Turn-design at turn-beginnings: Multimodal resources to deal with tasks of turn-construction in German

Arnulf Deppermann*

Institut für Deutsche Sprache, PF 101621, D-68161 Mannheim, Germany
Received 11 November 2011; received in revised form 24 July 2012; accepted 27 July 2012

Abstract

Based on German speaking data from various activity types, the range of multimodal resources used to construct turn-beginnings is reviewed. It is claimed that participants in talk-in-interaction need to deal with four tasks in order to construct a turn which precisely fits the interactional moment of its production:

1. Achieve joint orientation: The accomplishment of the socio-spatial prerequisites necessary for producing a turn which is to become part of the participants’ common ground.
2. Display uptake: Next speaker needs to display his/her understanding of the interaction so far as the backdrop on which the production of the upcoming turn is based.
3. Deal with projections from prior talk: The speaker has to deal with projections which have been established by (the) previous turn(s) with respect to the upcoming turn.
4. Project properties of turn-in-progress: The speaker needs to orient the recipient to properties of the turn s/he is about to produce.

Turn-design thus can be seen to be informed by tasks related to the multimodal, embodied, and interactive contingencies of online-construction of turns. The four tasks are ordered in terms of prior tasks providing the prerequisite for accomplishing a later task.

Keywords: Turn construction; Turn-beginnings; Multimodal interaction; Understanding in interaction; Projection

1. Research question and claims

This paper deals with the question: How do turns begin and which factors account for the way turn-beginnings are designed? I will claim that there are two major determinants:

- There are four tasks which participants have to deal with when designing a turn as a contribution to an ongoing interactional encounter: to achieve joint attention to the upcoming turn, to display uptake of prior turn(s), to deal with projections emanating from them, and to project properties of the upcoming turn.

* Tel.: +49 621 1581309; fax: +49 621 1581200.
E-mail address: deppermann@ids-mannheim.de.

1 The author would like to thank two anonymous reviewers from the Journal of Pragmatics for detailed and helpful comments on a prior version of the paper.
The interactional structure achieved at the moment a turn is to be launched, in particular the immediately prior turn, provides for options and restrictions of how a next turn can be designed as a locally sensitive continuation.

Taken together, these factors are seen as structural, interactional and functional motivations of what Schegloff (1996:63) has called “positionally sensitive grammars”. The paper will sketch some general properties of such grammars with respect to German data. It will be shown that an adequate treatment of turn-construction has to adopt a multimodal perspective, because dealing with tasks of turn-construction necessarily involves the simultaneous and sequential combination of various modal resources.

In what follows, I will first review the literature and develop a multimodal perspective on turn-construction (1.). After a short discussion of the notions “turn” and “turn-beginning” (2.) and observations about the distribution of grammatical forms at turn-beginnings (3.), I will try to show that there are four generic tasks speakers have to deal with when beginning a turn (4.). With reference to German data, various multimodal practices are shown to be specialized to deal with one (or more) of these tasks respectively. The paper closes with a discussion of the occasioned relevance of the four tasks and of basic properties of a grammar of turn-beginnings (5.) and it summarizes the findings with respect to a multimodal framework of turn-construction (6.).

2. Turn-beginnings and tasks of turn-construction in multimodal interaction

Starting with the seminal paper on turn-taking by Sacks et al. (1974), turn construction has been a major topic in CA. Since the advent of interactional linguistics, the topic has gained even more attention, with researchers now asking which linguistic resources are used to construct turn constructional units (TCUs) and how transition relevance places (TRPs) are projected (see, among many others, Auer, 1996a; Selting, 2000, 2005; Ford et al., 2002; Ford and Thompson, 1996; Steensig, 2001; Couper-Kuhlen and Ono, 2007). For the linguistic resources used, Schegloff (1996) coined the term “positionally sensitive grammars”. The role of gestures (Streeck and Hartge, 1992; Bohle, 2007; Mondada, 2007; Streeck, 2009) and gaze (Goodwin, 1981) for turn-transition has also been studied in some depth.

Turn-beginnings, however, have received less attention. The most prominent contribution is Schegloff (1996), who aims at discovering “general contingencies of talking in interaction (...) that in a recurrent, orderly way seem to shape the organization of a turn-at-talk” (Schegloff, 1996:54). He discusses the relationship of types of units making up turns to their positions in the turn and to the turn’s position within larger sequences. Most importantly, Schegloff argues why turns “sentence” is an improper category to start with when turn-construction is to be accounted for. Instead, he argues that the TCU is the basic unit of turn-construction. Still, Schegloff (1996:69–80) suggests that turns sometimes start (and end) with structures other than TCUs. The distinction between turn-beginnings and TCU-beginnings has been elaborated on by Lindström (2006). He argues that turns may consist of presegments not being TCUs which can precede the contribution proper of the turn. In order to account for positionally sensitive structures of turn-construction, Schegloff (1996) repeatedly points to functions and “generic jobs” (Schegloff, 1996:81) speakers have to deal with when constructing a turn, like showing understanding, dealing with issues of speaker-selection, stance-taking, producing a sequentially appropriate next action, etc. However, as Schegloff (1996) is rather occupied with the relationship of the shape of (linguistic) structures used for turn-construction and their position within a turn (“positionally sensitive grammars”), he does not aim for a systematic account of the interactional, i.e., sequentially and turn-externally motivated, tasks which seem to inform turn-construction in general. This, in contrast, will be the approach taken by this paper.

Relatively early, Charles Goodwin (1980, 1981) studied the importance to achieve mutual gaze for turn-beginnings. Only more recently, this line of multimodal research has been advanced by Mondada (2009) with respect to bodily coordination in space, pointing (Mondada, 2007), breathing (Schegloff, 1996), moving the body into the visual field of prospective recipients (Mortensen, 2009) and coordination of gaze (Rossano et al., 2009). These studies have shown that establishing observable recipiency first is the prerequisite for taking the turn at the next TRP. The prosodic design of turn-beginnings has been researched by Couper-Kuhlen (2001, 2004), with respect to the role of high onsets for the initiation of new topics and sequences, and Walker (2004) and Raymond (2010) for responsive actions. Research on the grammar of turn-beginnings is still rare. Lindström (2006) points to the fact that non-clausal presegments, which precede sentence starts, in Germanic languages mostly consist of pre-front field elements which are not syntactically integrated within the frame of the following sentence. Auer’s study of pre-front field structures in German (Auer, 1996b) is most important in terms of a grammar of turn-beginnings, although he does not discuss the relevance of this
Topological position specifically with respect to turn-beginnings, but concerning the internal structuring of turns. Studies devoted to more specific phenomena have dealt with devices of linking back to non-adjacent prior turns (Local, 2004; Mazeland and Huiskes, 2001; de Stefani and Hortacher, 2008), with beginnings of dispreferred turns (Pomerantz, 1984; Schegloff, 2010), with the linguistic design of formulations as displays of understanding (Deppermann, 2011) and with prosodic and recycling practices in the context of competitive turn-taking (Schegloff, 1987, 2004; French and Local, 1983). Although some practices relating to specific actions and problems have been studied in detail, we still lack a more comprehensive picture of the factors which matter for the design of turn-beginnings and how they impinge on linguistic choices.

For a turn to become an intersubjectively grounded contribution to an ongoing interaction, it must be designed as to suit the precise interactional moment and environment. To do this, the upcoming speaker has to take into account several factors:

a) What went on before in the interaction, most importantly, in the turn(s) and activities of other participants immediately prior to the turn to be produced (Schegloff, 2007). This crucially includes projections for next turns and activities established by them and their linguistic structure. It also includes co-participants’ simultaneous non-interactional involvements with other tasks, objects, etc.

b) The participation framework (Goodwin and Goodwin, 2004), i.e., participants’ mode of presence in the interaction (only as speakers and listeners in a telephone conversation or as visible actors in face-to-face encounters), their spatial line-up and orientation (vis-à-vis each other), their availability as recipients in terms of auditory and visual perception.

c) Properties of the interactional site, such as opportunities for and constraints on movement, visibility and audibility (Mondada, 2009).

d) The prospective speaker’s entitlement to take the turn, which may depend on being selected as a next speaker, but also on (topic-, activity-related, etc.) speaking rights arising from identities and knowledge relevant to the interaction (Enfield, 2011).

This paper is to show that turn-construction is multimodally organized and thus needs to be studied taking multimodal interaction into account. Multimodal turn-construction crucially involves talk, but additionally can imply the relevance of gaze, gesture, facial expression, bodily posture and movement in space, and manipulation of objects (cf. Goodwin, 2000). Taken together with properties of what the upcoming turn is to accomplish, the factors sketched above inform four tasks participants have to deal with in order to produce a turn which precisely fits the interactional moment in which it is to be placed. The four tasks are:

1. Achieve joint orientation: The accomplishment of the interactional and spatial prerequisites necessary for producing a turn which is to become part of the participants’ common ground.

2. Display uptake: The retrospective and responsive design of the turn with respect to the current state of the interaction, in particular, relating to the immediately preceding turn of the prior speaker. Next speaker needs to display his/her understanding of the interaction so far as the backdrop on which the production of the upcoming turn is based.

3. Deal with projections from prior talk: The speaker has to deal with projections which have been established by (the) previous turn(s) with respect to the next turn(s), i.e., which concern the upcoming turn s/he is about to produce.

4. Project properties of turn-in-progress: The speaker needs to orient (the) recipient(s) to properties of the turn s/he is about to produce.

Turn-design thus can be seen to be informed by tasks arising from the contingencies of the online-construction of turns (Auer, 2009a; Linell, 2009). These contingencies are:

a) Interactional: Turns are designed for (multiple) participants who (have) perform(ed) activities and can be taken to expect and know certain things and to take certain stances. Recipients’ actions are continuously monitored and taken into account by adaption of the construction of the turn-in-progress,

b) Embodied: Turn-construction is both responsive to and uses spatial, kinesic, tactile, visual and auditory affordances and constraints of the participants’ bodies.

c) Multimodal: Turn-construction allows for, and sometimes requires, various modal resources for action and their simultaneous and sequential coordination (see Goodwin, this volume).

d) Temporal: Turn-construction is inevitably situated at interactional moments which are characterized by ‘unique’ interactional, spatial, bodily and linguistic constellations, providing for specific constraints and opportunities to produce a next action.
This paper does not present detailed analyses of one interactional practice. Instead, I discuss a large array of interactional practices, in order to obtain a larger picture of tasks of the construction of turn-beginnings and the practices to deal with them. The discussion of single practices will necessarily be restricted to their relevance for the tasks of turn-construction discussed in this paper, abstracting from much detail which is relevant for the precise contextual production of the turns discussed. In the same vein, data extracts presented serve as typical instances of major practices studied elsewhere in appropriate detail already.

3. What is a turn-beginning?

Since this paper is on turn-beginnings, a few words on the concept ‘turn’ are in order. I take it that ‘turns’ and ‘actions’ have to be differentiated form each other: the concept ‘turn’ refers to vocal/verbal activities; actions can be accomplished by turns, but also nonverbally. Thus, e.g., a conditionally relevant response can be performed by a nonverbal action which is not a turn, like answering a polar question by a nod or head shake, or complying with a request, like passing a glass of water having been asked for it (cf. Levinson, in press). It is evident that basic phenomena and properties of turn-taking, such as ‘overlap’, the existence of a ‘floor’, the exclusively linear order of turn-constructed units produced by one speaker, do not apply to nonverbal activities, which are non-vocal and can be performed simultaneously (cf. Schmitt, 2005). Thus, I hold that there are no such things as ‘nonverbal turns’, although, as I will show, turn-construction is multimodally informed. I am well aware that this claim is debatable and requires more detailed justification, which is beyond the scope of this paper.

According to this conception of ‘turn’, when thinking about turn-beginnings, the point of departure is the question what a participant needs to do if s/he starts to produce a vocal contribution to the interaction. This means that turn-continuations, such as new TCUs of multi-unit-turns and increments, do not belong to the scope of this paper, unless they are produced by a second speaker (like collaborative completions). Turn-continuations after continuers are not regarded as turn-beginnings, if the speaker continues his/her line of action. In contrast, when response signals (particles, but also facial expressions, etc.) are taken up with a display of a specific interpretation of the response, i.e., by taking note of the evaluative or emotional stance of the response, via self-repair, clarification, emphatic insistence, etc., then the continuation amounts to a new turn which treats the recipient’s intermediate action as a turn itself. The same applies to same-speaker increments after turn-transition has occurred.

Vocal activities which are not ‘full’ turns are also included in this study, i.e., continuers and other response tokens which do not consist of more than one (short) TCU (mostly particles or subphrasal structures), which imply no claim to speaking rights beyond their production and which in many cases are not produced in precise coordination with TRPs (but see Barth-Weingarten, 2009). Nevertheless, they are crucial for interactional participation and the emergence of interactional structure and intersubjectivity (Schegloff, 1982; Stivers, 2008).

In this paper, the notion ‘turn’ will be used with respect to vocal activities with the above properties. Still, turn-beginnings are often hard to delimit. Whereas pre-turn in-breath, throat clear or lip parting, which prepare for the turn and index the claim to speaking rights, clearly belong to the pre-beginning phase (cf. Schegloff, 1996:92f.), other vocal displays, such as sighing, laughter or crying, rather are parts of the turn proper, because they already contribute to what the incipient turn is doing. So, while pre-beginnings need to be attended to when we inquire into how turns are launched in interaction, they are not considered as being parts of the incipient turn.

Demanding that a turn must include a vocal activity does not mean that turns are only vocal phenomena. Just to the contrary, it will be shown that turn-beginnings are multimodally organized phenomena. This implies the relevance of basic properties ranging beyond the organization of verbal conduct, such as simultaneity of embodied resources, the relevance of enduring configurations of space and displays of (bodily) participation, and non-identity of temporal extensions of verbal and related kinesic activities (cf. Oloff, this volume).

4. Vocal activities at turn-beginnings

It is trivial to note that turns a) do not begin in a uniform way and b) often do not contain sentences. Still, the structures which are precisely used to begin a turn have barely become an object of linguists’ attention (but see Auer, 1996b; Lindström, 2006). A first step indicating patterns of interest being in need for explanation may be to chart which kinds of vocal and verbal structures are produced at turn-beginnings.

German talk-in-interaction, randomly chosen from six interaction types (10 min each), was coded in terms of the phenomena produced at the turn-beginning of all turns occurring. The interaction types were: biographical research interview, doctor-patient interaction, behavior psychotherapy, emergency drills for paramedic staff, table football match, conversations among adolescents. The choice of the interaction types took factors into account which are known to impinge on structures of turn-construction. Thus multi-party vs. dyadic, institutional vs. leisure-time interactions and interactions
Non-verbal vocal activities:

- Laughter: 23
- Other (sighing, crying, sobbing, click on the tongue): 8

Verbal activities:

- Discourse particles: (hm, mhm, ja, nein, okay, gut, genau, achso, hoah, oh, pf, ts, ah, ach, naja): 427
- Discourse markers (also, ich mein, (h)ey, äh): 75
- Connectives (und, aber, weil, obwohl, wobei, dass): 29
- Question particles only (Was? Wie (bitte)? Wer? Ha?): 6
- Names only (vocative): 14
- One word-Imperatives/Exclamatives (komm, schau, auf): 11
- Anacolutha: 43
- Ellipses: 182
- Sentences: 149

Fig. 1. Distribution of vocal structures at turn-beginnings.

(a) Black and white

(N = 907 verbal activities)

- Sentences 16.4%
- Ellipses 20.0%
- Anacolutha 4.7%
- Particles 56.2%
- VOC, IMP, EXCL 2.7%

(b) Colors

(N = 907 verbal activities)

- Sentences 16.4%
- Ellipses 20.0%
- Anacolutha 4.7%
- Particles 56.2%
- VOC, IMP, EXCL 2.7%

Fig. 2. Relative frequency of verbal structures at turn-beginnings.

with focus on talk vs. focus on kinesic action were included. The 6 × 10 min contained a total number of 907 turns (as defined in section 3). Figs. 1 and 2 show the distribution of phenomena at the beginnings of the turns in the sample.3

Figs. 1 and 2 show that the overwhelming number of turns begins with vocal structures which are specialized in retrospection, i.e., they are responsive (response and question particles, some vocal sounds), relate the incipient turn

3 Each turn was only counted once, and only the first token occurring in the turn was counted. Within clusters of consecutive devices (like, e.g., 'laughter + response particle + discourse particle + sentence') at the turn-beginning, only the first item ('laughter') was coded. Coding categories refer to the maximal vocal/linguistic structure started at the turn-beginning. Ambiguous and incomprehensible cases were not considered.
to prior talk (discourse markers, connectives) or syntactically build on prior talk (most ellipses). This gross finding already indicates that responsivity to prior talk seems to be a major concern when beginning a new turn. However, in order to see how individual linguistic and other multimodal practices of turn-beginnings are designed to deal specifically with tasks of turn-construction, a close look at details of the design of turn-beginnings in their sequential context is needed.

5. Four tasks of turn-construction at the turn-beginning

This section is to substantiate the claim that there are four tasks speakers have to deal with when designing a turn in interaction:

1. Achieve joint orientation to the upcoming turn
2. Display uptake
3. Deal with projections from prior talk
4. Project properties of upcoming turn

We will deal with the four tasks in turn, discussing their interrelationships and pointing to multimodal practices specialized in dealing with the tasks.

5.1. Task 1: Achieve joint orientation to the upcoming turn

In order to produce a turn, prospective speakers have to do more than just wait to be selected as next speaker or use an opportunity for self-selection. The production of a turn designed to be taken up by one or more recipients needs multimodal preparation. Most importantly, this involves to secure the recipient's attention and to arrange a joint interactional space, i.e., the bodily co-orientation of the participants in the way necessary to collaboratively perform the tasks the upcoming turn is to accomplish (Mondada, 2009 and this volume).

How these tasks have to be achieved on a turn-by-turn basis will be discussed in this section by reference to extracts from emergency drills, an advanced training of professional paramedics to improve their skills of providing first aid to injured patients (see Fig. 3). This kind of data has been chosen, because emergency interactions make prerequisites for the production of a turn empirically visible which may remain hidden in situations where participants are more or less continuously and statically oriented to each other and (basically) to talk only (like conversing while sitting at a table, especially in dyadic interaction, or phone-calls). In contrast, emergency drills are multi-party interactions in which participants need to shift bodily positions in fast succession. The participation framework often changes in terms of who addresses whom and who is interactively available for whom. In addition, there are multiple, changing visual and kinesic foci. Therefore, both attention to talk and kinesic action and attention to possibly changing interactional partners have to be coordinated and updated continuously. Because of these properties, participants regularly have to perform multimodal activities in the service of bringing about the intersubjective prerequisites for the reception of turns to be produced in a very articulate way, not being able to just build on already established arrangements of intersubjective co-orientation (Figs. 4 and 5).

Availability of addressees for turns to be produced has to be sought for by specialized activities like in ex.1, where assistant 1 (A1) requests a compression bandage from assistant 2 (A2).
Fig. 4. A1 looks at A2 (hidden behind the officer in chief) (line 025).

025 A1: "<<looks at A2>> Tilmann,
Tilmann

Fig. 5. Mutual gaze between A1 and A2 (line 028).

027 A1: [<<addresses A2, all, p> gib mir nen druck;>]
give me a pressure

028 A2: [(looks up to A1)]

029 A2: <<looks to A1> brauchst DRUCK?> {---}
Need-2SG pressure?

Monitoring (see M.H. Goodwin, 1980) A2, A1 notices that A2 (hidden after OC in the still) is not available as a recipient, because he is occupied with the first-aid kit on the floor. In order to get A2's attention, A1 addresses him by a vocative ("Tilmann", line 025). Only when A2 looks up and establishes mutual gaze with A1, A1 goes on to produce his request (line 029). Before the request turn is produced, A1 thus takes care that interactional conditions for recipiency are established. While mutual gaze is treated as evidence for the relevant condition being met, the vocative is used as a servo mechanism to elicit the prospective recipient's alignment. We can note a systematics of how different modalities become relevant in establishing the conditions for producing a turn: The vocal-auditory mode is used to attract the addressee's attention while he is not available (and not easily reachable) by visual means, moreover, aided by the property of the vocative to pick out the

---

All excerpts are transcribed according to GAT2 (Selting et al., 2011).
intended recipient by his name; the kinesic-visual mode is providing the necessary infrastructure of reciprocal availability for turn-production and turn-reception.

Earlier, the patient (PA) had started telling the story of how he fell onto the stairs where the emergency team found him and where he is still lying. A1 notices that the patient is moving his head during his narration and requests him to keep his head steady.

*Ex.2 Emergency drill IDS-FOLK-Rettung3, 00:09–00:15*

008 ich hab die LAMP gemacht ghabt;=
   I attached the lamp

009 OC: "<<aspirated>aha;=
   oh
010 PA: =dort owwe und- ((points shortly downwards))
   up there and
011 dann- ((points shortly downwards))
   then
A1 does not address the patient as soon as he notices that the patient moves his head, which could be dangerous for the patient, given that he could be concussed. A1 first moves into the peripheral visual field of the patient (Fig. 6) and only then addresses him with the request to keep his head steady (line 013). Simultaneously with the onset of his request, A1 starts to fix the patient's head with his hands in an axial position (Fig. 7). Again, we can see that the request-turn is only produced after the multimodal prerequisites for recipiency (and correct understanding) are established. Mutual visual availability is again treated as the crucial precondition, which in this case demands from the producer of the prospective turn (A1) that he re-arranges interactional space so that he himself becomes visually available for the patient first. By producing simultaneously the request and fixing the patient's head, A1 makes both actions comprehensible, the verbal action giving the account for the kinesic operation, while the latter clarifies in which position the patient's head has to stay. Also, by establishing haptic contact with the patient, A1 makes clear that the request is addressed to him.

The establishment of joint interactional space thus is crucial not only for getting the turn and for securing recipiency, but also to provide for the necessary grounds for specific upcoming joint actions and for enabling spatial references to be made in the incipient turn (Mondada, 2009; see below). It can easily be overlooked that these multimodal prerequisites for turn-construction are inevitable. It is clear that these tasks have to be accomplished in (and before) the (official) opening of an interaction: spatial arrangements have to be coordinated; an "f-formation" (Kendon, 1990), mutual availability, attention and readiness for interaction, self-identification and recognition, etc. have to be accomplished (Mondada and Schmitt, 2010; Schegloff, 1972). Still, it is not so clear that these prerequisites have to be (re-)established for every next turn, because often there are no observable activities which are specifically designed to achieve them. Instead, next turns can build on coordinations achieved earlier, which, however, to need endure and to be adapted continuously in order to provide for the necessary grounds for upcoming turns. Most evident, however, is the orientation to joint attention in the coordination of gaze (Ch. Goodwin, 1980, 1981; Rossano et al., 2009): Incipient speakers try to catch the gaze of the intended recipient(s), which counts as a warrant for his/her attention to the upcoming turn. If future recipients are not already prepared for an upcoming turn, incipient speakers use specialized devices to claim the turn and to establish the necessary prerequisites of joint orientation to their upcoming turn.

- Various kinds of body movement can be used to attract the recipient's attention. Most importantly, prospective speakers move their body into the visual field of the prospective addressee, e.g., by leaning forward (Mortensen, 2009), pointing (Mondada, 2007), hand-raising (Sahliström, 2002; Mondada, this volume), or walking into sight (Deppermann et al., 2010), they touch the addressee (see above), or they manipulate objects in a way as to get the addressee's attention, such as when objects are associated with entitlements to speak, like a folder in a meeting (Deppermann et al., 2010). These devices can already be used while a prior turn is still underway in order to self-select as a would-be next speaker without overlapping or interrupting the current speaker (Mondada, 2007; Streeck, 2009).
- Audible in-breath is a phenomenon which is physiologically associated with preparing for speaking. It thus can be used as a conventional means for turn-claiming, because it signals that the speaker will take the turn immediately (Schegloff, 1996; Mortensen, 2009).
- There are also vocal and verbal means summoning the recipient's attention to the turn to be produced. Hesitation markers like *uh(m)* can be used for this business (Schegloff, 2010), attention- getters like *(h)ey* (Schlobinski et al., 1993) are more exclusively specialized to do the job, while vocatives (Schwitalla, 1995) are explicit means to pick out the addressee distinctively in multi-party-situations.

Turn-beginnings thus often need to be prepared by pre-beginnings (Schegloff, 1996:93f.), which are not part of the turn proper, but project its imminent onset. Still, the use of vocal means to establish the turn shows that the construction of turn-beginnings can be a self-reflexive enterprise, i.e., the turn-beginning is designed so as to achieve the grounds which are necessary for a successful intersubjective uptake of the emerging turn itself. Because of this reflexivity, it would be wrong to understand "prerequisites for turn-construction" as something which necessarily has to be in place before the turn starts or which, by definition, happens outside of the turn. Rather, because sounds cue attention, to start a turn is itself a major resource for securing the grounds for its own continuation. Just because of this
auditorily based organizational reflexivity, vocal materials produced in service of establishing the upcoming turn tend
to be either non-semantic, recycled, cut-off or specialized to attention-getting (like vocatives and attention-getters). So,
in this case, the initial stretch of a turn exhibits characteristics which make clear that it has this organizational function
of achieving joint orientation to the turn-in-progress.

The relevance of the task of securing joint orientation to the upcoming turn can most clearly be seen in deviant cases,
i.e., if a turn loses or did not gain its recipient. Hesitation, the production of discourse particles like uh(m), self-repairs,
and restarts are routine practices for attracting the recipient's attention, mostly involving a recycling of unattended
actions and semantic materials (Ch. Goodwin, 1980, 1981). The same applies to competitive turn-begins (cf.
Schegloff, 1987; French and Local, 1983), which in this respect suffer from the same problems of securing recipient's
uptake.5

The task to “achieve joint orientation to the upcoming turn” is basic for every turn which is to become part of an
interactional sequence. Still, the multimodal methods for achieving joint attention, the visual checks available, and
the relevance of spatial arrangements to be made in order to fulfill the task depend on the participation framework
(most importantly, dyadic vs. multi-party interaction), on spatial configurations of the participants and relevant objects
and on the mode of interaction (face-to-face vs. only auditive). Therefore, in dyadic telephone conversation, joint
attention matters equally for any turn to be taken up; however, it cannot be checked and established beforehand by
visual-kinesic practices, but needs to be presupposed unless communicative failure proves the presupposition to be
wrong.

While the activities discussed so far specifically deal with the task to secure joint attention, there are still other activities
designed to arrange interactional space in the way it is necessary for the upcoming action. In particular, this applies to all
activities which prepare for necessary shared visual fields for spatial reference. Mondada (2009) shows how participants
simultaneously delay their turn-production by self-repairs, hesitation markers and the like while moving their bodies into an
intersubjectively coordinated spatial position which allows them to refer successfully, i.e., recognizably for the listener, to
some spatial target.

5.2. Task 2: Display uptake

Turns are not only context-renewing, but also context-bound (Heritage, 1984a:241). Producing a turn for precisely this
interactional moment implies to show how it fits the interactional process so far. This involves three closely related things:

• The speaker needs to show that s/he has taken up his/her partner's prior activities and, more importantly,
• how s/he understands them;
• s/he needs to show how the upcoming turn relates to what has preceded.

To be sure, ‘understanding’ is not a homogenous notion. It can concern very different facets of prior turns, such as
reference, propositional content, allusions, kind of action, expectations regarding turn-continuation and recipient
reactions, interactional key, emotional and epistemic stance-taking, etc. Understanding displays typically treat several of
these facets simultaneously, albeit to different degrees of obviousness.

In social interaction, the prior turn is the natural point of departure for the next turn: by default, next turns are treated as
building on the speaker's understanding of the preceding turn, thus displaying this understanding to interactional partners
(Sacks et al., 1974; Schegloff, 1992). Understanding operates according to a logic of local contingency. This temporal
default logic of understanding in interaction has several corollaries:

• If the upcoming turn is not to be understood as relating to the immediately preceding turn, speakers need to use
specialized devices, such as misplacement markers (Sacks and Schegloff, 1973) and back-linking-devices.
• Next turns need not exhibit specific design features (like connectives, pro-terms referring back to the prior turn) in order
to show that they relate to the immediately preceding turn and not to some other prior activity.
• Speakers can, but need not, use devices which are specialized in displaying understanding. Any next action can, and
often needs, to be interpreted as displaying an understanding of the prior turn in terms of sequentially provided
pragmatic prerequisites needed for the current turn to be producible, responsive, relevant at that moment of the
interaction, make sense, etc. (Deppermann, 2008). In this way, a next action embodies implied understandings.

5 In addition, they are characterized by increased volume and high pitch. Systematic consideration of prosodic and phonetic details of turn-
beginnings is beyond the scope of this paper (see Couper-Kuhlen, 2001; Walker, 2004).
Because of this default conditional relationship between prior turn and next turn, which is constitutive of the interactional meaning of a next turn, listeners can be expected to infer which kind of understanding of a prior turn any next turn exhibits, even if there is no specialized index or even explicit uptake of the prior turn.

Implied understandings are regularly displayed in second pair parts of adjacency pairs, such as answers, acts of acceptance/denial, or (dis-)agreements. Producing the conditionally relevant next action counts as a sufficient display of understanding of the first pair part, because the raison d'être of the first pair part lies precisely in eliciting the appropriate response. An example is ex.3, where the patient starts telling how the accident happened.

Ex.3 Emergency drill IDS-FOLK-Rettung3, 00:03–00:11

003 OC: (--) was is pasSIERT?=
what happened
004 Al: <<looking at the patient lying on the floor>> Oh!>
oh!
005 PA: <<all,pp> ich wei-> (-) I kno
006 PA: <<all> ich WEIß net mehr ich bin->
I don't know anymore I am
007 (---) <<points upwards> ähm-> (---)
uhm
008 ich hab die LAMP gemacht ghabt;=
I attached the lamp
009 OC: aha.
010 dort owwe und- ((points shortly downwards))
there above and
011 dann- ((points shortly downwards))
then
012 Al: [[<positioned left to PA, fixes PA's head with both hands> halten sie den köpf mal ruhig [BIIItte;>]
keep your head PRT steady please
013 PA: [bin ich= (-) RUNner gfAlle.
I fell down

The patient's story does not explicitly display how he understood OC's question “was is pasSIERT?” (what happened, line 003), which as such is a generic formula, in terms of being a situated action (e.g., that the patient should deliver an account why he fell onto the stairs, impart facts which might give hints to his medically relevant condition and treatment, etc.). Still, these understandings are implied in his response, i.e., by producing a turn which can count as an answer, by selecting story-components which account for what he was doing before he fell, and by beginning his turn with an epistemic disclaimer (lines 005–006). This last component comes closest to an explicit display of understanding, because it exhibits how the patient's turn is recipient-designed in dealing with the expectation attributed to the questioner (OC) that he should give a precise and correct account.

Still, in absence of such clear cases of pragmatic dependency of adjacent turns, participants can resort to other, less conventional resources for inferring understandings implied by follow-up turns. Gricean maxims (cf. Grice, 1989; Levinson, 2000) are a major resource to draw such inferences about understanding and coherence in interaction.

Specialized displays of understanding are only necessary a) if a participant restricts him/herself to a merely responsive activity, b) if issues of understanding itself become the focus of the talk, c) if the specific way a turn-in-progress relates to what has happened before is to be displayed. We will deal with these cases in turn (see Deppermann and Schmitt, 2009).

(a) There are various response practices whose primary job is to display acknowledgement/attention, understanding, evaluation, and emotional or epistemic stance to prior turns. For this, discourse particles are used most regularly, their lexical, phonetic, and prosodic variants being more or less specialized in displaying different kinds and facets of understanding, such as change of state of information, epistemic superiority, scepticism, etc. (see e.g., Heritage, 1984b; Goodwin, 1986; Gardner, 2001; Jefferson, 2002; Fischer, 2006). Already Schegloff (1982) pointed out that they often do not only operate responsively, but also prospectively, e.g., as continuers or as bids to close a topic or a sequence (cf. Golato and Fagyal, 2008; Barth-Weingarten, 2011). Head nods and facial expression are additionally used, often tightly synchronized with verbal tokens (Bavelas and Gerwing, 2011). Although response tokens do not demonstrate but only claim understanding (cf. Sacks, 1992), they nevertheless display a basic analysis of what the partner's turn does, thus
showing the recipient's alignment with the speaker's activity (Stivers, 2008). This can be seen in ex.4 from a televised talk-
show. The guest KM tells how he became a presenter for folk music programs.

*Ex. 4 Talkshow IDS-AGD 4050.08, 00:16:00–00:16:21*

| 01 KM: und da sin ma mal EURTCHbar BEchern gegangen,= | 02 WB: mh; mhm |
| 02 =und da war ich> VOLL: beSOFFen= | 03 KM: | 04 -) und WÄHrend i so bSOFFen war muss ich irgendwas> |
| 03 da hat er mich> angFÜLLT? he filled me up | 05 KartSCHRIEBen haben und; signed something and |
| 04 (-) und WÄHrend i so bSOFFen war muss ich irgendwas> | 07 (-) <<p>der GRUNDstock dass ich;> |
| 05 unterSCHRIEBen haben und; | 07 the foundation that I |
| 06 (-) DES war eigentlich- that was actually | 08 WB: mh; mhm |
| 07 (-) <<p>der GRUNDstock dass ich;> | 09 KM: präsenTator fur volkstümlicheSENDungen war; (-) |
| 08 WB: mh; mhm | 10 WB: [mhm; ] |
| 09 KM: | 11 KM: [<<all>jetz] wird er glei wieder sogn> sigst jetz wo er HERkimmt; now he will say immediately again do you see now where he comes from |
| 10 WB: [mhm; ] mh | 12 WB: <<laughing>mh mh [.h MH,> |
| 11 KM: [<<all>jetz] wird er glei wieder sogn> sigst jetz wo er HERkimmt; now he will say immediately again do you see now where he comes from |
| 12 WB: | 13 KM: [<f>vertSTEHST??>] |
| 13 KM: | 14 WB: (-) [ja=] |
| 14 WB: (-) [ja=] | 15 KM: [S:0] fing es an. that's how it started |

The host (WB) shows his analysis of KM's story both by the sequential position and the vocal and lexical design of his responses. In lines 08 and 10, he produces acknowledgement tokens which are precisely timed with regard to the resolution of KM's story, i.e., the recognition of the concluding point (line 08) and of its closure (line 10). As KM adds a jocular self-ironic comment (line 11), WB aligns with this keying by responding with laughter particles (line 12). KM's overlapping prompt to return to a serious mode of dealing with the story (line 13), is again aligned with by WB producing a ratifying response token (line 14).

By their position in the sequence, their phonetic and prosodic design and their lexical type, response tokens both acknowledge the packaging of the partner's turn in terms of units of topical and pragmatic boundaries (such as completion of actions, relevant new information, punch lines, cf. Barth-Weingarten, 2009), and they can additionally display (empathic) uptake in terms of (sharing) emotional and evaluative stance and keying. Vocal displays (like laughter, sighs, etc.; see, e.g., Jefferson, 1984, 1985; Glen, 2003; Hepburn, 2004) are also often responsive. Indeed, the dividing line between (non-linguistic) vocal and verbal displays is somewhat arbitrary, because there are many "sound objects" (Gardner, 2001; Reber and Couper-Kuhlen, 2010) which have regular interactional uses, but which are not codified by grammars and dictionaries. This is also because phonetic and prosodic variation defines their function and meaning to an extent and in detailed ways which do not apply for the rest of the language.

Stivers (2008) shows that there may be a multimodal division of labor between acknowledgement tokens and nodding, the former displaying alignment of the recipient with the speaker's activity of story-telling, while the latter are reserved for the display of affiliation. While many response tokens are "generic" (Bavelas and Gerwing, 2011), i.e., merely registering the partner's turn and acting as continuers (Schegloff, 1982; Goodwin, 1986), facial expression displays evaluative and epistemic stance more specifically (Chovil, 1991; Peräkylä and Ruusuvuori, 2006). However, Bavelas and Chovil (2000) insist that facial displays become definite "acts of meaning" in most cases only if produced together with verbalizations.

Response tokens as such do not claim speaking rights beyond their production. If they establish some expectation for a next turn (like continuation of the partner's turn, topic-transition), then it is much more vague and less accountable than with full turns which establish some definite conditional relevance. Response tokens, however, often occupy turn-initial positions (cf. Sacks et al., 1974:722ff.). They may constitute a TCU of their own or precede the first TCU in the turn without being TCUs themselves, thus constituting presegments at the turn-beginning before the contribution the turn is to make
starts (Lindström, 2006; see also Schegloff, 1996). This can be seen in the next ex.5 from a psychotherapy session. Before the extract, the patient (PA) made clear that the therapist (TP) misunderstood that the patient has thought about suicide. In the transcript, therapist’s head-nods are indicated by %.

**Ex.5 Behavior Psychotherapy Ewald 22:20–22:54**

01 PA: des WAR im; (-) zusAMmenhang mit de:n (-) medikaMENte abzuSEtzen;
   this was in connection with dropping the drugs
02 TP: <p,creaky>o%KAY;
   okay
03   JA,
   yes
04   (-) dann HATT [ich_m ] (j)ah;
   so I had it (yes)
05 PA: [(eben)];
06 PA: (-) <<p,creaky>äh> <<dim>OB halt da des LEben noch (unmittelbar)>
   whether life was still immediately
07   (-} ah,
   uhm
08   (-) LEbenswert %wä%re %JA,
   worth living right
09 TP: % mh-
   uhum
10 PA: (-) <<all>das war> in DEM zusammehang mit den medikaMEN%te <<pp>ja.>
   this was in connection with the drugs
11 TP: % mkay- (-) SCHÖN dann muss ich das [korrigIERen.]
   okay fine so I will have to correct this
12 PA: [aber so- ] (-)
   but like
13 so dass ich jetzt (0.8) mir jetzt das LEben nehmen würde,=
   like that I would now take my life
14 =ne des net;
   no it’s not like that
15 TP: % mh- (0.7)
   uhum
16 mit der reaktion ham se jetzt meine frage auch schon beANTworet;
   with this reaction you have already answered my question

In lines 02, 03, 11, and 15, the therapist produces response tokens every time before he formulates upshots from the patient’s prior turn for his understanding of the psychological state the patient is currently in. Nods are closely synchronized with vocal response tokens: Nods regularly immediately precede them (lines 03, 09, 11, 15) and co-occur with them (line 02). Sometimes the therapists nods already as soon as the patient produces the lexical item with the main informational weight and prosodic stress in a TCU (lines 08, 10) to display that he registers the pragmatic point of the turn. Since this is a sequence in which a misunderstanding is repaired, the therapist seems to use response tokens more frequently and with additional care in order to index in an observable manner most systematically that he registers new information from the patient and takes it up correctly.

Responsive practices tend to occupy turn-initial positions, if not the whole turn. Responsive practices account for the overwhelming share of the large amount of turn-beginnings, turns and vocal activities without full turn-status which do not begin with sentential nor even phrasal linguistic structures (cf. section 4). If specialized responsive tokens are produced, they mostly precede progressive elements.

(b) Convergent with the finding that displaying understanding is a fundamental requirement for the construction of a next turn which has to be fulfilled either before or, implicatively, via the production of an expectable next action, Schegloff et al. (1977) point out that there is a preference for other-initiation of repair and other-repair to be produced immediately after the trouble source turn, i.e., before anything new is introduced. Since repair deals with problems of hearing and understanding, it is a basic operation of displaying understanding, destined to achieve intersubjectivity. Interestingly, repair operations also mostly use non-sentential forms like (elliptical) lexical substitution (Rae, 2008), question particles
(Drew, 1997), and (elliptical) partial repeats of prior turns (Selting, 1987; Schegloff et al., 1977). Depending on prosody, these devices are used to challenge, to check or to display understanding.

Challenges deal with problems involving the frustration of expectations (Selting, 1987). Open-class repair initiators (Drew, 1997) can be used to do the job, mostly with a distinctive prosody (Selting, 1995), but there are still other minimal devices which operate on lexico-syntactic grounds. Examples can be seen in ex.6 from an inbound-call to a call-center. The customer (CU) asks the call-center agent (AG) for a new modem, because he claims that the one he received from the company does not perform at the transmission rate it should.

**Ex.6 IDS Call-center 231 01:51-02:16**

<table>
<thead>
<tr>
<th>Line</th>
<th>Transcript</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>CU: (.) das heißt ich brauch eigentlich n neues MOdem; this means I actually need a new modem</td>
<td></td>
</tr>
<tr>
<td>02</td>
<td>AG: (-) ja das kann ich SO nich stehend frEIhändig einfach so ausLÖsen- yes I cannot release this (order) directly simply like that</td>
<td></td>
</tr>
<tr>
<td>03</td>
<td>(0.9)</td>
<td></td>
</tr>
<tr>
<td>04</td>
<td>AG: °h herr kunde° h° mister customer</td>
<td></td>
</tr>
<tr>
<td>05</td>
<td>(0.5)</td>
<td></td>
</tr>
<tr>
<td>06</td>
<td>CU: SONdern, but</td>
<td></td>
</tr>
<tr>
<td>07</td>
<td>(0.6)</td>
<td></td>
</tr>
<tr>
<td>08</td>
<td>AG: bitte? Pardon?</td>
<td></td>
</tr>
<tr>
<td>09</td>
<td>CU: (-) &lt;&lt;h&gt;SONdern?&gt; but?</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>(-) was DENN, what else?</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>AG: 'NEE! deswEgen sag ich jetzt- (.) no it is because of this that I say now</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>wenn es jetzt dAmit im Zusammenhang mit der geschWINdigkeit ist- if it is in connection with the speed</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>°h dann führe ich die notwendigen TESTS durch; gegEbenenfalls (.). then I will operate the necessary tests if required</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>AG: [äh (-) wEIterleitung an ein] tEchniker- uhm forwarding to a technician</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>CU: [des hat doch schon mal_n techniker geMACHT,] this has already been checked by a technician</td>
<td></td>
</tr>
</tbody>
</table>

The customer uses a grammatical item, i.e., a connective (“sondern”, but, lines 06 and 09) as a prompt (cf. Lerner, 2004) to elicit a self-repair from the agent. “sondern” (but) projects the syntax of a complex adversative sentence. Thereby the customer indexes that he expects the agent not only to produce a rejection of his request (as she did in line 02), but to offer an alternative how to proceed instead.7 The agent reacts to this with the open class repair initiator “bitte?” (pardon, line 08), displaying that she does not understand (or accept?) what the customer is up to with his repair-initiation. After repeating “sondern” (line 09), the customer specifies his expectation by “was denn?” (what else, line 10), making clear that he demands an alternative proposal from the call center agent. Problems of acceptance and understanding which arise from the frustration of expectations regarding the partner's turns are dealt with by devices of (next-turn) repair-initiation, which are produced according to a logic of minimization. One-item constructions (“sondern?”, “bitte?”) are used to signal the problem and – in the case of “sondern” – to project which kind of solution is expected. As the most minimal repair-initiation “sondern” fails to elicit the projected repair, the customer resorts to a next-to-minimal device (“was denn?”, what else?), a question-pronoun referring to the expected content and a modal particle (“denn”) which indexes that it had been due already before (Deppermann, 2009). No sentential or phrasal devices are used in order to display retrospective problems with acceptance and understanding of the partner's turns.

When understanding is checked (ex.1) or acknowledged (ex.2), we find the same preference for minimization of understanding-related activities.

---

6 “kunde”/customer is used as pseudonym for the caller's name.
7 Interestingly, although the adversative conjunct would be a syntactic expansion of the agent's original sentence from line 02, it would pragmatically amount to a self-correction of this sentence.
Ex. 1 Emergency drill IDS-FOLK-Rettung3, 00:22–00:25

027 A1: [<<addresses A2, all, p> gib mir nen druck;>]
        give me a pressure
028 A2: [((looks up to A1))]
029 A2: <<looks to A1> brauchst DRUCK?> (---)
        Need-2sg pressure?

In his understanding-check, A2 omits the second person pronoun (which does not produce ambiguity, because the inflection of the verb encodes the second person) and the indefinite article A1 had used in his request.

Ex. 2 shows very nicely how the minimal design of responsive, retrospective TCUs rests on adjacency.

Ex. 2 Emergency drill IDS-FOLK-Rettung3, 00:09–00:17

008 ich hab die LAMP gemacht ghabt:=
        I attached the lamp
009 OC: =<<aspirated> aha:=
        oh
010 PA: =dort owwe und- ((points shortly downwards))
        there above and
011 dann- ((points shortly downwards))
        then
012 A1: [<<positioned left to PA, fixes PA’s head with both hands>
013 halten sie den köpf mal ruhig [BITte;>]
        keep your head PRT steady please
014 PA: [<<positioned left to PA, fixes PA’s head with both hands>
        [bin i]ch- (-) RUNner gfAlle.
        I fell down
015 OC: runner gfAlle (.>)
        fell down
016 <<looks and points upwards with his right hand,
017 then grasps PA’s left wrist with his left hand >>>
018 !GANZ! Owwe ham> sie die lamp geMACHT ja?
        all on top you attached the lamp right?

The officer-in-chief (OC) first repeats the final TCU of the patient's small story of how he fell onto the stairs (“runner gfalle”, fell down, lines 014 and 015). This acknowledgement is elliptical, but it is not most minimal, because it recycles the phrasal verb from the patient's turn (instead of only producing an acknowledgement token like uhum). This non-minimal device does not only claim understanding, but demonstrates what was heard via repetition.8 This suggests that the accomplishment of mutual understanding is taken to be potentially fragile. Indeed, the continuation of OC's turn provides further evidence. OC points upwards before he reformulates the first segment of the patient's story in terms of a slightly more precise localization9: “ich hab die lamp gemacht dort owwe” (PA in 008 and 010) vs. “ganz owwe ham sie die lamp gemacht” (OC in line 018).

We can see here that in order to check understanding, the most minimal form possible, i.e., a repetition of the most important content words, is only used if the antecedent is adjacent (“runner gfalle”, fell down, lines 014 and 015). In contrast, a sentential format (like “ganz owwe ham sie die lamp gemacht”, OC in line 018) is used if the check traces back to a non-adjacent antecedent, which was not produced in an immediately preceding TCU. If such 'long distance' retrospective understanding displays are to be made, they cannot rely on the principle of local contingency, but antecedents need to be recontextualized by explicit co-reference calling for more elaborate syntactic formats like clauses and sentences.10

(c) Connectives and discourse markers are used to establish retrospective links between prior turns and the turn-in-progress, signaling the relationship of the upcoming turn to what has preceded (Mazeland and Huiskes, 2001; Local, 2004; de Stefani and Horfacher, 2008; Schiffrin, 1987; Deppermann, 2011). Often, they occur within a turn-initial TCU of their own not integrated into the following sentence (Auer, 1996b). Although their function is projective in announcing the kind of action to come, displaying an evaluative or epistemic stance to assertions to be made etc., they operate as well retrospectively by signaling how the upcoming turn coheres with prior turns. They index topic-continuity, -shift, or -resumption, project a formulation of prior turns, establish contrast, add another consonant aspect, etc. An example for an upshot formulation (cf. Heritage and Watson, 1979) is ex.7 from a psychotherapy session. The patient (PA) tells the therapist (TP) how well she gets along with her two daughters.

---

8 Understanding can also be displayed by repeating gestures (see Yasui, in this issue).
9 Note that the pointing gesture precedes the verbal reference by one TCU (cf. section 5.4 below).
10 Such explicit, phrasal and sentential formats of demonstrating understanding occur particularly often if gaining or testing understanding lies at the heart of the interactants' joint communicative project, as in psychotherapy sessions or in classroom interaction.
Ex. 7 Psychoanalysis "angepiekt", 17:41–17:51

01 PA: <<ch nu>> da SETzen wer uns hin und MÄchen das- (.)

02 und das is SEHR schön;

03 .hh und ich glaub das GÄB es nich sOnst.

04 TP: mh:m.

05 mh

06 TP: [mh:m.] mh


08 mh

09 TP: mh:m-

10 mh

11 TP: also da is so_ne verTRAUEnSebene-

12 so there is kind of a level of trust

13 TP: irgendwie so was,

14 PA: .HH <<f> JA:: (.) das IS:es-> (.)

yes there is

Using the turn-initial connective "also" (so, line 11), the therapist (TP) indexes how her upcoming turn relates to the preceding story of the patient, namely, as a notionalization (see Deppermann, 2011) of the upshot of what she takes the patient to mean. Thus, from the very outset of the turn, the therapist makes clear that an understanding display is to follow.

The turn-initial positions of responsive and retrospective devices, their preferences and restrictions as well as the preferences for other-initiated repair in the immediately following turn all attest to a temporal and pragmatic principle, which can be phrased as "understanding before progression". If speakers see the need to deal with issues of understanding, i.e., if understanding becomes problematic or if it cannot unproblematically, self-evidently, be implied in the design of progressive next actions, then understanding has to be dealt with first, i.e., before progression can go on (cf. Heritage, 2007). There is a complex relationship between progression and understanding in interaction. On the one hand, there is a preference for progressivity (Schegloff, 2007), providing for the fact that issues of understanding are dealt with implicitly, by minimal, non-sentential, elliptical forms and en passant (qua implication) whenever feasible, thus decelerating interactional progression the least possible by minimizing understanding-related activities.11 Crucially, this preference for minimizing understanding related activities is aided by the logic of local contingency laid out in the beginning of this section, because it allows participants to rely on default interpretations to a great extent, with regard to both the object of understanding (the preceding turn) and the kind of understanding. On the other hand, the accomplishment of intersubjectivity concerning the interactional past is the prerequisite for progression. Progression has to build on common understandings in order to be based on solid ground and to be accountable and comprehensible to all interactional partners (cf. Clark, 1992). The temporal precedence of understanding displays before progressive activities thus mirrors their pragmatic precedence.

5.3. Task 3: Deal with projections from prior talk

For many turns, to produce an action which fits the precise interactional moment at which it is produced means more than only to display how prior talk is understood. Besides this (and mostly implying it), the upcoming turn has to deal with projections established by prior talk. The creation of projections is one facet of the context-renewing nature of talk: It does not only do something in the moment it is produced, but also often prefigures more or less distinctly what should be done next.

Projections (Auer, 2005, 2009b; Schegloff, 2007) are expectations which concern either the continuation or completion of (the syntax, and/or semantics of) a turn-in-progress or the action type, design, topic, etc. of an anticipated next turn. Projection is a property of turns which is still rather poorly understood in many respects. This is because projection has

11 Types of interactions which might impose as exceptions to this rule at first sight, like psychotherapy sessions, where displays and negotiations of understanding regularly become topical, are actually not deviant, because their primary concern is to gain new, more detailed, etc. understandings of the participants' turns. Achieving, repairing, negotiating, etc. understandings in these interactions thus is not a prerequisite, but the very essence of interaction type task-related progression.
many dimensions, only some of them having been a target of in-depth research. There are at least three properties which
are characteristic of every projection:

a) What is projected?
b) How definite is the projected structure?
c) How obligatory is fulfillment of the projection?

a) What is projected? Projections can operate on several levels. Projections regarding turn-continuation, turn-taking,
action, and syntax have widely been studied (cf. Auer, 2005, 2009b; Selting, 2005; Barth-Weingarten, 2009), and the
interplay of these levels in forming the precise shape of a projection has moved into the focus of CA and Interactional
Linguistics during the last years (see, e.g., Raymond, 2003; Günthner, 2011a,b; Pekarek-Doehler, 2011). Still, there are
also projections regarding topic, genre, key, style, register, and code.

b) How definite is the projected structure? Projections vary in distinctness and granularity (cf. Auer, 2005).
'Distinctness' means that projections can be fixed or rather open. 'Granularity' here relates to the fact that a precise surface
structure, e.g., some lexical element, is expected, or, alternatively, that projections are rather schematic. So, the wedding
vow would be an example for a projection which is fixed and has a low granularity, because it is the precise response formula
that matters. Polar questions and declarative questions are examples of first pair parts which project a type-conforming
response (Raymond, 2003, 2010), i.e., a yes/no-answer or a confirmation respectively. A confirmation from the patient is due
in ex.8, as the officer-in-chief has formulated part of the patient's story (cf. Heritage and Watson, 1979):

Ex. 8 Emergency drill IDS-FOLK-Rettung3, 00:16–00:18

018 OC: !GANZ! Owwe ham> sie die lamp geMACHT ja?
all on top you attached the lamp right?
019 PA: ja;
yes
020 OC: <<looks upwards, len> aha>; (.)
oh

Type-conformity in the fulfillment of a projection in these cases implies to use particles which are specialized to producing
the types of relevant responses. What the particle precisely answers or confirms hinges completely on the turn which has
established the projection, i.e., the local relationship of projection together with the conventional type-conforming response
tokens saves the respondent from formulating the propositional content of the response him/herself.

Type-conforming relationships of a linguistically more complex kind hold between Wh-questions and projected
answers. Basically, the instantiation of the semantic role requested is due, like in ex.9:

Ex. 9 Doctor–patient AA BJ 01 246–247

246 DR: seit WANN wohnen sie alleine?
since when do you live on your own?
247 PA: seit noVEmber.
since November

A type-conforming answer is restricted to adjacent ellipsis, i.e., a phrase or a lexical item which is enough to fill the open
slot of the semantic role indexed by the question pronoun.

In general, the more definite the projection is, the more minimal is the default type-conforming response. It is restricted
to just those item(s) needed to convey what counts as locally relevant "new information" or as the minimally sufficient
responsive action. This observation is reflexive: If a respondent chooses to produce a more-than-minimal response, say, a
full sentence, then this has some pragmatic extra-value, even if it is propositionally equivalent to responding only by a
particle or a phrase.12

Projection is a matter of degree. It can allow for a wide range of options how to satisfy it, and, indeed, this may be
necessary for the projected action to do its job. Paramount examples are elicitations of narratives. They schematically
project a narrative, i.e., a descriptive multi-unit turn. They predefine a more or less broad scope of topics for the narrative
to come. The precise sequence of events, relevant actors and objects, etc., however, are not and cannot be projected, if the
narrative is occasioned in order to tell breaking news. This is the case in the beginning of our emergency drill. Having
arrived at the site, the officer-in-chief asks the patient what happened:

12 Space is too limited to show this in sufficient detail. But see, e.g., Heritage and Raymond (2005), who show that more elaborate second
assessments index a claim to independent or superior epistemic or normative authority in relationship to the first assessment.
**Ex. 2 Emergency drill IDS-FOLK-Rettung3, 00:09–00:15**

003 OC: (-) was is passIERT?=  
    what happened
004 Al: =Oh!  
    oh!
005 PA: <all, pp> ich wei-> (-)  
    I kno
006 <all> ich WEIS net mehr ich bin->  
    I don't know anymore I am
008 ich hab die LAMP gemacht ghabt;=  
    I attached the lamp
009 OC: <<= aspi atrated>> aha:=  
    oh
010 PA: = dort ovwe und- ((points shortly downwards))  
    there above and
011 dann- ((points shortly downwards))  
    then
012 Al: [<< positioned left to PA, fixes PA's head >]  
013 halten sie den kopf mal ruhig [BITte;>]  
    keep your head PRT steady please
014 PA: [bin i] ich (-) RUNner gFAlle.  
    I fell down

*What happened* makes a story relevant which accounts for the patient's current situation. However, the question does not specify which events should be talked about, which actors and objects will occur (apart from the patient), if and what kinds of subjective and emotional assessments will be made, etc.

(c) How obligatory is fulfillment? Projections also differ in terms of accountability and normative force. Fulfillment may be strictly expected (the case of the wedding vow), projected alternatives can be ordered according to preference structures (Pomerantz, 1984; Bilmes, 1988; see below), or projections can open up a (rather fuzzy) range of possible next turns without definite preferences. While first pair parts of adjacency pairs establish rather strict projections with high accountability, second pair parts mostly do not project much more than to return the floor back to the producer of the first pair part. Examples are patients' responses in doctor-patient interactions like in ex.10. Here, the patient closes his anamnestic account with an emotional assessment.

**Ex. 10 Doctor-patient IA_MR_03, 00:48–00:52**

44 PA: ich hab gleich ANGST  
    I immediately get afraid
45 <<dim> wenn ich so was dann HAB,  
    if I get such things
46 ne? n bisschen.>  
    right? A little bit
47 DR: wovor ham sie denn ANGST-  
    What are you afraid of

Following the patient's turn, the doctor may, as he does, produce a follow-up question, but he could also show empathy, produce a next agenda-based question, shift to bodily examination, impart a (preliminary) diagnosis or a prescription, etc. None of these activities is projected in particular by the patient's turn (and all occur after comparable patients' turns). The patient's turn-transition projects only that the doctor performs some professional action s/he considers to be in place at that very moment of the visit.

A note of caution regarding the term 'projection' and its relationship to 'expectation' might be in place with examples like these. To be sure, self-initiated actions like questions or requests often do not come unexpectedly, a fortiori in institutional contexts where they are constitutive of the discourse identities of professional agents performing situated professional tasks. Still, it is almost never the case that the type of action to be taken by the professional, its topic, etc. are projected by prior turns of clients. Rather, there is a general, interaction-type-related expectation that the institutional agent will perform certain actions which are part of his professional repertoire (cf. Heritage and Clayman, 2010), either in certain phases of the interaction or whenever s/he judges them to be relevant. This mutually assumed license makes itself felt by the absence of misplacement markers which display in one way or another that the institutional agent does something which was not projected by the client's preceding talk.

As a rule, dealing with projections implies a display of understanding (see section 5.2): By the way the turn is designed, the speaker displays his/her analysis of the expectations prior talk has created regarding the continuation of the interaction. This in turn almost always presupposes some understanding of propositional content, emotional and
evaluative stance, kind of action and other possibly relevant aspects of the meaning of prior talk, in short, at least of those which lay the grounds for the projections to be dealt with. Therefore, we mostly do not find activities specialized to displaying understanding before dealing with projections, but understanding displays are built into the way the projection is dealt with. Again, this implicational relationship of dealing with projections and displaying understanding reflects the preference for progressivity in interaction.

However, if understanding is not dealt with only by implication, it is displayed at the turn-beginning before the speaker moves on to deal with the projection. A routine way of doing this is to produce response tokens first, before the relevant next action follows. This often occurs if a dispreferred action is to follow. This is the case in ex.11 from a so-called "pitching" session, in which students discuss a self-authored film script with their teacher (TE). He criticizes it harshly, reproaching them to waste too much time before introducing the main characters. The reproach makes a response relevant, preferably acceptance, while a defense of the script is a relevant but dispreferred response.

Ex.11 Pitching Bernd 26:20–27:01

01 TE: äh also frau zur walDAU,\textsuperscript{13}
    uh well Mrs. zur Waldau
02 (-) vergEÜdet erZAHLzeit <<t,p>sozuagen->
    wastes story time so to say
03 (1.2)
04 ST: mh:: ja JA;
    mhm yes yes
05 also wenn man s im großen und GANZen sieht,
    well if you see it in total
06 (--) ich WEISS ja nicht;
    I don't know
07 also wir HAM uns natürlich;
    well of course we have
08 oder ich hab_s mit im hintergrund-
    or I have written this with
09 dieser <<len>struktur?> (--) geschRIebeen;
    this structure as a background
10 die ich WIRKlich ganz spannend finde-
    which I find really very exciting

The student (ST, who is responsible for the script) responds by first acknowledging the reproach ("mh") and producing what can be taken as an agreement token ("ja JA;", yes yes, line 04). However, it turns out that is only a preface to a rejection of the reproach by a justification of the student's concept of dramatic structure. In such non-affiliative cases, prefacing displays of understanding can be used in order to constrain the rupture in intersubjectivity which the non-affiliative, dispreferred action to follow threatens to cause. Apart from this, the dispreferred nature of the response is projected from early on in the turn (cf. Pomerantz, 1984; Schegloff, 2010), indexed by discourse markers ("also", well, lines 05 and 07), cut-off and restart with hesitation (lines 05 and 07) and an epistemic discourse marker formula ("ich weiß nicht", I don't know, line 08). The design of actions as dispreferred displays from the very beginning that the speaker orient to the partner's expectation, i.e., in the case of seconds, the projection established in his/her preceding turn.

The orientation of turn-construction to pending projections can also be seen by the fact that self-initiated expansions tend to follow only after the projection has been dealt with. A common pattern is the expansion of a type-conforming answer or confirmation. In ex.12, we see how a confirmation is expanded by further explanation.

Ex.12 Behavior Therapy Therapist black shirt 18:00–18:22

01 TP: also das meinen sie mit konzentrationSproblemen jetzt.
    so you mean this by concentration problems now
02 PA: ja ja- (--) yes yes
03 also ob ob ob ob ich jetzt wirklich ahm-
    well if if if if I am really uhm
04 (1.0)
05 äh (--) also gib- (--) uhm well there
06 s gibt momente da (kann ich) die wahrheit
    there are moments where (I can)not tell what is true
    von der unwahrheit nich SELber net unterschEi:de.
    from what is not true myself

\textsuperscript{13} "frau zur waldu" refers to a character from the film script.
Note that the patient (PA) does not expand his description right away as response, but he first delivers the projected, type-conforming confirmation (line 02). Thereby, the therapist's (TP) formulation (line 01) is first grounded, before the patient adds further details (lines 03-06), which, without the confirmation, could be understood as correcting the therapist's formulation. The fulfillment of projection thus precedes self-initiated action.

5.4. Task 4: Project properties of the turn-in-progress

The speaker needs to orient the recipient to properties of the turn-in-progress from its very beginning. Most important are those aspects of meaning which matter for adequate understanding and anticipation of the emerging turn-structure. This minimizes the need for later reanalysis and avoids misunderstandings and distractions from the emerging turn. Projection of the emerging turn involves a) establishing topic, b) projecting upcoming action and c) framing the turn.

5.4.1. Establish topic

If a speaker changes topic with respect to prior talk, turn-initial establishment of the new topic is most convenient for both production and reception, because attributions to be made in the turn are predicated upon the topic. Like many other languages, German offers syntactic structures like left-dislocation, free topics, clefts, existential clauses etc. specialized in introducing new topics turn-initially, before predications to the new referent follow. Especially when packaged within a separate TCU, these structures perfectly comply with the 'one new argument constraint' (cf. Du Bois, 2003), which implies that the production of more than one lexical argument per clause is to be avoided. Ex. 13 shows how left-dislocation is used. Denis has talked about a mother ("die", she) who publicly displayed her (alleged) sexual interest in minors by various actions (like singing). In order to introduce a new referent, her younger son, Denis begins the turn with a free lexical NP which establishes the new referent.

Ex. 13 Juk 17-23 “white trash slut”14

22 Denis: die is,=
she was,
23 Alex: =<<laughing> vorm haus gestanden und gesungen.>
standing in front of the house singing.
24 geil,=
 fucking great,
25 Denis: =der ihr kleiner sohn, (.)
her younger son
26 der, (.)
he,
27 als der kleine sohn elf war oder [so when the younger son was eleven or [sc
when the younger son was eleven or [sc
28 Bernd: das
that
29 is die [NOTgeile mutti, ]
is the [DESpolate14 mummy,]
30 Denis: [da is se an mir ] [.] [then she passed ] [.]
31 mit dem an mir vorbei gefahren,
me by with him,

In line 22, Denis starts a pronominal construction which topicalizes the mother ("die", she). He cuts himself off, and after Alex’ turn (lines 23–24), which quotes Denis’ prior report, Denis restarts his turn with altered reference, now introducing the mother’s son ("der ihr kleiner sohn", her younger son, line 25) as new referential topic in a separate TCU. This, however, is not expanded by a predication, but (after a pronominal co-reference “der”, he, line 26) Denis produces a new adverbial sentence (“als der kleine sohn elf war”, when the younger son was eleven, line 27), now attributing a property (age) to him, before shifting back to the prior topic (the mother, “se”, lines 30–31) and relating it to the intermediate topic (“is se mit dem an mir vorbei gefahren”, she passed me by with him). Constructions like the NP “der ihr kleiner sohn” in line 25 are infelicitously termed “left-dislocation”, not doing justice to the fact that they

14 “desperate” here means “desperately searching for a sexual partner”.


systematically arise from incremental online-production (Auer, 2009a): The speaker first establishes the (new) referent as a topic, which allows him/her to decide only afterwards what precisely will be attributed to the referent and how this will be done syntactically. Likewise, the recipient can first ground the new referent before processing additional bits of new information about it in subsequent TCUs. High onsets are an important prosodic practice to mark the beginning of a new, non-projected topic or sequence (cf. Couper-Kuhlen, 2001, 2004).

Pointing gestures are another device for establishing reference as early as possible. They regularly precede verbal reference, orienting recipients to relevant spatial domains and possible referential targets of the turn (Goodwin, 2003; Schegloff, 1984). In ex.3, we see how the patient uses pointing gestures to show where he tried to attach a lamp at the ceiling before falling.

Ex.3 Emergency drill IDS-FOLK-Rettung3, 00:03–00:11

003 OC: (-) was is passIERT?= what happened
004 Al: <<looking at the patient lying on the floor>=Oh!> oh!
005 PA: <<all,pp> ich wei-> (-) I kno
006 <<all> ich WEIß net mehr ich bin-> I don’t know anymore I am

007 (---) <<points upwards> ähm-> (---) uhm
008 ich hab die LAMP gemach ghabt;= I attached the lamp
009 OC: aha. oh.
010 dort owwe und- ((points shortly downwards)) there above and
011 dann- ((points shortly downwards)) then
012 Al: [<<positioned left to PA, fixes PA’s head with both hands>>]
013 halten sie den kopf mal ruhig [BITte;>]
    keep your head PRT steady please
014 PA: (bin i)ch- (-) RUNner gfAlle. I fell down

---

16 Fiehler (2005:1210) rightly calls such constructions “Referenz-Aussagesstrukturen” (Reference-Assertion Structures).
The officer-in-chief (OC) asks the patient for what has happened to him. Before the patient verbally refers to the lamp (line 006) he attached to the ceiling he points in its direction (line 007, see Fig. 8). In the same way, he points twice downwards to the staircase where he lies (lines 010–011) before adding that he fell down (line 014).

Pointing gestures never occur after verbal reference. Although they often start in pre-beginning position, it could be seen in ex.3 that this is not obligatory. This is because it is their contingent projective relationship to the upcoming verbal reference (by deictics, definite NPs, etc.) which matters, not their relationship to the turn-beginning. There seems to be a temporal metrics, according to which pointing gestures are produced in the transition space before the TCU which contains the verbal reference terms or before the referring term within the same TCU (cf. already Kendon, 1972). Referential pointing thus is performed in line with a more general logic of the multimodal production of spatial prerequisites of turn-construction.

5.4.2. Project upcoming action

Speakers may choose to turn to a new topic or action stepwise within the turn (Schegloff, 1996). But particularly in institutional interactions, institutional agents initiate new sequences or topics (which can be expected from them on behalf of role-related tasks and entitlements) directly (i.e., without prior understanding displays, discourse markers, etc.) with the beginning of a new turn. Examples include questions (like the elicitation of a narrative in line 003 of ex.14) and requests (lines 013 and 027).

Ex.14 Emergency drill IDS-FOLK-Rettung3, 00:01–00:25

001 OC: <leans on his left hand next to PA, orients to PA
002 <f> hallo guten TAG;
   hello good afternoon
003 OC: (-) **was is passIERT?**
   what happened
(...)
012 Al: [<<positioned left to PA, fixes PA's head >
013 halten sie den kopf mal ruhig [BITte;>]
   keep your head PRT steady please
(...)
027 Al: [<<addresses A2, all, p> gib mir nen druck;>]
   give me a pressure
028 A2: [[(looks up to Al))]
029 A2: <<looks to Al> brauchst DRUCK? (---)
   need pressure?

Like in many other languages, the linear structure of German syntax distinguishes several sentence-types, which are routine formats for corresponding actions (verb-first (polar) questions (line 029), W-questions (line 003), verb-first imperatives (lines 013, 027)). In these cases, sentence-type and corresponding action can be projected already at the beginning of the turn. In contrast, the sentential structure of declarative clauses and the type of action performed by them usually are not already projected with the first item, but become only increasingly clear during the turn-in-progress (Auer, 2005; for Danish see Steensig, 2001). But also gestures can project actions to come: Streeck and Hartge (1992) show how palm-up open hand gestures are used to project storytelling in Ilokano; Heath and Luff (this volume) analyze in detail how the auctioneer's raising of the hammer and gesturing with it foreshadows contracting by the imminent fall of the hammer (and thus occasions further (escalating) bids).

Of course, retrospective devices which create a relationship between the upcoming turn and prior talk (cf. section 5.2) also project pragmatic and/or topical properties of the upcoming turn, just as many discourse particles and formulae specialized in particular types of actions (like request formulae 'c'mon, please) do.

Nonverbal stance-markers of uncertainty like shrugs (Streeck, 2009) and facial displays of skeptical uptake (frowning) exhibit also this dual temporal indexicality, taking both a responsive stance towards the prior turn and projecting the kind of action to come (e.g., non-committing response, objection or repair-initiation). In emergency interactions, pointing is not only used to refer to objects (as in ex.3, see above section 5.4.a), but often projects addressees of requests and objects to be used for requested actions. We can see this in ex.15: The officer-in-charge uses pointing and gaze to pick out both the addressee of the request and the relevant objects.

---

Declarative syntax can be used for a much wider range of actions, though, depending on content and prosody. Elliptical turns are mostly used for fulfilling projections in a type-conforming way.
Ex. 15 Emergency drill IDS-FOLK-Rettung2, 00:29–00:37

001 PA: auf EINmal is des EINfach hat er da reinge%stochen.#
   All of a sudden he just stabbed me there
OC: >gazes at PA-----------------------------------------#
   %fig.9

Fig. 9. Officer-in-chief looks at patient (line 001).

002 (0.55)
   OC: shifts gaze from PA to A1
003 OC: %*du machs viTAL*par#amefeter; (.)
   You check the vital parameters
   OC: *points at A1--*
   #points at measurement device#
   *gazes at A1-----------------------------------------#
   %fig.10

Fig. 10. Officer-in-chief points and gazes at A1 (line 003).

004 OC: %*von dir will ich_n öh öh STÜTZverband# (.) mit öh öhm
   I want a support bandage from you
   with erm erm
   *points at first aid box standing before A2---------->
   *gazes at first aid box-------------------#gazes at PA's belly
   %fig.11
Fig. 11. Officer-in-chief points and gazes at A2’s first aid box (line 004).

Immediately before OCH addresses her assistant A1 in line 003, she turns her gaze to her and points to her (see Fig. 10). As the turn proceeds, she points to the device to be used for measuring the vital parameters she requests from A1. Before her next turn, which is addressed to A2 (line 004), she turns to gaze at the first aid box standing before A2 while simultaneously pointing at it (see Fig. 11). The box contains the support bandage OCH requests from A2. Gazing and pointing are closely synchronized with verbal turns so as to project who will be the addressee of the next turn and which object will matter for the upcoming request.

It is a matter of conceptual debate whether dispreferred actions are to be defined by structural or by normative properties (see Bilmes, 1988). Nevertheless, it is a robust finding that prototypical cases are (often abundantly) marked by devices projecting an action which runs counter to what is likely to be expected (see also ex. 11). Here is an example from a mediation session: The mediator (MED) asks the counter claim plaintiff (CCP) to make a statement about offences he is accused of; CCP responds by referring to (probably exonerating) documents he has brought along.

Ex. 16: Mediation session IDS-Schlichtung 3001.01

01 MED: sie ham sisch bisher noch nischt EINgelassen
   you have not yet made a statement
02 un deswegn sind sie also hIER. (.)
   and this is why you are here
04 dass wir [uns] mal darüber unterHALten.
   that we PRT talk about it
05 CCP: [ja,]
   yes
06 (1.4)
07 CCP: ja. (.) ä:h-
   yes uhm
08 (1.4)
09 CCP: um des:-
   in order to
10 (1.3)
11 CCP: um der GANzen- (0.3)
   in order to the whole
12 äh. (0.2) SÄche hier; (.)
   uhm thing here
13 Etwas-
   a bit
14 (1.7)
15 CCP: EInsichtischer zu werden,
   more transparent to become
16 für SIE vielleischt- (0.5)
   for you perhaps
17 isch hab HIER mal schon, (0.5)
   I have here PRT PRT
18 mein WENN sie_s interessiert. (0.6)
   I mean if you are interested
19 ein ä:h-
   an
20 MED: NET [net zu_n AKten] (. ) zu_n akten gehen.
   do not do not go to the documents
21 CCP: [ausZÜge-
   excerpts
Before he finally starts presenting the documents (in line 17), CCP's turn is replete with indices of a dispreferred action from its beginning: pauses (lines 06, 08, 10, 14), cut-off (09), self-repair (11), hesitation markers (07, 12), hedging ("vielleicht", perhaps, 15), anticipatory justification (09-16) and parenthetical offers to negotiate the relevance of the dispreferred action with the mediator (16, 18). All these phenomena foreshadow an action which is not in line with the expectation created by the request, i.e., a straightforward statement about what happened.

Like projections regarding next turns, it needs to be underscored that turn-initial projection of action can be more or less distinct in terms of defining the precise kind of action to come. E.g., turn-initial devices foreshadowing a dispreferred action often do not bespeak whether the turn will finally become, e.g., a refusal, a move to renegotiate the first pair part or whether it will initiate an insertion.

5.4.3. Frame the turn

Evaluative, emotional and epistemic stances are likely to be displayed at the turn-beginning. Displaying stance pre-orient the recipient's understanding in terms of assessment and type of action, may foreshadow which kind of affiliative action is expected from the recipient (e.g., laughter, co-indignation) and can serve to pre-empt certain reactions like criticism, repair-initiation, etc.

Ex. 3 Emergency drill IDS-FOLK-Rettung3, 00:09-00:16

003 OC: (-) was pasSIERT?= what happened
004 Al: =Oh! oh!
005 PA: <<all,pp> ich wei-> (-)
          I kno
006 <<all> ich WEIß net mehr ich bin->
          I don't know anymore I am
008 ich hab die LAMP gemacht ghabt;=
          I attached the lamp

In ex.3, the teller prefaces his story by an epistemic disclaimer, which can be seen to pre-empt criticism of inaccuracy and incompleteness, displaying that the speaker adopts only reduced accountability for his narrative compared to what could be expected under normal conditions.

Frames may also concern evaluative stance and keying. Projector constructions (Gü nthner, 2011a,b; Gü nthner and Hopper, 2010; Pekarek-Doehler, 2011; Auer, 2009b) like wh-pseudo-clefts (like "what happened was", "what really annoys me"), /¿-extraposition and 'NP is that' - sentences ("the thing is", "the problem is") are grammatical devices to project an upcoming action which will (mostly) require more than just one TCU. The specific lexical content of the projector clause expresses the stance attached to the things to be talked about in the projected turn components.

Vocal cues like laugh particles (cf. Jefferson, 1985) are further powerful means to embody stance towards matters to become topical in the turn-in-progress as well as devices to orient recipients to the expected kind of uptake. In ex.17, Denis comments on the fact that his mates glued a toy cow on the window pane of a caravan they were given by the municipality for their leisure time.

Ex. 17: Juk 13-1 "Kuh und Bürgermeister"

01 Mark: ihr [seid voll die A:schlöscher.] you [are real Assholes.]
02 Denis: [haha des is ja sEhr geil, ] .hh
          [huhu that is rEally hot, ] .hh
03 Denis: <<laughing> de BÜRGERmeister kommt mOrgen,
          <<laughing> the MAYOR will come tomORrow,
04 Otto: [sieht die kuh hier, ]
          [sees the cow here, ]
05 Denis: [,hh oder übermorgen, ] (.)
          [,hhh or the day After, ] (.)
06 Denis: [und ihr BABBT, (.) HA ha-]
          [and you STICK, (.){(laughs)}]
07 XM: [((chuckle))]
08 Denis: ihr BABBT, (.) mit sekundenkleber, (.)
          you STICK, (.) with crazy glue, (.)
09 was ja gar net Assig is,
          which is absolutely not trAshy,
10 auf des holzteil die kuh fest.> (.)
          the cow onto that wooden part.> (.)
11 XM: {(laughter)}
In line 02, Denis first produces laughter particles to preface a positive assessment used as a projector construction. Both serve as an evaluative and interpretive frame for a description of a scenario, which, in absence of these stance markers, could be (mis)understood as amounting to a reproach (which, in fact, Mark had made just before Denis’ turn in line 01). The recipients align with Denis’ projection of a positive evaluative stance and humorous key by responding with chuckles.

6. Discussion

6.1. Contingencies of tasks of turn-construction at turn-beginnings and their relationship

The basic claim of this paper is threefold:

- There are four tasks which are to be dealt with when designing a turn,
- these four tasks are intrinsically ordered according to a temporal and pragmatic logic,
- these four tasks explain how turn-beginnings are designed.

Still, this claim has to be qualified. It hinges on several conditions which vary according to when what kind of turn is to be produced. In addition, the basic consecutive logic of the tasks is not necessarily reflected by a series of activities specialized to perform them one after the other. So, the four tasks of turn-construction do not apply for every turn to be produced and they do not apply in the same way, independently of what the turn is to be doing and under what conditions it is produced.

First, not all tasks may be relevant. Displaying understanding and dealing with projections do not matter if there is no prior contribution to respond to (see ex.1). Many turns are restricted to responding: They display recipiency or understanding, but do not add anything more to the progression of the talk than a continuer, nor do they satisfy any projection. Turns can be so short that they do not have an internal dynamics characterized by turn-internal projection.

Second, not all tasks need to be dealt with in an obvious fashion by specialized means of talk and bodily action. Tasks can be dealt with implicatively, as in the case of displaying understanding by performing a projected action. Turns can also build on solutions to the task which have been achieved earlier in the interaction and are still in place (e.g., bodily arrangements) or can be presupposed as still latently available (this is mainly the case for displayed understandings or ellipsis). This is one facet of the local contingency of interaction, which inevitably adapts to the local conditions of the very moment of its production (cf. Mondada, this volume).

Third, tasks can be dealt with simultaneously. The multimodality of embodied action allows dealing with tasks simultaneously by using various modal means. E.g., gaze can be coordinated while producing a response or while projecting properties of the upcoming turn, a speaker can look at an object which is to become topical in a next TCU while still talking about something else, s/he may turn his/her head to a next addressee while completing a TCU which is still responsive to another participant (Schmitt and Deppermann, 2010).

6.2. The grammar of turn-beginnings

This paper is a sketchy attempt at outlining determinants and repertoires of “positionally sensitive grammars” (Schegloff, 1996) of turn-beginnings in German. I have tried to show how the choice of linguistic forms at turn-beginnings is occasioned a) by general tasks of turn-construction, which arise from the interactional, temporal and bodily nature of turn-construction and thus from language use in interaction and b) by spatial arrangements, preceding actions and their linguistic properties, which project certain next actions and which require or license certain linguistic structures to be produced, notably, offering the opportunity to build emerging talk on structures produced in the interactional past which are still available (cf. Goodwin this volume). Turn-beginnings, therefore, in this respect, are not one position. Rather, which grammatical choices are available at turn-beginnings depends on various orders of local contingencies. Still, the general point about these positionally sensitive grammars is that they are characterized by their orientation to the four tasks outlined in this paper, leading to different grammars by their local sensitivity and by the choices the speaker makes in terms of what the incipient turn is going to be. The finding that sentential structures play only a minor role in the construction of turn-beginnings derives from the fact that they are almost exclusively used in order to produce non-projected progression. This concern, however, is mostly tackled only later in the turn, after the four tasks outlined here have been dealt with (see Deppermann, 2012).

From a linguistic point of view of turn-beginnings, two kinds of functionally motivated grammatical organization stand out: a grammar of retrospection and a grammar of framing and projecting. The grammar of retrospection is specialized in displaying understanding and relating the upcoming turn to prior discourse. For the first task, displaying understanding,
response particles are the first option to do the job, i.e., structures which are not even phrasal. The second task, relating back, often combines with the third task, framing and projecting what is to come. Consequently, grammatical constructions from the same classes – discourse markers, phrasal formulae and connectives – are preferred turn-initial devices to establish coherence and to frame the incipient turn (only the latter opening up the possibility to begin the turn with a sentence). In addition, there is a grammar which makes use of structural latency (see Auer, 2006 for this term), which is not so much functionally motivated (except for reasons of economy), but rather structurally occasioned. These are structures of online-syntax responsive to the precise moment in interaction, i.e., the preconditions of interactional structure established so far which determine options and restrictions for verbal production. Ellipsis, completion, expansion, and substitution are linguistic phenomena which rest on the availability of the interactionally produced structure so far for a next TCU. Of course, adjacency plays a major role in this, but we still know very little about the precise determinants of which structures are available for how long, for which next action, depending on which kinds of competing and interpolated actions and we do barely know which restrictions hold for the linguistic design of these actions.

7. Conclusions

While prior research has mostly focused on turn-endings (projection of TRPs, continuation, incrementing, etc.), this paper contributes to an interactional linguistics of turn-beginnings. I tried to show that there are four general tasks of turn-construction, which account for the design of turn-beginnings:

1. Achieve joint orientation to the upcoming turn
2. Display uptake
3. Deal with projections from prior talk
4. Project properties of upcoming turn

The ubiquitous nature of the tasks and solutions relevant to them have sedimented in routine multimodal practices both in terms of visual conduct (coordination of body and gaze, etc.) and grammatical constructions, such as various kinds of response and discourse particles, connectives and adjacent ellipsis. The four tasks thus provide a functional explanation for the design of turn-beginnings. Tasks 1–3 explain why turns mostly start with linguistic structures which exhibit the sensitivity of turn-construction to the precise interactional moment by being responsive and/or depending in their situated design, function and meaning on prior structures. While the relevance of tasks 2–4 hinges on various interactional and linguistic factors, task 1 is inevitably relevant for every turn to be produced in face-to-face interaction; it explains why turns often cannot start right away when the floor is available, but need various kinds of multimodal preparation. The fulfillment of individual tasks often must build on the fulfillment of other tasks. In particular, establishing joint attention and displaying uptake structurally underlie the possibility of interactional progression, which needs to be rooted in retrospective understanding in most cases as well. Such relationships of precondition and enablement also account for typical patterns of the ordering of devices at turn-beginnings, like ‘response particle + projector construction’, ‘attention-getter + referential term’, etc. In sum, the four tasks seem to be both temporally and functionally ordered, the accomplishment of an earlier task basically being the prerequisite for accomplishing a later task. Deviations may either result in problems of interactional alignment or they may be used reflexively, giving rise to inferences based on the omission of dealing with one or more of the tasks posited above. Further research is needed in order to explore whether the hypothesis about the consecutive nature of the four tasks holds and how deviations may be exploited for pragmatic concerns.

Turn-construction is not only a verbal activity. It requires the coordination of various kinesic modalities which both enable and support, and at times also replace, verbal turns. There are at least four respects in which multimodality inevitably matters for turn-construction:

- **Embodied situated practice:** Turns are produced with respect to the local contingencies of spatial arrangements and opportunities, bodily action and participation frameworks. They require spatial and bodily preconditions for their production and reflexively play their role in organizing bodily action and object handling.
- **Coordination:** Multimodal resources have to be coordinated sequentially and simultaneously, both intrapersonally and interpersonally (Deppermann and Schmitt, 2007), in order to contribute to interaction.
- **Simultaneity** provides for several layers of action (see Goodwin, this volume) and even the possibility of performing several actions at once (see Heath and Luff, this volume). This makes it possible that visual action can complement vocal production, allowing to deal with several tasks of turn-construction simultaneously. Different actions can be performed for different recipients; peripheral perception can be mobilized to anticipate future courses of action while other events are still going on.
- **Functional equivalence:** Often, several modalities can be used interchangeably for dealing with one task (e.g., getting the recipient’s attention, referring, assessing). Therefore, it is possible to supplant devices from one modality with those
from another, depending on situational opportunities, anticipated rhetorical effects or efficiency for securing understanding.17

Viewing turn-construction with respect to interactional tasks necessarily leads to a “holistic” view of turn-construction within the framework of multimodal interaction. This paper has tried to show that turn-construction is informed by requirements of multimodal, spatially and temporally situated interaction from its very beginnings and that the various practices to deal with it precisely use and address multimodality both as task and as resource.

References


17 Future research should deal in detail with the precise (interactional) conditions and consequences of substituting resources from one modality with another for the same interactional task.