

## Research Article

Oliver Ehmer\* and Daniel Mandel

# Projecting action spaces. On the interactional relevance of cesural areas in co-enactments

<https://doi.org/10.1515/opli-2020-0172>

received December 23, 2020; accepted July 24, 2021

**Abstract:** This article investigates the interactional relevance of weak cesuras in multimodal transitions in enactments. Previous research has pointed out that enactments are multimodally accomplished phenomena in that they do not only consist of a quotation but usually involve changes in prosody and bodily conduct, too. Furthermore, it has been noted that an upcoming quotation may be projected in the preceding talk by phonetic cues. There is, however, little research on the precise multimodal realization of such transitions and their possible interactional relevance. Taking this as a starting point, we analyze a collection of co-enactments. Firstly, we show that quotations are projected not only by phonetic but also bodily cues, which often build up gradually in the preceding talk. These smooth transitions into enactment are analyzed as “cesural areas.” Secondly, we argue that such cesural areas and the cumulation of multimodal projections open up an opportunity space in the sense of Lerner (1991), whereby a joint enactment involving co-participants, i.e., a co-enactment, is possible. Thirdly, we show that participants jointly develop the meaning of the enactment in this space, mutually taking up and elaborating on their prior contributions. The data is taken from a corpus of collaborative storytellings in German.

**Keywords:** cesuras, segmentation of speech, talk-in-interaction, reported speech, projection, multimodality, collaborative storytelling, action space

## 1 Introduction

The notion of a “cesural area” (Barth-Weingarten and Ogden this issue, Barth-Weingarten 2013, 2016) captures the assumption that the boundaries between interactional phenomena are not necessarily created at a single point but sometimes extend across a stretch of talk involving multiple modalities. The aim of this article is to show that “cesural areas,” which occur at the beginning of an enactment, open up an action space that offers co-participants an opportunity to co-produce the enactment. By enactment, we mean that a person “pretends to inhabit another body” (Streeck 2002, 591) and animates this body by using different resources; such resources include the use of reported speech, the voice, and/or bodily behavior, such as gesture, gaze, facial expressions, posture shifts, and proxemics (Good 2015, Soulaïmani 2018). In our

\* **Corresponding author: Oliver Ehmer**, Albert-Ludwigs-University Freiburg, Romanisches Seminar, Platz der Universität 3, D-79085, Freiburg im Breisgau, Germany, e-mail: [oliver.ehmer@romanistik.uni-freiburg.de](mailto:oliver.ehmer@romanistik.uni-freiburg.de)

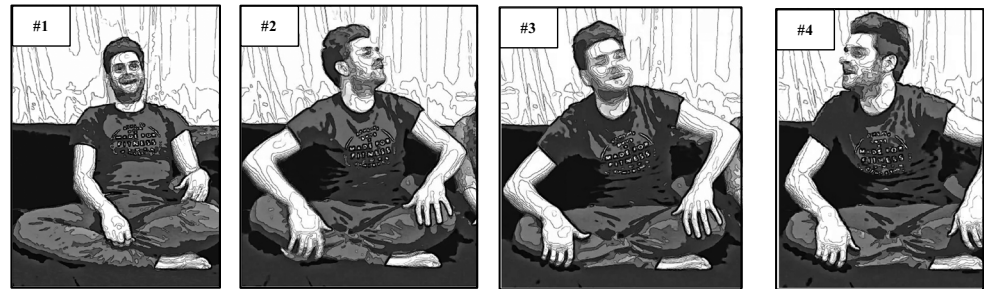
**Daniel Mandel:** Albert-Ludwigs-University Freiburg, Romanisches Seminar, Platz der Universität 3, D-79085, Freiburg im Breisgau, Germany, e-mail: [daniel.mandel@romanistik.uni-freiburg.de](mailto:daniel.mandel@romanistik.uni-freiburg.de)

approach, enactments are performances in the sense that they are “depictions” or “demonstrations” of an event rather than “tellings” or “descriptions” (cf. Clark and Gerrig 1990, 764, Sidnell 2006, 377, Wilkinson et al. 2010, 58). While enactments may be realized in a purely non-verbal form, they often contain “reported speech” or a “quotation.”

In this chapter, we present an analysis of co-enactments within the frameworks of Conversation Analysis (henceforth CA, Clift 2016, Mondada 2019, Sidnell and Stivers 2012) and Interactional Linguistics (henceforth IL, Barth-Weingarten 2008, Couper-Kuhlen and Selting 2018, Selting and Couper-Kuhlen 2001). Previous research from these domains has pointed out that the boundaries between an enactment and the surrounding talk may be fuzzy. With respect to the verbal dimension, for instance, it has been shown that voice quality may change before the quotation is even uttered, a phenomenon that has been called “foreshadowing” (Klewitz and Couper-Kuhlen 1999). Likewise, it has been shown that the enacted figure is sometimes already embodied on the bodily dimension before the quotation has begun (Ehmer 2011, 316–321). Example 1 presents such a case, which is transcribed according to GAT conventions (Selting et al. 2009). Lucas (on the left) and Renate (on the right) are talking about a beach vacation they went on together with another friend (Alex). Lucas is the main teller of the story; it was his first time surfing with friends. The punch line of the story is that Lucas always had problems getting onto the waves, while his experienced friend Alex was already a confident surfer. The following excerpt begins with Lucas quoting Alex’s request for Lucas to come closer to him as if this should be no problem at all.



Figure 1: Setting of Example 1.

**Example 1.** *come* (ger\_sft05, 7:33-7:37)

- 01 L: #1 °h und er sitzt da so #2 RUM==  
and he sits there like (that)
- 02 =und #3 guckt dich so AN <<all, dark voice> un(d) #4 mein(t)->  
and he looks at you like (that) and says



- 03 °h <<f> #5 ↑ja was #6 MACHSCH?>=  
so what are you doing
- 04 °h #7 <<f> !KOMM! ,>  
come
- 05 [<<pressed> h°> ((laughs))]
- 06 R: [<<pressed> h°> ((laughs))]

The quotation in l. 03–04 consists of the challenging question, °h <<f> ↑ja was MACHSCH? >= “so what are you doing” and the directive °h <<f> !KOMM! ,> “come.” In the preceding l. 01–02, the narrator describes the bodily behavior of his friend, who was sitting on his surfboard and looking around. At the end of l. 02, we can identify a quotative containing the verbum dicendi *meinen* “to say.” While on the verbal dimension, the transition from the quotative (l. 02) to the quotation (l. 03–04) is a rather clear one; here, we are interested in the phonetic and bodily cues that occur in the transitional area before the start of the quotation. Regarding the phonetic dimension, we can note that Lucas’s voice has already changed to a dark voice quality at the end of l. 02 and he speaks with relaxed articulation. In addition, Lucas activates several resources on the bodily dimension, i.e., leaning forward, shifting the body to the right, and gazing into the room (see stills #1 and #2). While this bodily behavior in some way “illustrates” the actions of the story

character as they are described verbally, it is also clear that Lucas has already started to enact that character at this point. This example thus illustrates the frequent phenomenon that, in an enactment, phonetic and bodily means are already in play before the start of the actual quotation. There is, however, little research on the precise multimodal realization of such transitions and their potential interactional relevance.

This is the point of departure of the present paper. We focus on the gradual emergence of enactments in talk-in-interaction by adopting a multimodal perspective that includes the verbal, phonetic, and bodily dimensions. The aim of this article is threefold: first, we analyze in detail how different modalities are finely coordinated in enactments. We show that enactments emerge gradually in talk-in-interaction and that bodily cues typically precede phonetic and verbal cues. This leads to a smooth transition into the enactment, which we analyze as a “cesural area” between the enactment and the preceding talk. Second, we argue that the gradual building up of projections in the “cesural area” is relevant for organizing the interaction: These projections open up a joint action space (Lerner 1991, 453) in which co-participants are given the opportunity to co-produce the enactment and hence take part in the local action. Third, we show that such co-enactments can take different forms, ranging from co-enactments with quotations produced by participants at the same time to quotations that are produced successively. We show, however, that regardless of this variation in form, co-enactments are typically characterized by a joint elaboration of the contents of the enactment by the participants. This co-production is facilitated by the preceding cesural area.

The remainder of this article is structured as follows. After presenting the theoretical background (Section 2), we present the data used in the study and the methodology (Section 3). The analysis can be found in Section 4, followed by a summary of the results in Section 5. This article concludes with a discussion of the results (Section 6).

## 2 Theoretical background

To analyze the cesural areas in enactments and their possible interactional relevance, we refer to three domains of research in CA and IL: *quotations and enactments*, *projection* as a basic principle of talk-in-interaction, and the *cesura approach*.

### 2.1 Quotations and enactments

Stemming from Volosinov’s research, studies in CA point out how “reported speech is speech within speech” (Volosinov 1929/1973, 149) and how speakers may use direct reported speech to quote what someone has said in the past (Holt 2000). It has been shown that speakers report not only events and lend their voice to the original speaker in the narration when quoting but may also involve themselves in stance-taking activities (Günthner 2002, Niemelä 2011). In addition to verbal means, phonetics has been proven to be relevant, for instance, with respect to stance-taking on a reported action (Günthner 1999, 2002). While many studies typically employ the terms (*direct*) *reported speech* and *quotation* to focus on the verbal and phonetic dimensions, there are a growing number of studies that also integrate the body, e.g., gaze, gesture, and posture shifts, into the analysis. Such studies often refer to the concept of *enactment* (Sidnell 2006, Stukenbrock 2012). For instance, Streeck defines enactment as follows:

During an enactment, the speaker pretends to inhabit another body – a human one or that of an alien, perhaps even a machine, or her own body in a different situation – and animates it with her own body, including the voice. (Streeck 2002, 591)

This definition highlights the fact that enactments involve a form of pretense, play, or change in footing (Holt 2007, Keevallik 2010). Moreover, the animated entity need not be a person (including the speaker him/herself in some other situation) but may even be an animal or an inanimate object. While some authors use the term re-enactment (Sidnell 2006, Thompson and Suzuki 2014), others emphasize the fact that

enactment can stage not only a past event but any kind of scene or situation, be it future-based, conditional, fictional, or generic (Cantarutti 2020, Ehmer 2011). Enactments are (selective) *demonstrations* of some aspects of a scene or event (Clark 2016, Clark and Gerrig 1990, Wilkinson et al. 2010).

With regard to quotations and enactments, a central task in interaction is to make them recognizable. Firstly, the enactment or quotation must be contextualized as a possible transition from talking “about” a scene to enacting it. Secondly, interactants must contextualize whose, or which, body is being animated, particularly when there are several possible referents. For this task, speakers may employ lexico-syntactic devices that contain a verbum dicendi or quotative particles, like German *so* in the phrase *und ich so* (English ‘and I’m like’, Golato 2000), for instance. Such “quotatives,” however, may not be present as introductory components (Holt 2007). Apart from lexico-syntactic means, speakers may make use of phonetics to make a quotation recognizable via changes in tempo, volume, pitch, and voice quality to set the quote off from the surrounding talk (Couper-Kuhlen 1999). Furthermore, the body may be utilized to provide further details about the enacted situation, such as the manner of action (Park 2009). The bodily dimension has been argued to be helpful in establishing intersubjectivity (Wilkinson et al. 2010) and to stimulate co-participation in enactments (Good 2015, Soulaïmani 2018).

Regarding the timing of the cues across different dimensions, Klewitz and Couper-Kuhlen (1999) have shown that changes in the phonetics may already occur in the talk that precedes a quotation, thereby “foreshadowing” and projecting the upcoming quotation. Similarly, phonetic features may extend beyond the quotation into the following stretch of talk, which is referred to by Bolden (2004) as “fading out.” With respect to the body, it has been noted that nonverbal resources can be employed before the quotation has even started (Ehmer 2011, 316–321). Thus far, however, we are not aware of any studies that focus specifically on the timing of bodily resources in enactments.

Before turning to projections more generally, we first clarify our understanding of some central terms. By *enactment*, we mean that a participant in an interaction “pretends to inhabit another body” and animates it by using verbal, phonetic, and/or bodily semiotic resources. *Co-enactments* are thus enactments of the same “other body” by several interactants. The term *quotation* is used to refer to the verbal part of the enactment, i.e., a verbally animated utterance by the “inhabited body.” By *quotative*, we mean all linguistic devices that frame a quotation prospectively and/or retrospectively.

## 2.2 Projection, action space, and multimodal compaction zone

To make an enactment recognizable, the concept of projection, a basic principle of talk-in-interaction, plays an important role. Following Auer (2005, 1), projection can be defined as “the fact that an individual action or part of it foreshadows another.” Projection depends on some observable phenomena that make the projecting element recognizable as a projection of a certain continuation, or closure of a temporally emerging gestalt. Projections play a central role in the organization of interaction since they provide other participants with a “certain premonition as to what this actor might be up to next” (Streeck 1995: 87). Such a “premonition” or “anticipation” of what is about to come is a central prerequisite for many different kinds of interpersonal coordination, such as the organization of relevant next actions (Dausendschön-Gay and Krafft 2009, Streeck and Hartge 1992), collaborative productions (Mori and Hayashi 2006), and, in particular, turn-sharing practices (Hayashi 2012, Lerner 2002, Pfänder and Couper-Kuhlen 2019). With respect to the interactional relevance of projections, Lerner (1991) coined the term “action space,” which he defines as follows:

Any action that foreshows [i.e. projects, OE/DM] a recognizable completion furnishes an action space. This space extends from the emerging utterance-so-far to the projectable completion of the unit. Units that foreshow a segmented action space are thereby completable. Completable units regularly have a projectable preliminary completion place from which the anticipatory completion can begin. (Lerner 1991, 453)

Lerner thus emphasizes the interactional property of projections: They can be completed by co-participants, who are therefore able to join the projected action. By also referring to these spaces as “opportunity spaces” (Lerner 1991, 450), Lerner highlights the fact that they provide an opportunity for joint action without necessarily demanding it or actually obliging co-participants to actively participate.

Projections may be established on different linguistic dimensions, including lexico-syntax (Auer 2005, 2015, Günthner 2011), phonetics (Couper-Kuhlen 2004, Wells and Macfarlane 1998, Walker 2012), as well as the body (Streeck 1995, 99, Streeck and Hartge 1992, Streeck and Jordan 2009, Kaukomaa et al. 2013, 2014, Schegloff 1984, 291). With respect to the relation between these dimensions, Schegloff (1984, 291) and Streeck (1995, 102) note that bodily behavior often precedes the verbal part of an action and “prepares the scene.” In a general sense, projections from different modal dimensions may build up successively, jointly preparing the ground for the same projected action during the emergence of a multimodal gestalt in interaction (Mondada 2014). Similarly, optimal completion points of emerging gestalts are reached “at a point when all syntactic, prosodic, and semanto-pragmatic projections have been processed” (Auer 2010, 12; our translation). To capture these moments of gestalt closure, Stukenbrock (2018, see also 2015) introduces the notion of a “multimodal compaction zone:”

“Multimodal compaction” means that projections cumulatively and in a temporally emergent way converge in putting those components on the spotlight that bear the highest informational value. (Stukenbrock 2018, 40)

We take a multimodal compaction zone to mean an area of talk-in-interaction where multiple projections—which have been opened up beforehand on possibly several semiotic dimensions—converge and accumulate in such a way that they culminate in the (possible) closure of the currently emerging gestalt. Crucially, Stukenbrock conceives of these moments of convergence not as “points” but rather as “zones” within talk-in-interaction.

While the notion of “multimodal compaction” refers to the zone where multiple projections *converge* and an emergent gestalt is (possibly) in the process of being closed, our analysis of enactments also focuses on how multiple projections across different dimensions are gradually *built up* before the action or action component, for which they prepare the scene. In this regard, we find the notion of the “cesura approach” extremely useful.

## 2.3 The cesura approach and the notion of “cesural area”

The cesura approach was developed by Barth-Weingarten (2013, 2016, also Barth-Weingarten and Ogden this SI) to address segmenting speech and interactional phenomena in a novel way. In CA and IL, the issue of segmenting speech and interaction has been continually addressed (Auer 2010, Deppermann and Proske 2015, the articles in this special issue). So far, different taxonomic criteria have been identified on which such segmentation could be based, among them syntax, prosody, and pragmatics. Auer (2010, 2015) argues that it is the task of the participants to recognize and process emerging gestalts at all times throughout the course of a conversation. Moreover, interlocutors may even negotiate completion points or gestalt closures in the ongoing interaction. Auer, therefore, takes the position that the segmentation of speech should not be “based on units but on points” (2010, 12) when emerging gestalts are (possibly) complete.

Taking a similar point of departure as Auer, Barth-Weingarten (2013, 2016) employs the notion of “cesura” to focus on what separates chunks of talk from one another. She emphasizes, however, further aspects: There may be several relevant parameters within one dimension, e.g., relevant parameters for the phonetic dimension include pitch, loudness, and voice quality. These parameters are thought of as being gradual, and different degrees of parameter changes create cesuras of variable strength: “the more parameter changes occur at one point in talk, the stronger the perception of the cesura will be” (Barth-Weingarten and Ogden 2021, 537, this issue). In addition, temporal interplay between parameters plays a



crucial role. Cesuras are often not created at a single point in time but across a stretch of speech. Empirically speaking, it is at times impossible to determine the exact location of a “boundary” and cesuras may actually emerge across a whole area of speech (Barth-Weingarten 2013, 100; 2016: 6, 75–76); the notion of “cesural area” captures this fact. The more changes in different parameters do not coincide at a specific point but diverge from one another over a stretch of talk, the more extended and possibly “weaker” the cesural area becomes.

While the cesura approach was originally developed with a focus on the verbal and the phonetic dimension, we aim to integrate the body into our analysis of co-enactments. We use the term “cesural area” as an analytical notion to capture the transitional area in interaction between one segment and the next, which participants orient to as being relevant. A cesural area is created by the interplay of parameter changes on different multimodal dimensions, including lexico-syntax, phonetics, and the body. The parameters themselves are gradual and may have indexical properties.

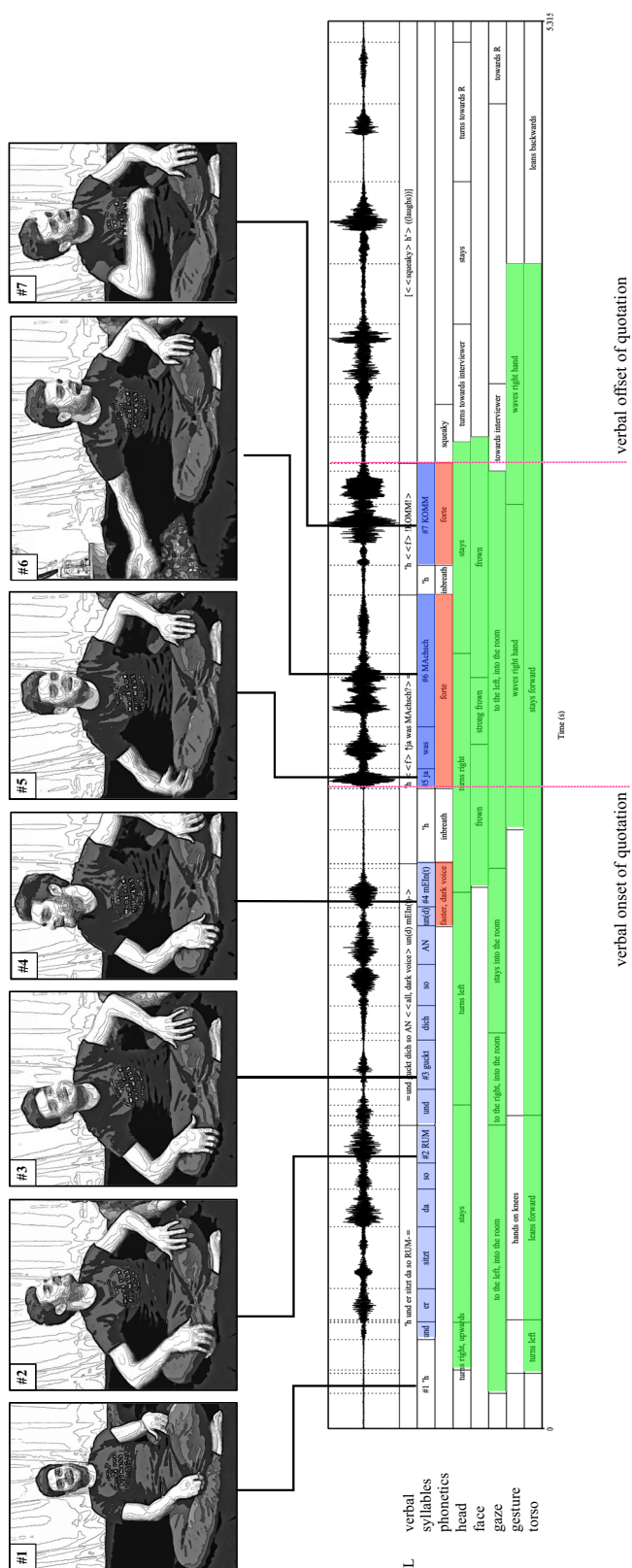
### 3 Data and methods

Our data stems from the Sofa-Talks Corpus (Dressel and Satti 2021, Pfänder et al., in prep.). The corpus comprises more than 200 video recordings of couples with a close personal relationship, who were asked to talk about episodes in their lives as friends, couples, and families. In addition to the couple, a third interlocutor is present, who acts as the interviewer. In most cases, (s)he is known to the couple but does not take part in the events narrated. While the two people on the sofa are sitting next to one another, the third participant is sitting in front of them, either behind or next to the camera. Both the interviewer and camera are addressed as recipients of the stories.

For the purposes of our study, we collected co-enactments in which a couple enacts the same story character with temporal overlap and at least one of the tellers produces one or more quotations. While we found a large number of enactments in our corpus that exhibit a low degree of co-participation, we focus on 15 instances in which there are high levels of co-participation. In the subsequent analysis, we present German examples wherein both participants verbally enact and start producing the first quotation more or less simultaneously.

To analyze the co-productions, we proceeded systematically by annotating the relevant resources. First, we transcribed the audio-material using *Praat* (Boersma and Weenink 2020) according to the GAT 2 conventions (Couper-Kuhlen and Barth-Weingarten 2011, Selting et al. 2009). Second, we annotated the use of different bodily resources in ELAN ([Computer software] (2020); Kendon 1980, 212, Bressemer and Ladewig 2011). For the bodily dimension, we identified head movements, facial expressions, gaze, gestures, and movements of the torso, which are annotated on separate tiers. Each bodily movement is annotated from its onset until its retraction. As a third step, we prepared visualizations of the enactments by plotting the annotations in *Praat* and visually enriching them by means of the design tool Runway ([Computer software] (2020)). This procedure allowed us to respect the different temporalities of the multimodal resources (Deppermann and Günthner 2015, 9, Deppermann and Streeck 2018), e.g., their temporal extension and timing relative to one another. Figure 2 displays an example of the resulting visualization for Example 1.

On top of the figure, stills from the video recordings are displayed. Below them, a plot of the annotations created in *Praat* is shown. On the first line of the plot, there is the waveform, which represents the sound. The following lines show annotations of the various multimodal resources in a time-aligned fashion. The first line represents the verbal dimension as found in the transcript. In the second line, colored in blue, the words are segmented into syllables; the quotation is in a slightly darker blue. The stills on the top of the visualization are temporally aligned with the syllables by black lines. The phonetic dimension is



**Figure 2: Visualization Example 1.**



highlighted in red, while the different bodily resources are represented in the lines highlighted in green. The vertical red lines signal the onset and the offset of the quotation.

These visualizations are used to analyze, and depict, convergences and divergences across the verbal, phonetic, and bodily dimensions that are present in the data set and which can be modeled with the “cesura approach.”

## 4 Analysis

The central aim of our analysis is to show that cesural areas occurring at the beginning of enactments in joint storytelling may open an action space that offers other participants the opportunity to co-produce the enactment. In the following analysis, we present three instances of enactments in which participants co-produce a quotation. While the first example recruits both verbal and bodily cues to project the enactment, the second example illustrates that bodily projections alone can be sufficient to co-produce the quotation. Moreover, while in these two examples, the participants co-produce the quotations in temporal overlap, the third example presents a case in which participants produce the quotations successively. We show that, despite this difference in timing, all of the examples analyzed are characterized by a very similar process of mutual uptake and co-elaboration of the enactment.

### 4.1 Quotations starting simultaneously

The following example illustrates that the projection of an enactment may lead to a precisely timed onset of the quotation by both participants. We first provide an analysis of the sequential development of the interaction and then, in a second step, analyze the bodily cues. Two friends, Lena and Rosa, are sitting side by side on a sofa (Figure 3). They are talking about a shared experience they had while working at a dance theater. Part of the job was to play a triangle from time to time to signal to the audience when to applause; both of them were sometimes uncertain about the particular moment at which they should play the triangle.



Figure 3: Setting of Example 2.

**Example 2.** *playing the triangle* (ger\_sft07, 12:49-12:55)

- 01 L: und manchmal w:usst ich halt <<laughing> AUCH nicht so-> h° (0.5)  
and sometimes I also just didn't know like
- 02 [<<p> (was IS\_n)?> ]  
what is PRT
- 03 R: [<<stylized dialect> SOLL isch jetzt;>]  
should I now
- 04 L: SOLL i[ch ↑jEtzt? ]  
should I now
- 05 R: [<<h> bimMELN?>]  
play
- 06 [<<f> Ode:r->> ]  
or
- 07 L: [ist es zu FRÜH,]=Ode:r-  
is it too early, or
- 08 (0.2)
- 09 NEE?=Ode:r-  
no or
- 10 (0.7)

At the beginning of the sequence, Lena describes from her hic-et-nunc perspective that she was sometimes unsure: und manchmal w:usst ich halt <<laughing> AUCH nicht so->h° (0.5) (l. 01) “and sometimes I also just didn’t know like.” The syntax is projective since the direct object, i.e., “what” she was unsure about, is missing. The syntactic incompleteness in combination with the semantics of the verbum pensandi *know* projects the production of an upcoming inner thought. This semantic and syntactic projection is further emphasized by the quotative particle *so* “like” (Auer 2006). The laughter particles at the end of the utterance frame the upcoming quotation as amusing (Günthner 2002). The temporal adjective *sometimes* characterizes the upcoming enactment as a depiction of not a one-off but a repeated event. The sequence continues with a series of co-produced quotations in the form of questions that take up these framings.

Lena’s first quoted question, <<p> (was IS\_n)?> (l. 02) “what is PRT,” is formatted with interrogative syntax that consists of the question word *what* and the copula verb *to be* in the third person singular. The phonetically reduced particle *denn* as\_n in this position indexes some form of astonishment, which is emphasized by the rising intonation at the end of the turn construction unit (TCU; Couper-Kuhlen and Selting 2018, 31–111). Simultaneously with Lena’s question, her interlocutor Rosa co-produces a further question. The utterance [<<stylized dialect> SOLL isch jetzt;>] (l. 03) “should I now” is marked as a question by verb-first syntax and rising intonation, in addition to being phonetically stylized as dialectal. Since the verb *know* is in first person singular, Rosa adopts the perspective of the story character “Lena” and quotes her inner thoughts, which likely also depict her own thoughts since she was often in the same “habitual” situation, too. The third quotation is produced by Lena again: SOLL i[ch ↑jEtzt?] (l. 04) “should I now.” Lena repeats Rosa’s immediately preceding contribution literally but nondialectally. Lena thereby ratifies Rosa’s co-produced quotation as adequate. Rosa’s contribution overlaps with Lena’s next utterance [<<h> bimMELN?>] (l. 05) “play.” Rosa incrementally continues her prior quotation with the infinitive form of the verb, spelling out the verb complement that had been left unsaid. By adding the turn-final marker

[<<f> Ode:r->>] (l. 06) “or” to her quotation, she further indexes her uncertainty (Drake 2015). In overlap with this quotation, Lena produces another quotation that elaborates on this uncertainty by asking if it was “too early”: [ist es zu FRÜH,] = Ode:r- (l. 07) “is it too early or.” Interestingly, by doing so, Lena takes up Rosa’s prior use of turn-final *oder*. The co-production ends with Lena’s last quotation NEE?=Ode:r- (l. 08) “no or,” which can be seen as a negative response to the preceding question and is once again marked as uncertain via the tag “or.”

With respect to the production of the verbal part of the enactment, both participants can be observed to co-produce quotations. On the syntactic dimension, the participants not only incrementally extend their own contributions but also re-use structures introduced by their co-participant. In terms of the semantics of the enactment, the participants collaboratively elaborate on this “insecurity” that characterized their experience in the workplace.

After the verbal dimension, we now focus on the bodily cues that occur before the onset of the quotation (l. 01). Figure 4 illustrates Lena’s bodily displays before the actual quotation starts.



Figure 4: Bodily displays before the quotation, Example 2.

At the beginning of l. 01, Lena starts forming an iconic hand gesture in front of her upper body as if she were holding a triangle and preparing to play. Lena freezes her hands in this position, which is visible to Rosa (#1). While at the beginning of l. 01, Lena gazes into the middle distance, she then directs her gaze toward Rosa and simultaneously starts smiling (#2). Her gaze then wanders into the middle distance again. At this moment, Lena inhales and produces a pause before the upcoming quotation. Simultaneously, Lena drops her lower jaw and relaxes her facial expression slightly (#3).

Lena’s use of her bodily resources can be interpreted as a projection of the upcoming enactment. Most importantly, the iconic hand gesture and its freezing can be interpreted as pantomimic since Lena partially adopts the bodily configuration of the narrated scene, i.e., waiting to play the triangle. Lena’s gaze can also be interpreted in this way: While looking into the middle distance when starting to embody the story character, her shift of gaze toward Rosa can be interpreted as “enlisting” (Sidnell 2006) her into the corresponding role of a coworker in the story world to recruit her help. At the same time, this shift of gaze toward Rosa may be interpreted as inviting her to take part in the projected enactment. Lena’s facial expression similarly frames the upcoming enactment as amusing by her smile. At the same time, the drop of her mouth co-indexes the uncertainty of the story character.

To focus on the organization of the co-production of this enactment, we now consider Rosa’s bodily displays before the quotation (Figure 5).

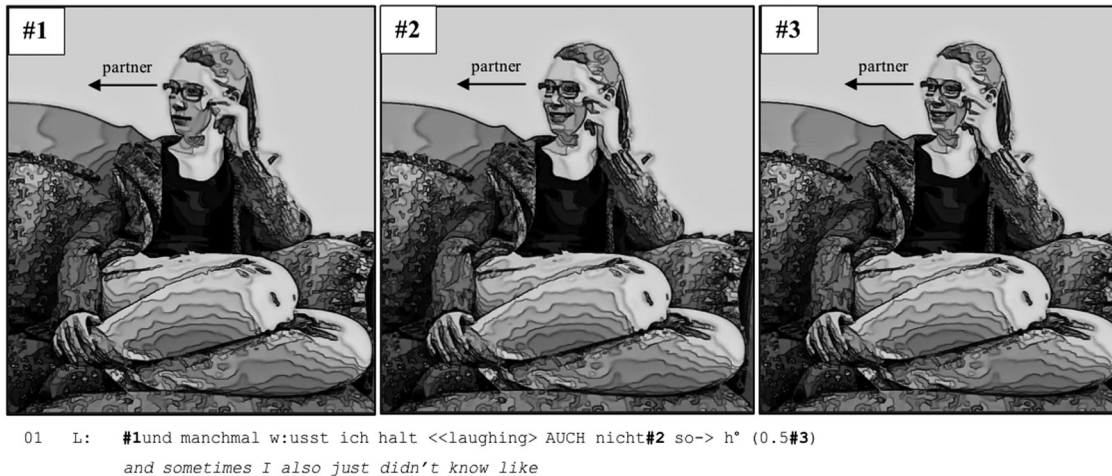


Figure 5: Rosa's bodily display during Lena's projection (Example 2).

Rosa closely monitors Lena by constantly gazing at her face. She changes her facial expression while sitting upright: Almost at the same time as Lena starts laughing, Rosa starts smiling (#1 and #2), ending in a big smile at the end of the quotative (#3). Thus, Rosa already takes up the framing of the enactment as “amusing” before the actual quotation starts. Figure 6 presents a visualization of the multimodal resources as they occur throughout the co-enactment.

With respect to the verbal dimension, the red line in Figure 6 illustrates the simultaneous onset of the co-produced quotation. Furthermore, the figure shows that Lena has already started to enact before the onset of the quotation via bodily and then phonetic resources. Hence, when the timing of the different dimensions is taken into account, there is no sharp boundary between the enactment and the preceding talk but rather a smooth transition into the enactment, which can be understood as a cesural area. Within this area, the different resources gradually accumulate and establish projections, which are then resolved in the multimodal compaction zone. It is within this multimodal compaction zone that different multimodal resources culminate and where, with the production of a quotation, the (possible) completion of the emerging enactment is reached. The first quotation presents only a “possible” completion of the emerging gestalt since the enactment is extended over several quotations. Importantly, the projections established by Lena within the cesural area allow her co-participant Rosa to not only anticipate the upcoming quotation but also to co-produce it with a precisely timed onset (cf. *joint-action*, Clark 1996). We argue, therefore, that weak cesuras and gradually accumulated projections at the beginning of enactments in joint storytellings open an interactional space in which co-participants have the opportunity to co-produce the enactment. Furthermore, within this action space, participants mutually take up elements from each other's contributions and jointly work out the semantics of the enactment. While these uptakes are most prominent on the verbal level within the compaction zone, they may take place before the quotation even starts, as evidenced by Rosa's uptake of Lena's smile.

Summing up the results of the analysis of example 2, we would like to emphasize three points. First, the projections established by Lena in the cesural area allow Rosa to join and co-produce the quotation in a micro-timed onset and thereby participate in the action of enacting. Indeed, Rosa co-produces a quotation that fits syntactically, pragmatically, and semantically into that projected space. She thus displays not only her understanding of the emerging action but also co-constitutes that action. Secondly, Rosa takes up Lena's smile before the actual quotation to show not only her involvement but also her alignment with Lena's perspective and affective display (Stivers 2008). Thirdly, we find a mutual elaboration of the semantics of the enactment, which is reflected in the mutual uptakes of the verbal structures in the quotations.

While in this example the projection involves a projecting verbal structure, we show in the next example that already bodily projection on its own can create a space in which the partner-in-interaction can co-produce the enactment.

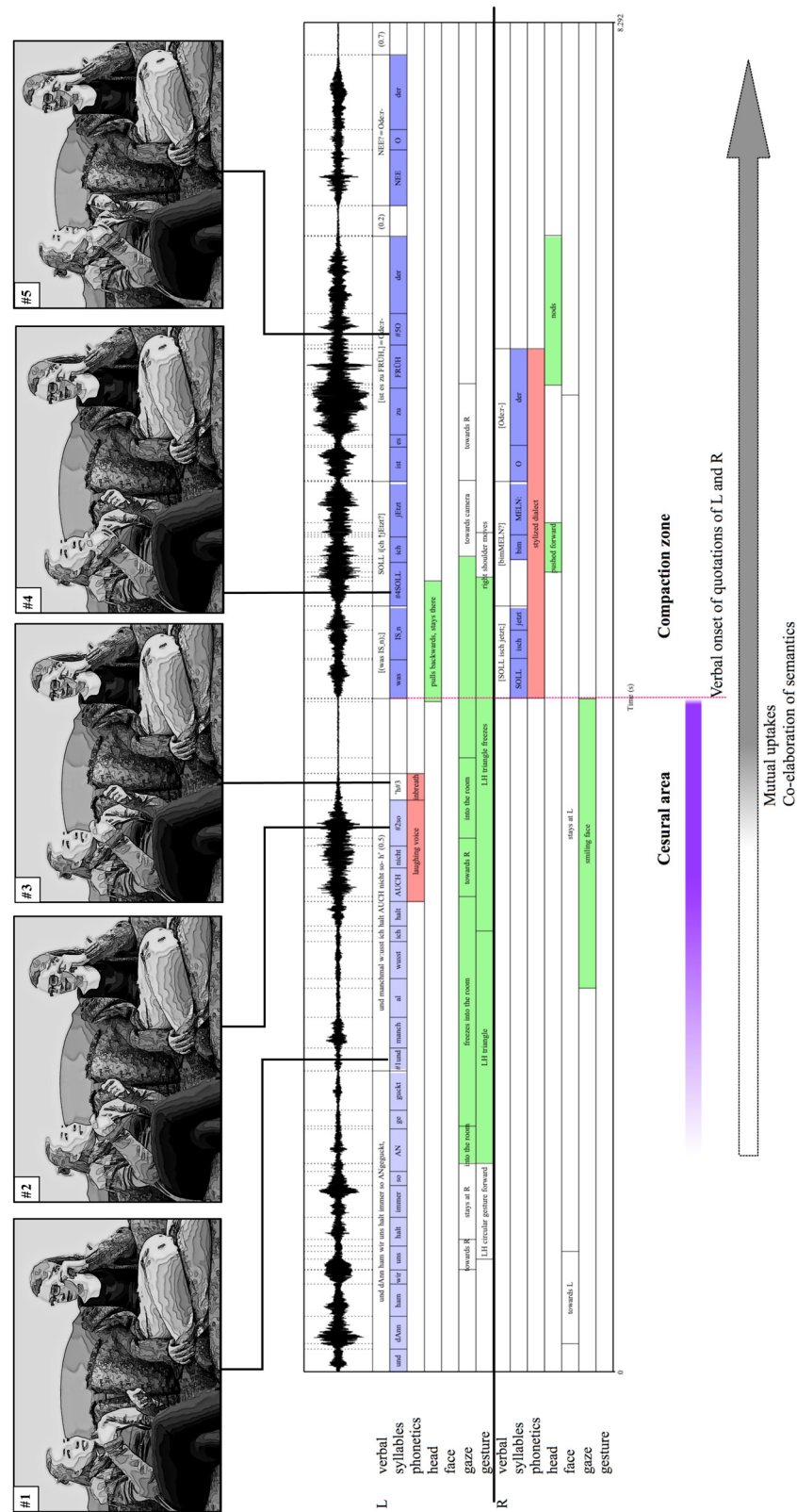


Figure 6: Visualization Example 2.



## 4.2 Co-constructed quotation following bodily projection

Two friends, Luise and Rebecca, are sitting side by side on a sofa (Figure 7). Luise is recounting a past event that they experienced together: When they studied at university, they had to submit documents before taking the final exams. When Luise asked her friend whether all her documents were complete, Rebecca realized that she had forgotten to fill in the application form that was necessary for successful registration. The main teller of the story, Luise, stages the past dialogue between the two by first enacting herself and then Rebecca's reaction. Here, we are especially interested in the second quotation since the transition between the story characters, "Luise" and "Rebecca," is realized without any quotative and it is at this point that Rebecca joins in the enactment. Again, we first focus on the sequential development of the example.



Figure 7: Setting of Example 3.

**Example 3.** *Rebecca's application* (ger\_sft11, 1:22-1:28)

- 01 L: <<all, :-)> (un)\_ich\_gSAGT>-=  
and I said
- 02 =<<all> und den ZUlassungsantrag;>=  
and the application form
- 03 =den hast du AUCH ausgefüllt;>  
you filled out that too
- 04 (0.6)
- 05 <<creaky, pp> öh->  
aehm
- 06 R: <<:-), pressed> °h\_w:As ist der ZUlas[sungsauf<<laughing>\_trag?>> ]  
what is the application form
- 07 L: [wAs ist der <<:-)> ZUlas]  
[sungsantrag;>] what is the application  
form
- 08 R: [((laughs)) ]
- 09 <<:-), pressed, h> ich WUSSte es nicht;>  
I didn't know about it



With <<all, :-)> (un)\_ich\_gSAGT>= (l. 01) “and I said,” Luise utters a quotative with the verbum dicendi *say* to introduce the first quotation. Luise quotes herself directing a question at Luise: =<< all> und den Zulassungsantrag;>= | =den hast du AUCH ausgefüllt;> (l. 02–03) “and the application form | you filled that out too.” Phonetically, the quotation is stylized as an “innocent” question given the fast pace, a smiling voice, and a flat intonation curve. As a first pair part, this question projects an answer in the story world. The projected answer, however, is not produced immediately afterward. Rather, it follows a relatively long pause of 0.6 seconds (l. 04), which contextualizes a dispreferred response. In line with this, the quotation of the response produced by Luise starts with the particle [<<creaky, pp> öh->] (l. 05). This particle functions not only as a hesitation marker but also as a change-of-state token, which “is used to propose that its producer has undergone some kind of change in his or her locally current state of knowledge, information, orientation or awareness” (Heritage 1984, 299). Hence, the quotation starts with a vocal preface. In general, prefaces have the function of establishing an “indexical frame” (Auer 1996) for the interpretation of the upcoming turn on different levels of discourse (Heritage 2013, Schegloff and Harvey, 1973). On the phonetic dimension, Luise’s creaky voice indexes that the quoted story character is stunned and perplexed: The vocal beginning of the quotation is thus formatted to provide an interpretative frame for the following verbal part. Moreover, note that the change from one quoted character to the other is not explicitly framed by a quotative. Nevertheless, Rebecca is perfectly capable of correctly interpreting the upcoming quotation since she co-produces it with <<:-), pressed> °h\_w:As ist der Zulassungsantrag?>>] (l. 06) “what is the application form.”<sup>1</sup> The quotation is a nontype-conforming response to the preceding quoted question since it does not directly (dis-)confirm but presents a counter question (Raymond 2003), thereby signaling a deviation from the presuppositions of the preceding quoted question. Rebecca’s astonishment is further contextualized by Luise’s stylized lengthening and hesitation at the onset of the quotation on the question word *w:As* (l. 06). Luise thus forms her quotation in a way that reflects the same semantics as Rebecca’s astonished *öh*-preface but in a more verbally explicit way. Furthermore, Lena’s laughter particles mark the quoted question as the punch line of the story (Sacks 1974). Rebecca thus signals that she is able to co-construct not only the content of the quotation but also its significance for the overall ongoing activity of the storytelling. Subsequently, Luise recycles Rebecca’s quoted question with: [wAs ist der <<:-)> Zulassungsantrag;>] (l. 07) “what is the application form.” She hence ratifies Rebecca’s co-produced quotation as appropriate.

To sum up the analysis up to this point, we note that both speakers co-produce the enactment, mutually elaborating on and taking up potentially meaningful elements produced by the other participant. We can thus conclude that they jointly accomplish the enactment and its importance for the overall course of the telling. As previously stated, the action space of the co-produced enactment in this example is not projected by verbal means. Rather, as Figure 8 shows, Luise activates multiple bodily resources to organize the transition between the two story characters.

As she produces the first quotation, Luise smiles widely, her eyes are widely open, and she holds her head upright (#1). In the subsequent pause, Luise abruptly changes her bodily and facial expressions and starts embodying her friend “Rebecca” in the story world. Luise nods in the direction of her gaze, which is directed into the middle distance, away from her co-participant Rebecca. Simultaneously, she drops the corners of her mouth, stops smiling, and begins staring (#2). She then slightly moves her head backward, dropping the corners of her mouth even further (#3). In this position, she produces the vocal particle [<< creaky, pp> öh->] (l. 05). Thus, Luise first changes her facial expression, alters her head position, and diverts her gaze to index and project the hesitation and perplexity of the story character “Rebecca,” which she then expresses more explicitly with the particle *öh* (l. 05) at the beginning of the quotation.

Up until this point, Rebecca was sitting still and looking at Luise. Only when Luise begins to bodily enact the story character, “Rebecca” does she start smiling. Simultaneous with Luise’s <<creaky, pp> öh-> (l. 05), Rebecca quickly turns her head toward the interviewer and then starts co-enacting the scene herself, also

<sup>1</sup> Rebecca’s slip of the tongue, ‘Zulassungs**auftrag**’ (in English: “order form”) instead of the version ‘Zulassungs**antrag**’ (in English: “application form”), is not relevant for the purposes of our analysis.



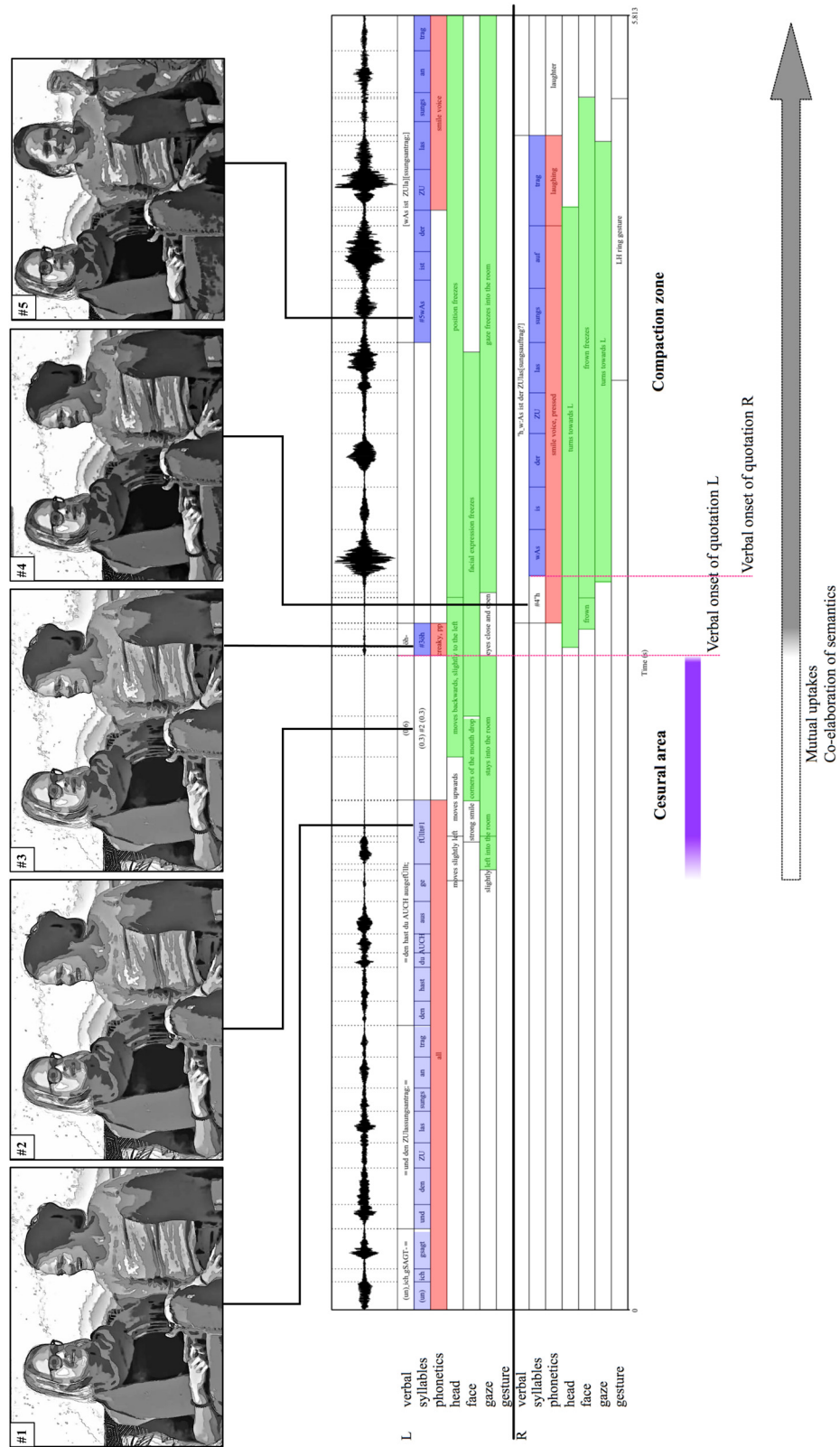
**Figure 8:** Preparation phase and first particle of the quotation by Luise (example 3).

proceeding first on the bodily and then on the verbal dimension. By using smile voice followed by laughter, she indexes the quotation as the punch line of the narration (l. 08). Figure 9 summarizes the timing and trajectories of the co-produced enactment.

Figure 9 shows that Rebecca starts her co-produced quotation immediately after Luise's *öh*-preface and that Rebecca then continues to co-produce the quotation in temporal overlap. Furthermore, Rebecca has already started embodying the Figure "Rebecca" by frowning and turning toward the Interviewer while Luise is still producing the *öh*-preface. Figure 9 also illustrates that Luise, as the main teller, employs gradually accumulating bodily cues to index the change from embodying one story character to another, before starting the quotation. We argue that it is these bodily cues, together with the question and projected answer sequence on the verbal dimension, which allow her co-participant to anticipate the upcoming enactment. These projections, which are established in the cesural area, open—within the activity of joint storytelling—the action space of a co-enactment.

While we have thus far focused on the opening of the action space, we now turn to an analysis of how the participants "work out" and jointly construct the meaning of the enactment. The first quotation by Luise does not start with a lexical element but rather the vocal particle [<< creaky, pp> *öh*->]. This turn preface—together with the phonetic and bodily resources—serves to index some sort of hesitation and perplexity. This indexed perplexity is then formulated more explicitly by Rebecca's quoted question: <<:-), pressed> °h\_w:As ist der Zulas[sungsauf <<laughing> \_trag (l. 06) "what is the application form." Rebecca thus transfers the indexed semantics to the symbolic level. By speaking with a smiling and pressed voice, she also takes up the humorous footing of the narration, which has already been established by Luise, and presents the quotation as the punch line of the story. Rebecca's co-enactment is then ratified by Luise, who takes up the semantics and the lexico-syntactic structure of Rebecca's quotation in her own following quotation. It is to be noted that these mutual uptakes do not take place in clear succession but are produced in partly overlapping speech instead. Therefore, this can be viewed as a stepwise and collaborative multimodal joint construction of the enactment.

The analysis of Examples 2 and 3 has shown that the gradual build-up of multimodal projections in an enactment may lead to a more or less simultaneous onset of a co-produced quotation. In addition, and more importantly, both examples are characterized by the joint development of the semantics and a particular stance taken in the co-enactment as proceeding in a stepwise fashion, taking up and transforming the cues present in the projection. This is also the case in Example 4, although the quotations by the two participants are not produced in overlap but successively.



### 4.3 Joint meaning making in successive quotations

Two friends, Leslie and Resi, are sitting on a sofa (Figure 10). Leslie is telling a story about a trip she went on with Resi to the United States. When the two friends arrived in New Orleans, they consulted their travel guidebook to look for further information about the city, in which they read that they should stay in a specific area of the city for their own safety. The women, however, did not take the warning seriously. The following excerpt starts with Leslie quoting the travel guidebook. Here, however, we are interested in the quotation that starts in l. 06 when the participants collaboratively enact their stance toward the advice.



Figure 10: Setting of Example 4.

**Example 4.** *Narrow minded* (ger\_sft13, 18:45-18:52)

- 01 L: und es HIEß so==  
and it sounded like
- 02 =ja: \_also in dEm corner <<creaky> könnt ihr euch AUFhalten;=  
yes so you can stay in that corner
- 03 =aber !NIR!gendwo <<:-)> ANders,>  
but nowhere else
- 04 R: [hm\_HM; ]  
hm hm
- 05 L: [un\_wir DACHten] <<creaky> so;>  
and we thought like
- 06 <<creaky> ah::;>  
ah
- 07 R: <<f> jA pff>  
yes pff
- 08 R: [((laughs)) ]
- 09 L: [<<f> voll !SPIE!Bi:g;>]  
totally narrow-minded

Leslie produces an impersonal quotative = und es HIEß so- = (l. 01) “and it sounded like,” which includes both a verbum dicendi and the TCU final so “like” to project a quotation from the travel guidebook. With the following quotation, Leslie changes her perspective and presents the guidebook’s warning as if it were directly addressed toward the two travelers: ja:\_also in dEm corner <<creaky> könnt ihr euch AUFhalten; = | = aber!NIR!gendwo <<:-)> ANders,> (l. 02-03) “yes so you can stay in that corner | but nowhere else.” In the verbal design of the quote, we hear not only the “original” voice of the guidebook but also the negative stance that Leslie takes toward the warning in a polyphonic way (Günthner 1999, 2002). This stance is indexed by phonetic means (accentuation and lengthening of the beginning of the TCU = ja:\_also (l. 01) “yes so”) and the strong accent on the adverb !NIR!gendwo “nowhere.” This negative stance builds up an interpretative background as to how the story is going to continue, i.e., the enactment that we are interested in.

When Resi produces the confirming continuer [hm\_HM; ] (l. 04), Leslie introduces the quotation of an inner thought with [un\_wir DACHten] <<creaky> so;> (l. 05) “and we were thinking like.” The quotative in the 1st person plural contextualizes this projected quotation as “joint thought,” claiming that both women were thinking the same thing regarding this piece of advice. Leslie starts the quotation with the lengthened vocal particle <<creaky> ah;> (l. 06) “ah,” a turn-prefacing element. As in Example 3, the vocal element is used as a vehicle for phonetic stylization since it is uttered with a creaky voice, contracting the vocal tract and retracting the tongue. These phonetic parameters on the particle express an affective display of “mockery,” which is further supported by a “dropping” hand gesture, an observation that will be returned to.

After Leslie’s production of the preface to the quotation, Resi continues the quotation with <<f> jA pff> (l. 07) “yes pff.” The quotation consists of the response particle yes, which acknowledges the advice given in the guidebook in a concessive way. The following sound object *pff* (Reber and Couper-Kuhlen 2010) is a conventionalized way of expressing either “helplessness” or “depreciation,” of which the latter applies in this sequential context. Resi thus explicitly takes up Leslie’s prior affective display expressed by the vocal particle <<creaky> ah;> (l. 06) “ah” and transfers it to her continuation of the quotation with <<f> jA pff> (l. 07) “yes pff,” elaborating it further with added symbolic material.

Leslie continues the quotation further with [<<f> vOll!SPIE!Bi:g;>] (l. 09) “super narrow-minded” so that the women’s stance toward the guidebook is now explicitly expressed. In addition, Leslie puts extra stress on the adverb and lengthens the last syllable of the adjective, thereby marking the utterance as the punch line of the story. Hence, the continued quotation serves not only to elaborate the prior quotations in terms of the semantics but also with regard to the display of affection. This is further evidenced by the subsequent appreciation of this elaboration by Resi, expressed in her laughter in the next line (l. 08).

As in the previous examples, the quotation is preceded by the use of several multimodal resources (Figure 11).

At the beginning of l. 05, Leslie tilts her head backward and shakes it slightly. In addition, she assumes a grumpy face by bunching her eyebrows and pulling the corners of her mouth slightly sideward (#1). While maintaining this facial expression, she turns her gaze toward Resi (#2) and directs an open-palm-up pointing gesture toward her (#2). This gesture can be seen as an invitation for Resi to take the turn, both as co-present interlocutor and as an “enlisted” story character. These bodily cues are micro-timed with the quotative particle “so” (l. 05). As she utters the *ah*-preface, Leslie abruptly bends her wrist downward, transforming the palm-up gesture into a throw-away gesture (Bressemer and Müller 2014), which co-indexes the woman’s disrespect for the guidebook’s warning in a bodily manner. Simultaneously, Leslie starts smiling and thereby contextualizes the mockery as amusement (#3).

This analysis reveals that Leslie projects the quotation not only verbally by employing a quotative (l. 05) but also by using bodily cues, with which she begins embodying the enacted figure. Importantly, Leslie’s grumpy face already establishes the framing of the enactment as mocking the warning. In addition, Leslie also gives Resi the opportunity to co-produce the upcoming quotation by turning and gazing toward her; Resi responds to this invitation in a remarkable way (Figure 12).

Immediately after Leslie’s quotative so (l. 04) “like,” Resi establishes mutual gaze (#4). She then slightly shakes her head and recycles Leslie’s prior throw-away gesture with her left hand (#5). When



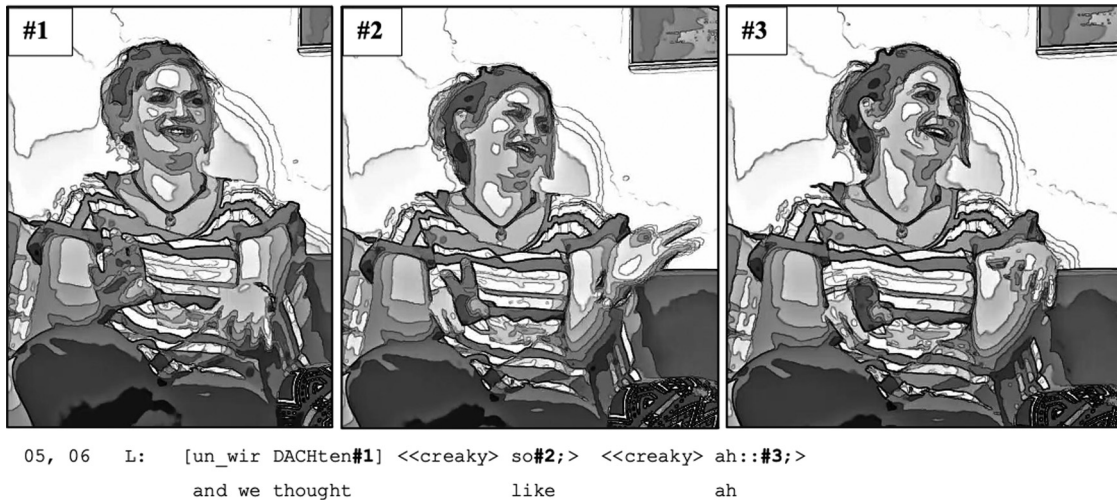


Figure 11: Bodily displays and first particle of the quotation (Example 4).

uttering <<f> ja pff> (l. 07) “yes pff,” she rolls her eyes, emphasizing the negative evaluation. This eye-roll is later taken up by Leslie, more specifically, when she continues the quotation with [ <<f> v011!SPIE!Bi:g;>] (l. 07) “super narrow-minded” and additionally throws her head backward and opens both hands, just like Resi (#6).

As in the previous examples, the participants jointly elaborate the enactment. When looking at the whole verbal expression from a post hoc perspective [un\_wir DACHten] <<creaky> so;> | <<creaky> ah;> | <<f> ja pff> | [<<f> v011!SPIE!Bi:g;>] (l. 04, 06, 07, 08) “and we thought like | ah | yes pff | totally narrow-minded,” one could get the impression that only one speaker produced the whole utterance. We

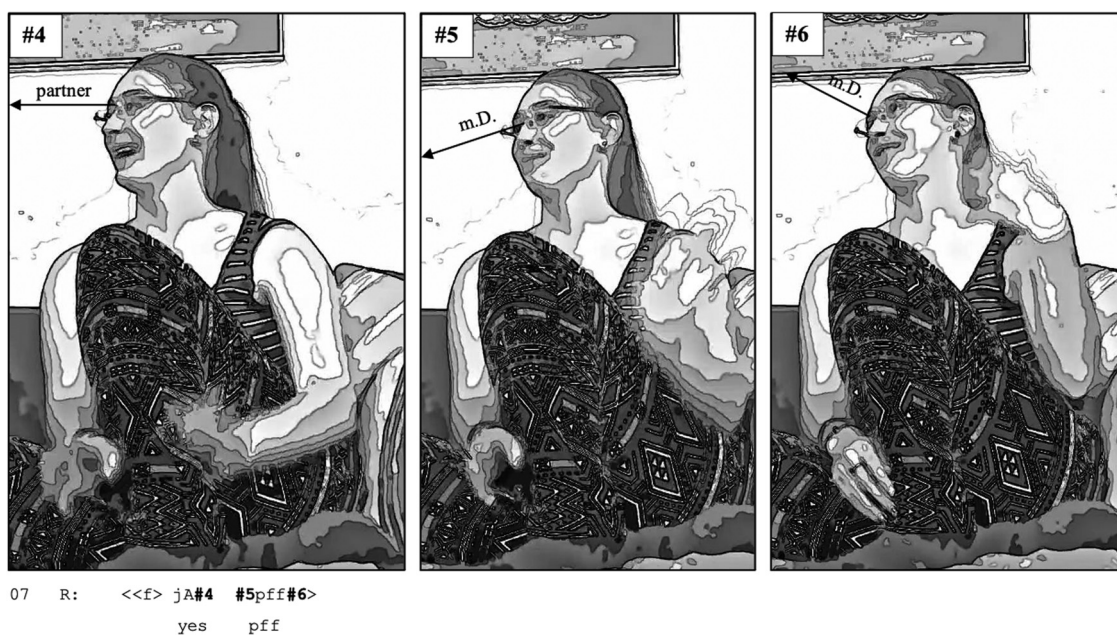


Figure 12: Resi's uptake (Example 4).



would like to emphasize, however, that the participants rather collaboratively “work out” and gradually develop the meaning of the enactment. This is realized by building on both bodily and verbal aspects of the interlocutor’s preceding contribution and its meaning potentials. Regarding the bodily dimension, the participants mutually pick up, recycle, and transform expressive elements, such as the throw-away gesture and eye rolling. On the verbal and phonetic dimensions, indexical cues are taken up and gradually transformed into more symbolic means. In doing so, the participants not only co-produce the semantics of the enactment but also reassure each other as to their previous evaluation of the guidebook’s warning as sharing the same stance (Niemelä 2011).

The following Figure 13 illustrates the timing and trajectories of the different resources in the co-produced enactment.

On the verbal dimension, Figure 13 illustrates that the first quotation produced by Leslie starts with a particle preface, and Resi begins to co-produce the quotation immediately afterward. The development of the enactment on the verbal level is thus the same as in Example 3. Example 4 differs from the preceding examples, however, since the first enactor (Leslie) does not continue her quotation in overlap with the co-enactor (Resi) but only once Leslie has finished the quotation.

With respect to the phonetic and bodily dimensions, Figure 13 illustrates that the indexical cues on these dimensions precede the verbal level. Thus, as in the preceding examples, the enactment is characterized by a smooth fading in and gradual cumulation of indexical cues, all of which form a cesural area. The projections established in this cesural area are interactionally relevant in the sense that they open up a joint action space for co-enactment. This joint development is particularly prominent in two respects. First, the two interlocutors elaborate and maintain the action of “quoting” over three TCUs. By mutually taking up and elaborating on elements of the co-participants’ prior contributions, they jointly *work out the semantics* of the enactment. This process of mutual up-taking starts in the cesural area before the actual quotation. Second, the two interlocutors become involved in a similar manner and *develop a shared stance*. The affect display of mockery is produced and mutually signaled by the employment of both phonetic (lengthening, stressing) and bodily resources (throw-away gesture, eye rolling), which indicate the participants’ mockery of the guidebook’s warning. In addition, they share their stance by smiling at each other. Such co-productions thus serve joint sense-making in interaction in terms of both the semantics of the enactment as well as the participants’ joint stance toward it.

## 5 Results

We can sum up the results of our analysis in three respects: the *multimodal realization* of enactments, the interactional *opening of an action space*, and the dynamics of *joint elaboration on the content* of the enactment.

A first result with respect to the *multimodal realization of enactments* is that enactments are indeed realized as multimodal gestalts (Mondada 2014). Enactments emerge over time via the combination and accumulation of multiple indexical cues on not only the verbal and phonetic dimensions but also the bodily dimension, which includes facial expression (smiles, frowns), gaze, hand and body movements, and orientation. With respect to the temporal onset of resources across the different dimensions, we have shown that, for the most part, they are not aligned: bodily cues typically precede phonetic and the verbal cues. An enactment is thus characterized by a gradual increase of indexical cues, which leads to an increase in expressiveness and culminates in a multimodal embodied expression in the compaction zone. The gradual accumulation of these resources at the beginning of an enactment leads to a cesural area between the preceding talk and the enactment itself.

The *cesural areas analyzed are interactionally relevant* in that they open up an opportunity space for joint action: the projections that are established in the cesural area create an action space in which participants other than the main teller may (but do not have to) co-produce the enactment. In his seminal paper, Lerner (1991) points out that the first part of a compound turn-constructive format projects the

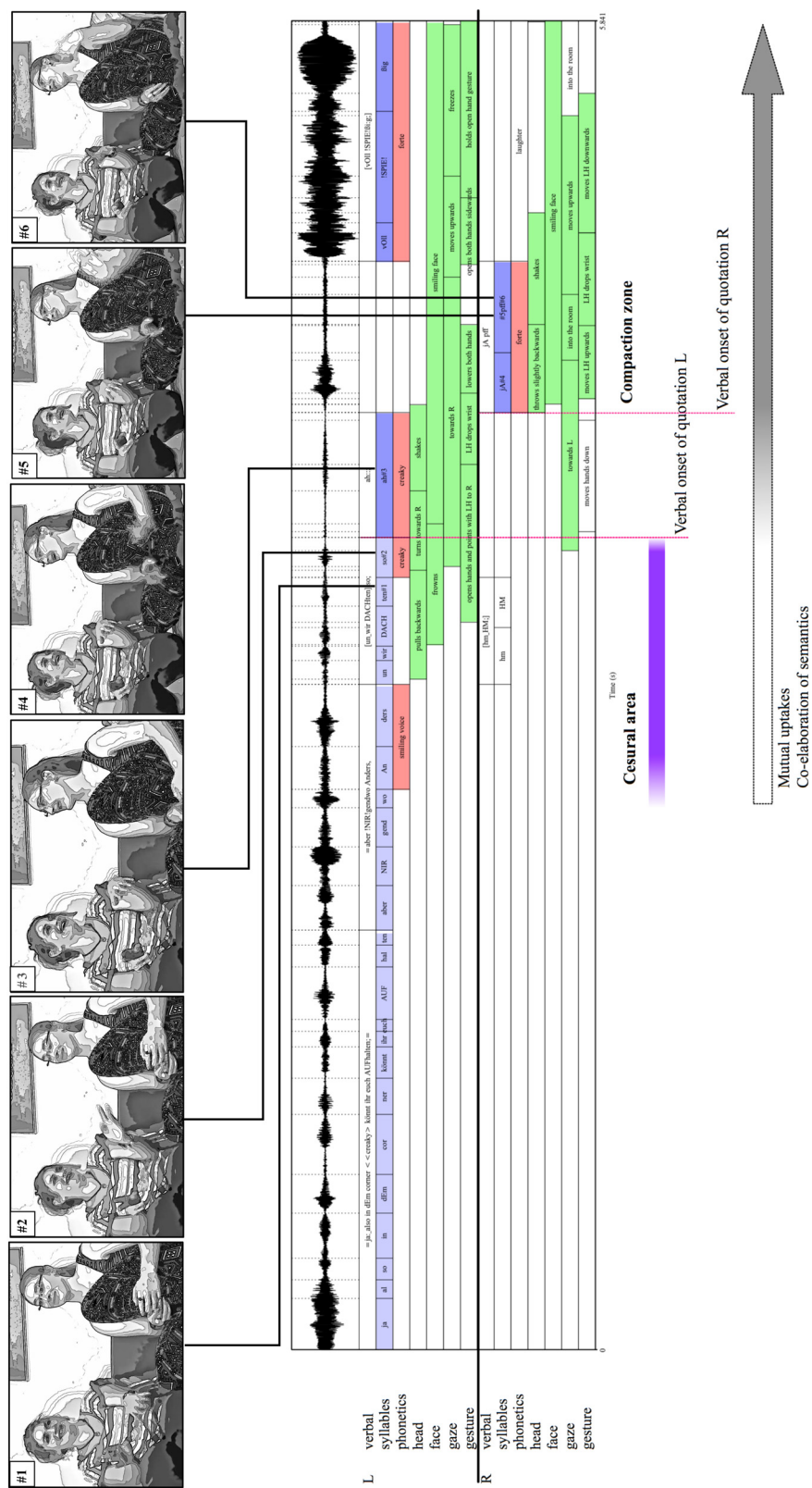


Figure 13: Visualization Example 4.

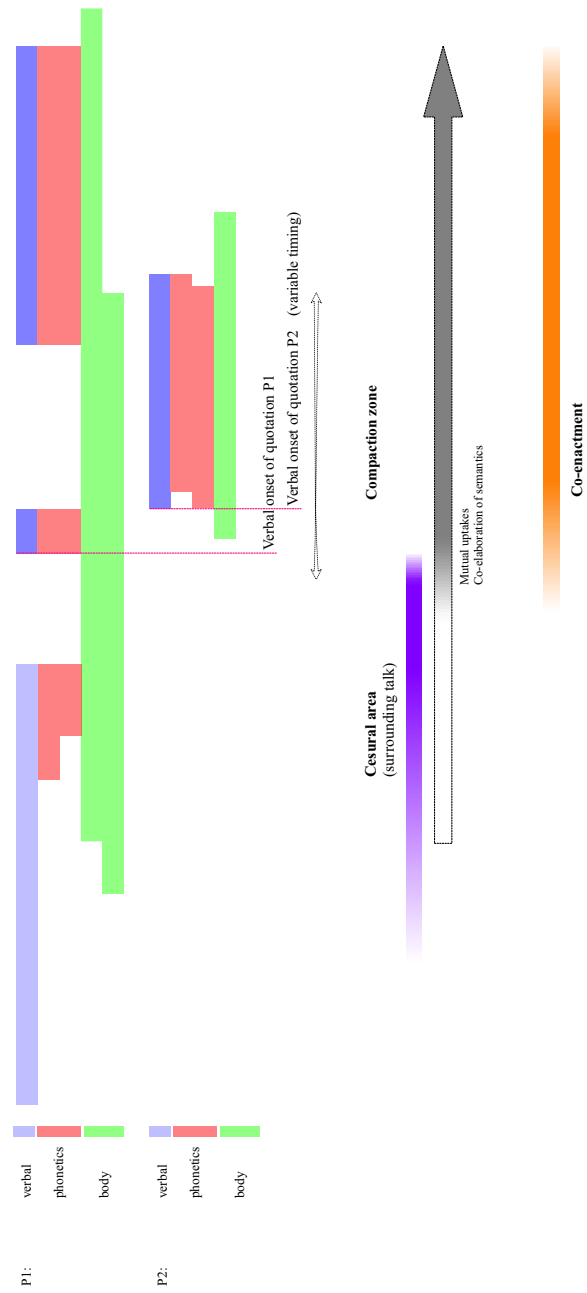


Figure 14: Schema co-enactments.

second part and thereby opens an opportunity space for joint action. Our results support this claim. In particular, the data reveals that the projection of the second part may also heavily rely on bodily resources. For cases in which there is no verbal projecting element as a “first part,” the projection may be established solely on the bodily dimension. Importantly, the action space in our data is also used by the participants to fulfill tasks that are central to the overarching activity of joint storytelling. The participants jointly reconstruct what has happened or has been said (epistemic), negotiate a stance (deontic), create involvement (emotional alignment), build up a certain footing (e.g., humorous), and fashion the tellability of the story (punch lines, etc.). Co-enactments seem particularly well suited for the joint accomplishment of storytelling in situations in which both tellers have experienced the situation together. While narrations in general involve a specific participation framework where there is only *one* participant as the main teller, our data shows that co-participation can be invited into such spaces. Cesural areas in enactments are thus a resource that opens up an action space or “opportunity space” in which a co-enactment can take place. In general, co-enactments are spaces in joint storytelling where participants may collaboratively complete tasks that are central for this activity (Figure 14).

As Figure 14 shows, participants use action spaces in different ways to co-produce an enactment in the compaction zone (see the graded orange line). The onset of quotations by co-participants can be precisely timed (see Example 2) or the second participant may start just a little later (see Examples 3 and 4). The further continuation of the co-production may overlap (see Examples 2 and 3) or occur successively (see Example 4). Despite such differences in the timing of the co-produced quotations, the co-development of the meaning of these enactments is very similar.

In the opportunity space, we find *elaboration on the meaning* of the enactment. Both participants jointly work out, and elaborate on, what is being expressed in the enactment by mutually taking up indexical cues and verbal material from prior contributions made by the respective co-participant. These mutual uptakes have typically already started before the quotation itself. Here, there is usually a framing of the upcoming enactment (such as verbs describing the events, laughter particles indexing the footing, and changes in voice quality) and bodily indicators of the inceptive enactment of the story character (changes in posture, mimic gestures, changes in voice quality). These cues are frequently responded to even before the actual quotation. In the multimodal compaction zone, we find several versions of a quotation produced by both participants. Throughout the course of the enactment, quotations are elaborated on by participants, often starting in fragments and ending in full clauses. What is explicitly expressed in the quotation has, for the most part, already been indexed beforehand, either bodily or phonetically. In addition, quotations often start with a nonlexical prefacing element (e.g., *öh*, *ah*). Such prefaces within the quotation serve as further framing devices or as a “continuation” of the preceding indexical cues. While the quotations in many cases could be read—from a post hoc perspective—as one syntactic project, they are actually produced collaboratively by the co-participants. In this sense, these enactments are polyphonic and polykinetic in a dual sense: polyphony and polykinetics do not only hold between figure and enacter (Günthner 2002, Keevallik 2010) but also between different enactors, who elaborate on the quotation by building upon previously given material.

## 6 Discussion

This article has analyzed co-enactments taken from a corpus of joint storytellings. Enactments are understood to be animations of an entity (human or other, the speaker him/herself or other) or “demonstrations” of an action (rather than “tellings”). The analysis was based on a collection of co-enactments in which two tellers co-produce the verbal part of an enactment, i.e., a quotation, and thus involve themselves in a specific turn-sharing practice.

This article focuses on the transition between the preceding talk and a quotation. As noted in the previous research, there is often no clear-cut shift from the preceding talk to the quotation but a smooth transition, whereby the enactor already begins to embody features of the enactment before the quotation

itself. A general conclusion to be made from our study is that all of the analyzed co-enactments are characterized by a cesural area in which the current main teller of the story projects the enactment. This projection is realized not only verbally (if at all) but also phonetically and, in particular, bodily.

We, therefore, argue that cesural areas in enactments are interactionally relevant in the sense that they facilitate co-production. In other words and in a more general sense, transitions and projections can open up a joint action space (Lerner 1991) in which other participants are given the opportunity to take part in the local action, i.e., to co-produce the enactment in joint storytellings. This argument is in line with the previous research: enactments more generally are interactional spaces that are characterized by “heightened displays of reciprocity and attention” (Sidnell 2006: 406, also Holt 2000). Co-enactments are a particular way of displaying attention and signaling understanding, and more specifically, the co-participant becomes more than a mere recipient but an active producer. In addition, we have shown that co-enactments are interactional spaces in which the participants jointly construct the meaning of an enactment in a sequential fashion by taking up and building upon previous material. By taking part in such joint productions, the participants also negotiate the closure of the emerging *gestalts* by extending them further.

While previous research on cesural areas in interaction has focused on phonetic parameters, we have highlighted the importance of bodily resources. By applying a parametric approach, several resources on the bodily dimension have been systematically considered with regard to their interplay on the verbal and phonetic dimensions. In our view, the *cesura* approach can be applied fruitfully to the bodily dimension, too. Due to its parametrical and gradual character, it also seamlessly integrates with Stukenbrock’s (2018) notion of *multimodal compaction zones* and the concept of emerging *multimodal gestalts* (Mondada 2014).

Of course, how the limits of co-participation should be defined remains an open question. In this study, we centered our analysis on instances in which a co-participant produces a quotation. In our data, however, we also find many enactments in which a co-participant does not produce a quotation but nevertheless co-enacts “nonverbally.” For instance, co-participants often move just their mouths without producing any sound, shadowing the articulatory movements of the main enactor simultaneously to his/her quotation.<sup>2</sup> Moreover, we often find cases in which the co-participant does not produce any articulatory movement but simultaneously co-enacts bodily and moves as if (s)he were in the enacted situation, e.g., by adopting a correlating body posture, moving the torso accordingly, or even by co-performing gestures. Of course, we also find cases in which the bodily displays of an enactment are difficult to interpret due to their subtlety. We, therefore, argue that co-enactments are not only limited to participants actually co-producing a quotation but rather that co-enactments can be solely in the form of bodily co-productions and that there are “degrees” of co-production (see Mandel, In prep.). In this sense, co-enactments that are not simultaneously but consecutively produced (as in Example 3) should be considered as one of a variety of different ways in which co-participation can occur in an enactment. Such types of co-enactments, in our opinion, deserve further investigation and could broaden the overall picture with regard to which phenomena should be included in the realm of joint action. What we currently take for granted is the fact that bodily projections and cesural areas are regularly found at the transition between the preceding talk and the quotation of the enactment.

The transitions analyzed in this article can be considered “weak” cesuras that form a cesural area between one segment and the next in interaction. Such cesuras are created by the interplay of various multimodal parameter changes, e.g., on the verbal, phonetic, and bodily dimensions. The results of this study support the hypothesis that weak cesuras and cesural areas are important for the organization of talk-in-interaction. Weak cesuras in enactments are interactionally relevant in that they contribute to opening an opportunity space in which participants may involve themselves, even if the degree of involvement varies, and act together.

**Acknowledgements:** We would like to thank the editors of this special issue, Richard Ogden and Dagmar Barth-Weingarten, in particular, for their very fruitful comments on a prior version of this article. A special

<sup>2</sup> Example 1 is an instance of such a case: close observation of the video reveals that the co-participant, Renate, voicelessly articulates the quotation while Lucas verbally quotes the enacted figure.

thanks also goes to our anonymous reviewers for their valuable comments. In addition, we are indebted to Stefan Pfänder and his team for providing us with the corpus. We also thank Laura Cuthbertson for proof-reading. Of course, all remaining shortcomings are our own responsibility.

**Funding information:** This work has been partially funded by the Deutsche Forschungsgemeinschaft under grant 391351163, which we gratefully acknowledge.

**Author contributions:** All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

**Conflict of interest:** The authors state no conflict of interest.

## References

- Auer, Peter. 1996. "The pre-front field in spoken German and its relevance as a grammaticalization position." *Pragmatics* 6(3), 295–322.
- Auer, Peter. 2005. "Projection in interaction and projection in grammar." *Text and Talk* 25(1), 7–36.
- Auer, Peter. 2006. "Construction grammar meets conversation: Einige Überlegungen am Beispiel von 'so'-Konstruktionen." In: *Konstruktionen in der Interaktion*, eds. Günthner, Susanne and Wolfgang Imo, p. 291–314. Berlin: de Gruyter.
- Auer, Peter. 2010. "Zum Segmentierungsproblem in der gesprochenen Sprache." *InLiSt* 49, 1–19.
- Auer, Peter. 2015. "The Temporality of language in interaction: Projection and latency." In: *Temporality in Interaction*, eds. Deppermann, Arnulf and Susanne Günthner, p. 27–56. Amsterdam: Benjamins.
- Barth-Weingarten, Dagmar. 2008. "Interactional linguistics." In: *Handbook of interpersonal communication*, eds. Antos, Gerd, Ventola Eija and Tilo Weber, p. 77–106. Berlin: de Gruyter.
- Barth-Weingarten, Dagmar. 2013. "From 'intonation units' to cesuring – an alternative approach to the prosodic-phonetic structuring of talk-in-interaction." In: *Units of talk – units of action*, Series "Studies in language and social interaction", eds. Szczepek Reed, Beatrice and Raymond, Geoffrey, p. 91–124. Amsterdam: Benjamins.
- Barth-Weingarten, Dagmar. 2016. *Intonation units revisited – cesuras in talk-in-interaction*. Amsterdam: Benjamins.
- Barth-Weingarten, Dagmar and Richard Ogden. 2021. "'Chunking' spoken language: Introducing weak cesuras." *Open Linguistics* 7(1), 531–48. [Special Issue: Weak cesuras: what fuzzy boundaries can accomplish in talk-in-interaction, edited by Dagmar Barth-Weingarten & Richard Ogden]. <https://doi.org/10.1515/opli-2020-0173>.
- Boersma, Paul and David Weenink. 2020. Praat: Doing phonetics by computer [Computer program]. Version 6.1.37, [http://www.praat.org/\(16.12.2020\)](http://www.praat.org/(16.12.2020)).
- Bolden, Galina. 2004. "The quote and beyond: Defining boundaries of reported speech in conversational Russian." *Journal of Pragmatics* 36(6), 1071–118.
- Bressem, Jana and Silva Ladewig. 2011. "Rethinking gesture phases: Articulatory features of gestural movement?." *Semiotica* 2011(184), 53–91.
- Bressem, Jana and Cornelia Müller. 2014. "A repertoire of German recurrent gestures with pragmatic functions." In: *Body – Language – Communication*, eds. Müller, Cornelia, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill and Jana Bressem, p. 1575–91. Berlin: De Gruyter Mouton.
- Cantarutti, Marina N. 2020. *The multimodal and sequential design of co-animation as a practice for association in english interaction*. PhD thesis, University of York.
- Clark, Herbert H. 1996. *Using language*. Cambridge: Cambridge University Press.
- Clark, Herbert H. and Gerrig, Richard. 1990. "Quotations as demonstrations." *Language* 66(4), 764–805.
- Clark, Herbert H. 2016. "Depicting as a method of communication." *Psychological Review* 123(3), 324–47.
- Clift, Rebecca. 2016. *Conversation analysis*. Cambridge: Cambridge University Press.
- Couper-Kuhlen, Elisabeth. 2004. "Prosody and sequence organization in English conversation." In: *Sound patterns in interaction*, eds. Couper-Kuhlen, Elisabeth and Cecilia Ford, p. 355–76. Amsterdam: Benjamins.
- Couper-Kuhlen, Elisabeth and Dagmar Barth-Weingarten. 2011. "A system for transcribing talk-in-interaction: GAT 2. English translation and adaptation of Selting, Margret et al. (2009): Gesprächsanalytisches Transkriptionssystem 2." *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 12, 1–51.
- Couper-Kuhlen, Elisabeth and Margret Selting. 2018. *Interactional linguistics. Studying Language In Social Interaction*. Cambridge: Cambridge University Press.



- Couper-Kuhlen, Elizabeth. 1999. "Coherent voicing: On prosody in conversational reported speech." In: *Coherence in spoken and written discourse. how to create it and how to describe it*, eds. Bublitz, Wolfram, Uta Lenk and Eija Ventola, p. 11–33. Amsterdam: Benjamins.
- Dausendschön-Gay, Ulrich and Ulrich Krafft. 2009. "Preparing next actions in routine activities." *Discourse Processes* 46(2–3), 247–68.
- Deppermann, Arnulf and Susanne Günthner (eds.). 2015. *Temporality in interaction*. Philadelphia: Benjamins.
- Deppermann, Arnulf and Nadine Proske. 2015. "Grundeinheiten der Sprache und des Sprechens." In: *Handbuch satz, Äußerung, schema*, eds. Dürscheid, Christa and Jan Georg Schneider, p. 17–47. Berlin: de Gruyter.
- Deppermann, Arnulf and Jürgen Streeck (eds.). 2018. *Time in embodied interaction. Synchronicity and sequentiality of multi-modal resources*. Amsterdam: Benjamins.
- Drake, Veronika. 2015. "Indexing uncertainty: The case of turn-final *Or*." *Research on Language and Social Interaction* 48(3), 301–18.
- Dressel, Dennis and Ignacio Satti. 2021. "Embodied coparticipation practices in collaborative storytelling." *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 22, 54–86.
- Ehmer, Oliver. 2011. *Imagination und animation. Die herstellung mentaler Räume durch animierte Rede*. Berlin: de Gruyter.
- ELAN (Version 6.0) [Computer software]. 2020. *The language archive*. Nijmegen: Max Planck Institute for Psycholinguistics. Retrieved from <https://archive.mpi.nl/tla/elan>
- Golato, Andrea. 2000. "An innovative German quotative for reporting on embodied actions: *Und ich so/und er so* 'and i'm like/and he's like'." *Journal of Pragmatics* 32(1), 29–54.
- Good, Jeffrey S. 2015. "Reported and enacted actions: Moving beyond reported speech and related concepts." *Discourse Studies* 17(6), 663–81.
- Günthner, Susanne. 1999. "Polyphony and the 'layering of voices' in reported dialogues. An analysis of the use of prosodic devices in everyday reported speech." *Journal of Pragmatics* 31, 685–708.
- Günthner, Susanne. 2002. "Stimmenvielfalt im Diskurs: Formen der Stilisierung und Ästhetisierung in der Redewiedergabe." *Gesprächsforschung. Online-Zeitschrift zur verbalen Interaktion* 3, 59–80.
- Günthner, Susanne. 2011. "Between emergence and sedimentation: Projecting constructions in German interactions." In: *Constructions: Emerging and emergent*, eds. Auer, Peter and Stefan Pfänder, p. 156–85. Berlin: de Gruyter.
- Hayashi, Makoto. 2012. "Turn allocation and turn sharing." In: *The handbook of conversation analysis*, eds. Sidnell, Jack and Tanya Stivers, p. 167–90. Oxford, U.K.: Wiley-Blackwell.
- Heritage, John. 1984. "A change-of-state token and aspects of its sequential placement." In: *Structures of social action: Studies in conversation analysis*, eds. Atkinson, J. and Maxwell, John Heritage, p. 299–345. Cambridge, Cambridge University Press.
- Heritage, John. 2013. "Turn-initial position and some of its occupants." *Journal of Pragmatics* 57, 331–37.
- Holt, Elizabeth. 2000. "Reporting and reacting: Concurrent responses to reported speech." *Research on Language and Social Interaction* 33(4), 425–54.
- Holt, Elizabeth. 2007. "'I'm eyeing your chop up mind': Reporting and enacting." In: *Reporting talk: Reported speech in interaction*, eds. Holt, Elizabeth and Rebecca Clift, p. 47–80. Cambridge University Press.
- Keevallik, Leelo. 2010. "Bodily quoting in dance correction." *Research on Language and Social Interaction* 43(4), 401–26.
- Kendon, Adam. 1980. "Gesture and speech: Two aspects of the process of utterance." In: *Nonverbal Communication and Language*, ed. Key, Mary Ritchie, p. 207–88. The Hague: Mouton.
- Klewitz, Gabriele and Elizabeth Couper-Kuhlen. 1999. "Quote – Unquote? The role of prosody in the contextualization of reported speech sequences." *Pragmatics. International Pragmatics Association* 9(4), 459–85.
- Kaukomaa, Timo, Anssi Peräkylä, and Johanna Ruusuvuori. 2013. Turn-opening smiles: Facial expression constructing emotional transition in conversation. *Journal of Pragmatics* 55, 21–42.
- Kaukomaa, Timo, Anssi Peräkylä, and Johanna Ruusuvuori. 2014. Foreshadowing a problem: Turn-opening frowns in conversation. *Journal of Pragmatics* 71, 132–47.
- Lerner, Gene H. 1991. "On the syntax of sentences-in-progress." *Language in Society* 20, 441–58.
- Lerner, Gene H. 2002. "Turn-sharing. The coral co-production of talk in interaction." In: *The language of turn and sequence*, eds. Ford, Cecilia E., Barbara Fox, and Sandra Thompson. p. 225–56. Oxford: Oxford University Press.
- Mandel, Daniel. In preparation. *Jemandem eine Stimme geben – polyphone und polykinetische Äußerungen im Sprechen von Angesicht zu Angesicht*. Dissertation. Freiburg i.Br.
- Mondada, Lorenza. 2014. "Pointing, talk, and the bodies. Reference and joint attention as embodied interactional achievements." In: *From Gesture in Conversation to Visible Action as Utterance*, eds. Seyfeddinipur, Mandana and Marianne Gullberg, p. 95–124. Amsterdam: Benjamins.
- Mondada, Lorenza. 2019. "Contemporary issues in conversation analysis: Embodiment and materiality, multimodality and multisensoriality in social interaction." *Journal of Pragmatics* 145, 47–62.
- Mori, Junko and Makoto Hayashi. 2006. "The achievement of intersubjectivity through embodied completions." *Applied Linguistics* 27, 195–219.
- Niemelä, Maarit. 2011. *Resonance in storytelling: Verbal, prosodic and embodied practices of stance taking*. Oulu: University of Oulu.

- Park, Yujung. 2009. "Interaction between grammar and multimodal resources: Quoting different characters in Korean multi-party conversation." *Discourse Studies* 11(1), 79–104.
- Pfänder, Stefan and Elizabeth Couper-Kuhlen. 2019. "Turn-sharing revisited: An exploration of simultaneous speech in interactions between couples." *Journal of Pragmatics* 147, 22–48.
- Pfänder, Stefan, Federica D'Antoni, Luis Ignacio Satti and Elke Schumann. in preparation. Sofa-Talks.
- Raymond, Geoffrey. 2003. "Grammar and social organization: Yes/no interrogatives and the structure of responding." *American Sociological Review* 68(6), 939–67.
- Reber, Elisabeth and Elizabeth Couper-Kuhlen. 2010. "Interjektionen zwischen Lexikon und Vokalität: Lexem oder Lautobjekt?." In: *Sprache intermedial. Stimme und Schrift, Bild und Ton*, eds. Deppermann, Arnulf and Angelika Linke, p. 69–96. Berlin: de Gruyter.
- Runway (Version 1.9) [Computer software]. 2020. Retrieved from <https://celestialteapot.com/runway/>.
- Sacks, Harvey. 1974. "An analysis of the course of a joke's telling in conversation." In: *Explorations in the ethnography of speaking*, eds. Bauman, Richard and Joel Sherzer, p. 337–53. Cambridge: Cambridge University Press.
- Schegloff, Emanuel and Harvey, Sacks. 1973. "Opening up closings." *Semiotica* 8, 289–327.
- Schegloff, Emanuel. 1984. "On some gestures' relation to talk." In: *Structures of Social Action: Studies in Conversation Analysis*, eds. Atkinson, Maxwell J. and John Heritage, p. 266–95. Cambridge: Cambridge University Press.
- Selting, Margret and Elizabeth Couper-Kuhlen (eds.). 2001. *Studies in interactional linguistics*. Amsterdam, Benjamins.
- Selting, Margret, et al. 2009. "Gesprächsanalytisches Transkriptionssystem 2 (GAT2)." *Gesprächsforschung: Online-Zeitschrift zur verbalen Interaktion* 10, 353–402.
- Sidnell, Jack and Tanya Stivers (eds.). 2012. *The handbook of conversation analysis*. Malden: Wiley-Blackwell.
- Sidnell, Jack. 2006. "Coordinating gesture, talk, and gaze in reenactments." *Research on Language and Social Interaction* 39(4), 377–409.
- Soulaimani, Dris. 2018. "Talk, voice and gestures in reported speech: Toward an integrated approach." *Discourse Studies* 20(3), 361–76.
- Stivers, Tanya. 2008. "Stance, alignment and affiliation during storytelling: When nodding is a token of affiliation." *Research on Language and Social Interaction* 41, 31–57.
- Streeck, Jürgen. 2002. "Grammars, words, and embodied meanings: On the uses and evolution of so and like." *Journal of Communication* 52(3), 581–96.
- Streeck, Jürgen and Hartge, Ulrike. 1992. "Previews: Gestures at the transition place." In: *The Contextualization of language*, eds. Auer, Peter and Aldo Di Luzio, p. 135–58. Amsterdam: Benjamins.
- Streeck, Jürgen and Scott Jordan. 2009. "Projection and anticipation: The forward-looking nature of embodied communication." *Discourse Processes* 46(2–3), 93–102.
- Stukenbrock, Anja. 2012. "Zur Beredsamkeit des Körpers. Figurendarstellung und Figurenwissen als multimodale Alltagsinszenierung." In: *Figurenwissen. funktion von wissen bei der narrativen figurendarstellung*, eds. Jappe, Lilith, Olav Krämer and Lampert Fabian, p. 345–85. Berlin: De Gruyter.
- Stukenbrock, Anja. 2015. *Deixis in der face-to-face-Interaktion*, Berlin: de Gruyter.
- Stukenbrock, Anja. 2018. "Forward-looking: Where do we go with multimodal projections?." In: *Time in embodied interaction: Synchronicity and sequentiality of multimodal resources*, eds. Deppermann, Arnulf and Jürgen Streeck, p. 31–68. Amsterdam: Benjamins.
- Thompson, Sandra and Suzuki, Ryoko. 2014. "Reenactments in conversation: Gaze and reciprocity." *Discourse Studies* 16(6), 816–46.
- Volosinov, Valentin. 1929/1973. *Marxism and the philosophy of language*. Cambridge, MA: Harvard University Press.
- Walker, Gareth. 2012. "Phonetics and prosody in conversation." In: *The handbook of conversation analysis*, eds. Stivers, Tanya and Jack Sidnell, p. 455–74. Malden: Wiley-Blackwell.
- Wells, Bill and Sarah Macfarlane. 1998. "Prosody as an interactional resource: Turn-projection and overlap." *Language and Speech* 41(3–4), 265–94.
- Wilkinson, Ray, Suzanne Beeke and Jane Maxim. 2010. "Formulating actions and events with limited linguistic resources: Enactment and iconicity in agrammatic aphasic talk." *Research on Language and Social Interaction* 43(1), 57–84.