegloff, E. A. (2007a). Categories in action: person-reference and membership categorization. *Discourse Studies, 9*(4), 433-461.

Schegloff, E. A. (2007b). *Sequence Organization in Interaction: a Primer in Conversation Analysis*. Cambridge, England: Cambridge University Press.

Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The Preference for Self-Correction in the Organization of Repair in Conversation. *Language, 53*, 361-382.

Schegloff, E. A., & Sacks, H. (1973). Opening Up Closings. *Semiotica, 8*, 289-327.

Searle, J. R. (1970). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.

Selting, M. (2000). The construction of units in conversational talk. *Language in Society, 29*, 477-517.

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology, 18*, 668-671.

Senju, A., & Johnson, M. H. (2009). The eye contact effect: mechanisms and development. *Trends in Cognitive Sciences, 13*(3), 127-134.

Sidnell, J. (2004). There's risks in everything: Extreme case formulations and accountability in inquiry testimony. *Discourse and Society, 15*(6), 745-766.

Sidnell, J. (2006). Coordinating Gesture, Talk, and Gaze in Reenactments. *Research on Language and Social Interaction, 39*(4), 377-409.

Sidnell, J. (2011). *Conversation analysis: An introduction*. Oxford, UK: Wiley-Blackwell.

Sidnell, J. (Ed.). (2009). *Conversation Analysis: Comparative Perspectives*. Cambridge: Cambridge University Press.

Simmel, G. (1969). Sociology of the Senses: Visual Interaction. In R. E. Park & W. Burgess (Eds.), *Introduction to the Science of Sociology* (pp. 356-361). Chicago: University of Chicago Press.

Simon, H. A. (1990). Invariants of Human Behavior. *Annual Review of Psychology, 41*, 1-19.

Smith, G. E. (1924). *The evolution of Man*. London: Oxford University Press.

Sommer, R. (1959). Studies in Personal Space. *Sociometry, 23*, 247-260.

Sommer, R. (1962). The distance of comfortable conversation: a further study. *Sociometry, 25*(1), 111-116.

Sommer, R. (1967). Sociofugal Space. *American Journal of Sociology, 72*(6), 654-660.

Spelke, E. S. (1985). Preferential looking methods as tools for the study of cognition in infancy. In G. Gottlieb & N. Krasnegor (Eds.), *Measurement of audition and vision in the first year of postnatal life* (pp. 323-363). Norwood, NJ: Ablex.

Sperber, D. (2006). Why a deep understanding of cultural evolution is incompatible with shallow psychology. In N. Enfield & S. Levinson (Eds.), *Roots of Human Sociality* (pp. 431-449). Oxford: Berg.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and Cognition*. Cambridge, Mass.: Harvard University Press.

Spezio, M. L., Huang, S., Castelli, F., & Adolphs, R. (2007). Amygdala Damage Impairs Eye Contact During Conversations with Real People. *The Journal of Neuroscience, 27*(15), 3994-3997.

Stern, D. N. (1974). Mother and infant at play: The dyadic interaction involving facial, vocal, and gaze behaviors. In M. Lewis & L. A. Rosenblum (Eds.), *The effect of the infant on its caregivers* (pp. 187-213). New York: Wiley.

Stivers, T. (2005). Modified repeats: One method for asserting primary rights from second position. *Research on Language and Social Interaction, 38*(2), 131-158.

Stivers, T. (2008). Stance, alignment and affiliation during story telling: When nodding is a token of preliminary affiliation. *Research on Language and Social Interaction, 41*, 29-55.

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., et al. (2009). Universals and cultural variation in turn-taking in conversation. *PNAS, 106*(26), 10587-10592.

Stivers, T., & Majid, A. (2007). Questioning Children: Interactional Evidence of Implicit Bias in Medical Interviews. *Social Psychology Quarterly, 70*, 424-441.

Stivers, T., & Robinson, J. (2006). A preference for progressivity in interaction. *Language in Society, 35*(3), 367-392.

Stivers, T., & Rossano, F. (2010). Mobilizing Response. *Research on Language and Social Interaction, 43*(1), 3-31.

Streeck, J. (1993). Gesture as Communication I: Its Coordination with Gaze and Speech. *Communication Monographs, 60*(4), 275-299.

Streeck, J. (1994). Gestures as Communication II: The Audience as Co-Author. *Research on Langauge and Social Interaction, 27*(3), 223-238.

Streeck, J. (1995). On Projection. In E. N. Goody (Ed.), *Social Intelligence and Interaction: Expressions and Implications of the Social Bias in Human Intelligence* (pp. 87-110). Cambrdige: Cambridge University Press.

Streeck, J. (2003). The Body Taken for Granted: Lingering Dualism in Research on Social Interaction. In P. Glenn, C. LeBaron & J. Mandelbaum (Eds.), *Studies in Language and Social Interaction* (pp. 427-440). Mahwah, NJ: Lawrence Erlbaum.

Streeck, J. (2009). *Gesturecraft: The Manu-facture of Meaning*. Amsterdam: John Benjamins.

Streeck, J., Goodwin, C., & LeBaron, C. D. (Eds.). (2011). *Embodied Interaction: Language and Body in the Material World*. Cambridge: Cambridge University Press.

Sussman, R. W. (1991). Primate origins and the evolution of angiosperms. *American Journal of Primatology, 23*, 209-223.

Tanenhaus, M. K., Magnuson, J. S., Dahan, D., & Chambers, C. (2000). Eye movements and lexical access in spoken-language comprehension: evaluating a linking hypothesis

between fixations and linguistic processing. *Journal of Psycholinguistic Research, 29*(6), 557-580.

Terasaki, A. (2004 [1976]). Pre-announcement sequences in conversation. In G. Lerner (Ed.), *Conversation Analysis: Studies from the First Generation* (pp. 171-223). Amsterdam: John Benjamins.

Thayer, S., & Schiff, W. (1974). Observer Judgment of Social Interaction: Eye Contact and Relationship Inferences. *Journal of Personality and Social Psychology, 30*(1), 110-114.

Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. Dunham (Eds.), *Joint attention: its origins and role in development* (pp. 103-130). Hillsdale, NJ: Erlbaum.

Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Cambridge, Ma: Harvard University Press.

Tomasello, M. (2009). *Why we cooperate*. Cambridge, MA: The MIT Press.

Tomasello, M., & Farrar, J. (1986). Joint attention and early language. *Child Development, 57*, 1454-1463.

Tomasello, M., Hare, B., Lehmann, H., & Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: the cooperative eye hypothesis. *Journal of Human Evolution, 52*, 314-320.

Tomasello, M., & Todd, J. (1983). Joint attention and lexical acquisition style. *First Language, 4*, 197-212.

Tomkins, S. S. (1963). *Affect, Imagery, Conciousness* (Vol. II). New York: Springer.

Torres, O., Cassell, J., & Prevost, S. (1997). Modeling Gaze Behavior as a Function of Discourse Structure. *First International Workshop on Human Computer Conversations*.

Van Dijk, T. A. (1979). Pragmatic Connectives. *Journal of Pragmatics, 3*(5), 447-456.

von Grünau, M., & Anston, C. (1995). The detection of gaze direction: A stare-in-the-crowd effect. *Perception, 24*(11), 1297-1313.

Watson, J. B. (1913). Psychology as the behaviourist views it. *Psychological Review, 20*, 158-177.

Watson, J. B. (1924). *Psychology from the standpoint of a behaviourist*. Philadelphia: Lippincott.

Weber, M. (1978 [1922]). *Economy and Society*. Berkeley: University of California Press.

Yamagiwa, J. (1992). Functional Analysis of Social Staring Behavior in an All-male Group of Mountain Gorillas. *Primates, 33*(4), 523-544.

Yngve, V. H. (1970). On Getting a Word in Edgewise. *Papers from the Sixth Regional Meeting, Chicago Linguistics Society*, 567-578.

## Samenvatting

Wat betekent het om te kunnen zien en om gezien te kunnen worden door anderen? Wat doen we met onze ogen als we met elkaar praten? Dit proefschrift onderzoekt hoe mensen hun ogen en hun lichaam gebruiken tijdens onderlinge ontmoetingen, en het zet de patronen in hun gedrag uiteen. In het bijzonder laat het zien hoe personen uit een bepaalde cultuur (Italianen uit Emila-Romagna, een streek in Noord-Italië) kijken tijdens natuurlijk voorkomende tweegesprekken in hun huis. Het doel is om enkele van de gewoonten en normen uiteen te zetten die deze personen hanteren, en zo te komen tot een gedetailleerde beschrijving van de structuren van kijkgedrag in menselijke interacties en ons begrip te verfijnen van de wijze waarop mensen deelnemen aan sociale situaties. Bovendien leidt een onderzoek naar kijkgedrag in onderlinge ontmoetingen ertoe dat we beter begrijpen hoe deelnemers omgaan met hun rechten en plichten in interactie, vooral met betrekking tot samenwerken, agency en de verantwoording van sociale handelingen. Het belicht ook de fijnmazige afstemming en organisatie van beurtwisselingen en sequenties van sociale activiteiten. De nadruk in dit proefschrift ligt op de manier waarop deelnemers kijken aan het begin, in het midden en aan het eind van een spraaksequentie. De resultaten hiervan dragen bij aan een algemener begripskader van de wijze waarop mensen in interacties hun ogen gebruiken, vooral als we ze in een vergelijkende en ontwikkelingscontext bezien. Bovendien verschaffen deze resultaten computerdeskundigen en deskundigen op het gebied van kunstmatige intelligentie belangrijke informatie over hoe voorspelbaar de blikrichting van gespreksdeelnemers is en in welke mate een robot of avatar dit zou kunnen nabootsen.

### Data en methode

Om te kunnen analyseren en begrijpen hoe gespreksdeelnemers omgaan met de grenzen van talige handelingssequenties heb ik een reeks video-opnames gemaakt van natuurlijk voorkomende interacties waarbij mensen samen eten, kaartspelen, koken, koffiedrinken of in een auto reizen. De deelnemers hebben geen script of instructie gekregen met betrekking tot het verloop van hun gesprekken. De opgenomen interacties ontstonden op natuurlijke wijze, en hadden dus ook plaatsgevonden onafhankelijk van mijn verzoek om de deelnemers te mogen filmen, waardoor deze situaties en het verloop van de gesprekken spontaan zijn en niet

369

geënsceneerd. In de meeste situaties was er sprake van tweegesprekken, hoewel er ook enkele driegesprekken zijn gebruikt om te kunnen komen tot fundamentele inzichten over de organisatie van het kijkgedrag. Details over de deelnemers, de situatie, de locatie, en de hoofdactiviteit van elke opgenomen interactie worden vermeld in Appendix A.

Gelet op de nadruk op de systematische organisatie van kijkgedrag in relatie tot sociale handelingen, die vooral tot stand komen door middel van talige interacties, had ik een methodologie nodig die mij in staat zou stellen om specifieke conversatiestructuren te analyseren en te identificeren. Conversatie Analyse (hierna: CA) verschafte dit raamwerk. Deze methode richt zich meer op actie en interactie dan op taal of taalkundige categorieën. CA stelt de onderzoeker in staat om recht te doen aan de fijnmazigheid van interacties en details te analyseren zonder de macrostructuren, zoals sequentieorganisatie en de globale *overall*-organisatie van de activiteit uit het oog te verliezen. Bovendien voorkomt de focus op het perspectief van de gespreksdeelnemers, hun oriëntatie op en interpretatie van de handelswijze van de andere gespreksdeelnemers, een ad hoc-gebruik van categorieën en labels, en een analyse die niet terug te voeren is op de data doordat ze niet ondersteund wordt door de principes waar de gespreksdeelnemers zelf zich op richten. Deze convergentie tussen het perspectief van de onderzoeker en dat van de deelnemers, de kern van conversatieanalytisch onderzoek, is in dit proefschrift volledig opgenomen: ik heb niet eenvoudigweg beschreven wanneer mensen hun ogen bewegen, maar ook vastgelegd welke gevolgen voor de interactie een verschuiving van de blikrichting kan hebben.

De hier gerapporteerde data zijn echter niet alleen kwalitatief geanalyseerd. Inzichten uit andere disciplines zijn gebruikt in de algemene analyse van de data. Ik ben begonnen met een kwalitatieve analyse van bepaalde delen van de interacties, waarbij ik vergelijkingen heb gemaakt van interacties waar bepaalde interactionele handelingen en praktijken tentoongespreid leeken te worden. Na het uiteenzetten van de basale interactionele kenmerken van elke activiteit, ging ik terug naar de data en codeerde systematisch wat er in een bepaald aantal minuten (meestal tien minuten, zoals in ieder hoofdstuk aangegeven), die willekeurig uit elke interactie gekozen werden, gebeurde. De codering werd toegepast op de variabelen die eerder waren geïdentificeerd als cruciaal voor de onderzochte interactionele handeling. Vervolgens heb ik statistische analyses (vooral logische regressies) uitgevoerd om significante correlaties tussen de variabelen en belangrijke predictoren van bepaalde resultaten vast te stellen. De kwantitatieve

analyse diende ertoe om de organisatie van bepaald gedrag te identificeren en te kwantificeren, door de empirische hypothese over de functie van bepaald kijkgedrag die uit het kwalitatief onderzoek naar voren kwam te falsificeren (Popper, 1959 [1934]).

Om de details van de spraak en het zichtbare gedrag in iedere interactie te annoteren, heb ik gebruik gemaakt van de software ELAN (http://www.lat-mpi.eu/tools/elan/), die is ontwikkeld op het Max Planck Instituut voor Psycholinguïstiek om onderzoekers in staat te stellen om van dezelfde interactie tegelijk meerdere opnames (audio en video) te bekijken en te annoteren.

Om in geschreven transcripties van de conversaties ook het kijkgedrag van de deelnemers weer te geven, heb ik een nieuw transcriptiesysteem ontwikkeld, dat in Appendix B is opgenomen.

### De resultaten

Dit proefschrift onderzoekt hoe mensen hun ogen en hun lichaam gebruiken tijdens onderlinge ontmoetingen. Eerder onderzoek naar kijkgedrag in gesprekken stelde dat het kijkgedrag beïnvloed werd door factoren van buiten het gesprek (bijvoorbeeld geslacht, leeftijd en sociale status; zie bijvoorbeeld Ellsworth & Ludwig, 1972; Kleinke, 1986), dat het kijkgedrag de mate van betrokkenheid van de deelnemers aan gesprek aantoont (bijvoorbeeld Argyle & Cook, 1976; Goffman, 1963; Goodwin, 1981), dat het een regulerende rol heeft bij beurtwisselingen (bijvoorbeeld Duncan, 1975; Kendon, 1967) en, ten slotte, dat het een rol heeft bij het in gang zetten van sociale handelingen, afhankelijk van de context en de uitvoering ervan (e.g., Kidwell, 2005, 2009; Sidnell, 2006). Dit proefschrift bevat diverse belangwekkende uitkomsten die een heroverweging van een aantal van deze stellingen noodzakelijk maken. Die uitkomsten worden uiteengezet in drie empirische hoofdstukken (de hoofdstukken 2, 3 en 4).

Hoofdstuk 2 toont aan dat de recipiënten van een vertelling hun blik naar de verteller richten gedurende de eerste beurt van de vertelling. Wanneer een persoon echter recipiënt is van de eerste beurt van een aangrenzend paar, laten ze geen bijzondere tendens zien in hun kijkgedrag. Als zodanig worden blikken anders gebruikt, afhankelijk van de sociale handelingen en activiteiten waarbij gespreksdeelnemers betrokken zijn. Dit plaats ons voor een raadsel met betrekking tot de wijze waarop luisteraars nagaan of ze bijvoorbeeld te maken hebben met de eerste beurtopbouweenheid (BOE) van een langere vertelling  of met een aankondiging bestaande uit één BOE. Het antwoord ligt, ten minste gedeeltelijk, in het feit dat luisteraars

vertrouwen op zogenaamde "fast and frugal heuristics" (Gigerenzer en Goldstein, 1996), gebaseerd op de semantische domeinen in de uiting, om te herkennen wat voor handelingsproject de spreker begint. Beurtopbouweenheden die een telling initiëren, bevatten bijvoorbeeld vaak een onderwerp in de eerste persoon (enkelvoud of meervoud) en een verwijzing naar het verleden of de toekomst. Ook geven sprekers hun epistemische toegang aan met betrekking tot de gebeurtenissen die worden gemeld en/of maken zij melding van een derde persoon die niet eerder is besproken. Eerste beurtopbouweenheden die een aangrenzend paar inleiden, bevatten daarentegen vaak een onderwerp in de tweede persoon (enkelvoud of meervoud), alsook deiktische woorden en modale werkwoorden. De verschillende woorden waarmee gespreksdeelnemers hun sequentie initiërende handelingen vorm geven, maken de ontvangers duidelijk of het hier gaat om het begin van een uitgebreide vertelling of om het eerste paardeel van een aangrenzend paar (e.g. Schegloff, 2007). De resultaten uit dit hoofdstuk verschaffen ons een verfijning van Goffman's (1981a) analyse van de verschillende typen luisteraars. Goffman onderscheidt 'officiële' en 'officieuze' gespreksdeelnemers. Binnen de 'officiële' deelnemers maakte hij een nader onderscheid tussen 'aangesproken' en 'niet aangesproken' recipiënten. Hoofdstuk 2 laat zien dat binnen de categorie 'aangesproken recipiënten ' ook nog onderscheid kan worden gemaakt tussen recipiënten van een uitgebreide vertelling en recipiënten van het eerste paardeel van een aangrenzend paar. Deze nadere precisering verfijnt ons inzicht in het interactionele participatiekader en herkenning van het gedrag dat iemand geacht wordt te tonen om een aandachtig en 'aangesproken' recipiënt te zijn. Deze uitkomsten wijzen dan ook op de noodzaak van een hercontextualisatie van veel van de eerdere stellingen met betrekking tot het kijkgedrag. Vooral die stellingen die in een experimentele setting tot stand zijn gekomen, aangezien die studies doorgaans geen rekening hielden met de inhoud van de conversatie en de sociale handelingen die uitgevoerd worden in gespreksbeurten (dat wil zeggen, of de deelnemers elkaar eenvoudige vragen stelden of elkaar een verhaal vertelden).

Hoofdstuk 3 beschrijft de manier waarop het kijken kan worden gebruikt om in een gesprek de ontvanger er toe aan de zetten een reactie te geven. Waar eerder werk de nadruk legde op verbale middelen en hoe die ingezet kunnen worden om een reactie van de ontvanger te ontlokken, laten de bevindingen in dit hoofdstuk zien dat alleen een blik al genoeg is om een reactie te krijgen. Tevens laat de analyse zien dat dit niet alleen kan worden gedaan op een *plaats-relevant-voor-beurt-overdracht*, maar ook in de context van het vertellen van een verhaal,

na iets grappigs of als een 'try-marker'. In het eerste deel van dit hoofdstuk laat ik zien dat de blikken van sprekers en recipiënten gedurende een eerste paardeel verschillende effecten hebben ten aanzien van de productie en het timing van de recipiënt reactie. Het kijken van de spreker naar de recipiënt is bedoeld om een respons uit te lokken terwijl het kijken van de recipiënt naar de spreker bedoeld is om een tijdig respons te bevorderen. Als de gespreksdeelnemers elkaar aankijken is een tijdig response zelfs nog waarschijnlijker. In het tweede deel van dit hoofdstuk analyseer ik hoe het kijken gebruikt kan worden om een ontbrekende recipiëntreactie te ontlokken op een *plaats-relevant-voor-beurt-overdracht* na een eerste paardeel. In het bijzonder laat ik zien hoe de timing van het kijken naar de recipiënt is georganiseerd. Ook laat ik de effectiviteit ervan zien, net als de mogelijke begrenzingen ervan en hoe het zich verhoudt tot andere (verbale) manieren om een recipiëntreactie te ontlokken. Ten slotte, laat ik in het derde deel van dit hoofdstuk zien hoe de kennis die in de eerste twee delen van het hoofdstuk is opgebouwd toepasbaar is in andere gebruiksomgevingen waarbij een respons van de recipiënt relevant is (bijvoorbeeld na een grap of na een verwijzing naar iets bekends zoals de naam van een gezamenlijke vriend), ook in omgevingen waar het respons normaal gesproken niet als conditioneel relevant wordt gezien. Vervolgens laat ik zien hoe de blikrichting van de spreker een rol speelt in het verkrijgen van een recipiëntreactie in deze sequentiële contexten.

Hoofdstuk 4 laat zien hoe sprekers het einde van een sequentie bepalen. Reeksen van beurten vormen doorgaans handelingssequenties die vaak verder gecombineerd worden tot een uitgebreid handelingsproject. Als een sequentie eenmaal een mogelijk punt van voltooiing bereikt heeft, wordt deze dus soms uitgebreid met een reeks beurten terwijl dit in andere gevallen niet gebeurd. In dit hoofdstuk onderzoek ik hoe het kijkgedrag bijdraagt aan de interactionele handeling van het beëindigen van een sequentie. Ik laat zien dat gespreksdeelnemers zich richten op de relevantie van wegkijken aan het einde van een sequentie als dit samenvalt met het beëindigen van het grotere handelingsproject waar de sequentie deel van uitmaakt. Dit in tegenstelling tot eerdere studies die stelden dat het kijken vooral te maken had met beurtwisselingen. Als gespreksdeelnemers een mogelijk sequentie-einde naderen en zij het handelingsproject willen afsluiten, zullen zij de neiging hebben om weg te kijken. In het geval er niet wordt weggekeken wordt het handelingsproject en dus de sequentie uitgebreid tot ze kunnen worden afgesloten zonde oogcontact tussen de deelnemers. Daarmee breidt dit hoofdstuk ons inzicht uit in niet alleen het kijkgedrag maar ook in de sequentiële opbouw van sociale

handelingen. Een algemene vraag die in dit hoofdstuk wordt opgeworpen is of wegkijken in zichzelf beëindiging betekent. Door weg te kijken en niet terug te kijken naar de andere gespreksdeelnemer neemt iemand een houding aan ten opzichte van de gang van zaken waar de andere deelnemer zich wel of niet op kan afstemmen. Het wegkijken van één gespreksdeelnemer maakt geen bepaalde reactie conditioneel relevant, maar het is wel een poging tot en voorstel voor beëindiging van de handeling. Als de andere gespreksdeelnemer zich afstemt op het wegkijken kan het beëindigen van het handelingsproject daadwerkelijk plaatsvinden en doorgaans vindt dit dan ook plaats. Als er niet wordt afgestemd, volgen er doorgaans nog enkele sequentie-uitbreidingen. De stelling is dat het kijkgedrag sequentiebeëindiging bewerkstelligt in gevallen waarbij de beëindiging in principe problematisch is; terwijl het in meer routineuze gevallen vooral een blijk is van begrip dat de sequentie compleet is. Met andere woorden het kijkgedrag kan zowel blijk geven van als een symptoom zijn voor sequentiebeëindiging.

Ten slotte verschaft dit proefschrift een methodologisch raamwerk dat gebruikt kan worden om de hier besproken resultaten uit te breiden naar verwante gesprekscontexten of interacties die voorkomen in andere culturen of bij dieren. De bevindingen zouden een nuttige verbetering moeten verschaffen van onze huidige kennis met betrekking tot oogbewegingen en kijkgedrag tijdens sociale interacties en zijn mogelijk in het bijzonder nuttig voor onderzoekers die geïnteresseerd zijn in toepassingen voor kunstmatige intelligentie of het bestuderen van cross-culturele interacties.

## Curriculum Vitae

Federico Rossano studied Semiotics and Communication Studies at the University of Bologna and spent one year at the University of California, Los Angeles, studying Linguistics, Sociology and Conversation Analysis. He then moved to Nijmegen for his Ph.D., as part of the Multimodal Interaction Project within the Language and Cognition group of the Max Planck Institute for Psycholinguistics. His dissertation focuses on the organization of gaze behavior in social interaction. He is currently working as a post-doctoral researcher in the Department of Developmental and Comparative Psychology at the Max Planck Institute for Evolutionary Anthropology in Leipzig. His current research focuses on the following topics: sequence organization and the organization of action in talk-in-interaction, gestural communication and visible behavior in great apes and in humans (in particular gaze behavior), joint attention in young infants, and the understanding of property in humans and great apes.

# MPI Series in Psycholinguistics

1. *The electrophysiology of speaking. Investigations on the time course of semantic, syntactic and phonological processing.* Miranda van Turennout

2. *The role of the syllable in speech production. Evidence from lexical statistics, metalinguistics, masked priming and electromagnetic midsaggital articulography.* Niels O. Schiller

3. *Lexical access in the production of ellipsis and pronouns.* Bernadette M. Schmitt

4. *The open-/closed-class distinction in spoken-word recognition.* Alette Haveman

5. *The acquisition of phonetic categories in young infants: A self-organising artificial neural network approach.* Kay Behnke

6. *Gesture and speech production.* Jan-Peter de Ruiter

7. *Comparative intonational phonology: English and German.* Esther Grabe

8. *Finiteness in adult and child German.* Ingeborg Lasser

9. *Language input for word discovery.* Joost van de Weijer

10. *Inherent complement verbs revisited: Towards an understanding of argument structure in Ewe.* James Essegbey

11. *Producing past and plural inflections.* Dirk Janssen

12. *Valence and transitivity in Saliba: An Oceanic language of Papua New Guinea.* Anna Margetts

13. *From speech to words.* Arie van der Lugt

14. *Simple and complex verbs in Jaminjung: A study of event categorization in an Australian language.* Eva Schultze-Berndt

15. *Interpreting indefinites: An experimental study of children's language comprehension.* Irene Krämer