Social Interaction

Video-Based Studies of Human Sociality

_____

Co-animation and the Multimodal Management of Contextualisation Problems when Jointly 'Doing Being' Others[1]

_____

Marina N. Cantarutti

The Open University

Abstract

In everyday interaction, participants speak on their own behalf but may temporarily speak as or on behalf of a figure (i.e. past or fictional self, others or objects). This practice of 'animation' can be continued or extended by co-participants in responsive position, resulting in co-animation (Cantarutti, 2020) of the same figure. Animation relies on the successful ascription of roles, participation framework shifts and projected stances to either the here-and-now of interaction or the there-and-then of animated content. In turn, the recognition of a response as a co-animation requires the creation of similarity between animated contributions. Through a multimodal interactional linguistic analysis of 89 cases of co-animation, this paper discusses how participants jointly solve these interactional contextualisation 'problems' smoothly through multimodal gestalts of lexico-grammatical, prosodic and gestural detail.

Keywords: animation, co-animation, multimodality, participant problems, 'doing being'

_____

[1] This article is based on my doctoral research at the University of York.

1. Introducing co-animation

In everyday interaction, participants routinely speak on their own behalf (Lerner, 1996b). However, it is frequently the case that during complex activities such as tellings, participants may introduce the actions of others or past versions of themselves, whose words or behaviour get staged 'live' in the here-and-now of the current interactional space. Literary, discourse-pragmatic and interactional approaches have called (aspects of) this practice 'reported speech', 're-enactment', 'voicing', 'quotation', 'role shift' (among others, in spoken and signed languages; see Buchstaller, 2013, and Cantarutti, 2020, for a review). This practice will be here referred to as 'animation', adopting Goffman's (1981) idea of animators as 'sounding boxes'[2] for others, temporarily speaking or acting *as* or *on behalf of* a figure (i.e. past or fictional self, others or objects) in the development of an activity. In the here-and-now, animation is successfully recognised as such by co-participants, who orient to the action carried out *through* the animation rather than acting as addressees to the animated content per se.

An excerpt illustrating animation in ordinary English interaction appears below, with animated contributions in **bold**. Friends Jesse and Fiona are discussing ideas for Jesse's radio programme. Jesse announces he intends to repeat the structure of the prior year (lines 28-29), but now exercising the achievement of knowing people who can come in as presenters, unlike in the past, where recruitment was more challenging (line 35). After prefacing a contrast ('as opposed to…', line 38), Jesse animates himself doing a request for participation to absent parties ('please, come on my show', line 38). Jesse frames this as playful with interspersed laughter, and embodies it through an abrupt looking-up gesture and higher pitch and volume (line 38; Figs. 1, 3), thus offering a caricaturised past version of Self recruiting participants for the programme.

---

[2] The production model has been criticised as insufficient to cover the laminated reality of roles that animation in interaction brings about (see C. Goodwin, 2007; Levinson, 1988). Our work subscribes to this criticism but still adopts the Goffmanian label as it captures well the 'doing being' nature of the practice.

Excerpt 1.1: MCY04SANDWICH 'Desperate' (GAT-2 + Multimodal Transcription conventions; * delimits gestures by Jesse)

```
27 JES:[hh°]
28      its ˆBAsically gonna be whAt i ‾DID (0.6)
29      ˇLAST year, (.)
30      but ‾NOW that (0.9)
31      its `NOT jus:t; (0.4) ˇME:,=
32      =there ll be a `pOol of pre[ˇSEN  ]ters,
33 FIO:                          [`YE:AH.]
34 JES:<<l,cr> ˇHOPEfully,> (0.6)
35      ‾A:ND (.) ´nOw that we Actually ´KNOW some people? hh°
36      (0.4)
37 FIO:`YE:AH.
38 JES:as *oˇPPOSED to:;* (.)
   jes    *looks away   *eyes up
39      *<<f,h>`PLE::ASE = c(h°)ome *on my `SH(h°)O:   *W> [°hh ] ha
        *head & eyes up----,,,,,,,,,*head to home pos  *looks down
40 FIO:                                              [ʔah°] ʔah° hahaha
```

Figure 1. Jesse's (right) embodied self-animation (lines 38-39)
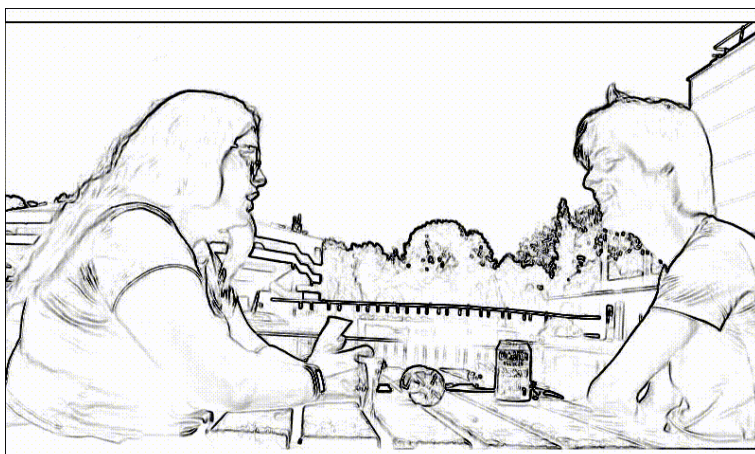
Jesse's version of Self begging for participation is recognisable as an animation as it is not treated by Fiona as addressed at her, and it is produced as an 'overdone' (Drew, 1987) form of behaviour that looks incongruent with the prosodic and gestural design configurations of the ongoing telling.

Our example also shows how animation can be an option for co-participants at a responsive slot. If we focus on the whole of Fiona's response (1.2 below), we can see that she demonstrates understanding (Sacks, 1992) of this being playfully-framed by first providing matching laughter to Jesse's self-mocked past version of Self (line 40), and then by continuing the animation of Jesse's figure, unpacking the 'begging' tone in a further mocking way (lines 42-43). Fiona produces this version with a modalised and velarised voice quality that resembles that of a child, while visibly pouting with the sides of her lips pressed (Figs. 2, 3). In other words, Fiona has engaged in **responsive co-animation** (Cantarutti, 2020), which

results in Fiona and Jesse jointly animating a playful version of a past begging Jesse. Fiona's animated response is acknowledged by Jesse with a smile (line 43).
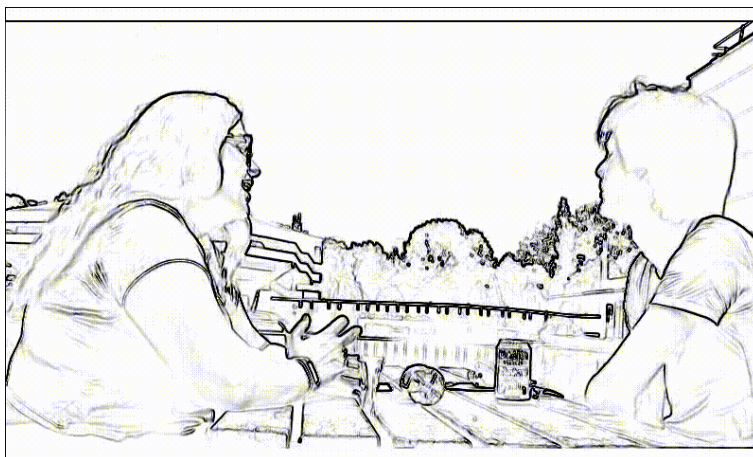
Excerpt 1.2: MCY04SANDWICH 'Desperate' (lines 38-43; *‡∞ delimit gestures by Jesse)

```
 39      *<<f,h>`PLE::ASE = c(h°)ome *on my `SH(h°)O:    *W> [°hh ] ha
         *head & eyes up----,,,,,,,,,*head to home pos  *looks down
 40 FIO:                                                 [ʔah°] ʔah° hahaha
 41 JES:*°hh hh° °hh *
    jes *looks away L*looks away R ->>
 42 FIO:<<mod, raised lrx, velarised, pouting> im `DESperate.(.)
 43      ∞↑`PLE:::ASE,> (0.5)
    jes ∞smiles ->>
 44      ↓`YE:AH.
 43      (0.9) + ‡ (0.8)
    jes         ‡nods(x3)
```

Figure 2. Fiona's co-animation of the begging Jesse figure (lines 42-43)

This paper will focus on both sides of the practice of **co-animation** – i.e. first animations like Jesse's, as well as responsive co-animations like Fiona's above, by which a participant continues or (re-)completes an animation by a prior speaker in an adjacent position (Cantarutti, 2020; cfr. Guardiola & Bertrand, 2013; Mathis & Yule, 1994; Niemelä, 2011).

Extract 1 shows how this is a successful practice, with no repair or sanctioning. In order for this to have proceeded so smoothly, participants ought to have made a number of things interpretable for each other, that is, they should have solved a number of interactional 'problems'. Fiona should have been able to recognise that Jesse was 'doing being' himself in a different place and time (line 39), and that his request to come to the show was not directly addressed to her, and at the same time, that there is a particular humorously self-deprecating stance being

projected onto his own animated request. Jesse, in turn, should recognise that Fiona is not professing any level of despair towards him in her response. Rather, she demonstrates understanding by continuing his playful self-mockery version while intensifying the playful tone.

Our work contributes to prior multimodal studies on the contextualisation of animation (e.g. Blackwell et al., 2015; Bolden, 2004; Buchstaller, 2013; Couper-Kuhlen, 1998; Fox & Robles, 2010; Günthner, 1997, 1999; Klewitz & Couper-Kuhlen, 1999; Lampert, 2018; Mandel & Ehmer, 2019; Niemelä, 2010; Reber, 2020; Sidnell, 2006; Stec et al., 2015; Tannen, 2007) by looking at what the previously understudied practice of co-animation tells us about the design features co-participants pick out as relevant for the creation of similarity between animated contributions. What is more, by problematising how co-animation weaves in together two interactional situations with role shifts (Stec et al., 2016) and changes in participation frameworks (Goffman, 1981) as well as stance displays that sway between the here-and-now of the ongoing course of action and the there-and-then of animated content, this paper addresses how participants jointly solve these interactional contextualisation and ascription 'problems' smoothly through multimodal gestalts (Mondada, 2018) of interacting lexico-grammatical, prosodic and gestural detail.

Therefore, this paper provides an overview of the types of interactional problems that co-participants manage when deploying a first animation, and those that emerge and that recipients orient to when responding through co-animation (section 3). Through a detailed multimodal interactional linguistic analysis[3] of 89 cases of co-animation, this paper offers a 'catalogue' of the most frequent multimodal gestalts bearing a role in managing these problems. Contextualisation problems are divided into: a) the marking of disjunction from the here-and-now and of coherence between animated contributions, and b) the overlaying of a stance towards the animated content in the here-and-now, in particular around two of the most frequent displays in the collection: (mild) indignation, and self-mockery. The next section will describe our data and methods.

2. Data and methods

This study is based on a sequential and multimodal analysis of 89 cases of co-animation; that is, of adjacent contributions where participants are 'doing being' the same figure as part of an ongoing social activity, in our collection those being troubles talk (n=38), teasing episodes (n=28) and joint fictionalisation (n=23). The cases were identified in 10 hours of video-recorded everyday interaction in English from the MCY (Cantarutti, 2018) and the MPI RCE (Rossi, 2011) corpora,

---

[3] In order to centre the analysis on multimodal design, the discussion of the social consequences of co-animation have been restricted in this paper. For a full discussion of co-animation as a practice and its social consequences in different social activities, see Cantarutti, 2020, in press.

both collected via direct recruitment (with ethics approval by the University of York) and with informed consent from participants to be recorded and for their anonymised data to be presented. This paper features a selection of representative analyses where the report of multimodal detail was foregrounded over the discussion of other aspects of social organisation that co-animation makes relevant, the latter having been developed in detail elsewhere (e.g. Cantarutti, in press).

Methodologically, the study aligns with multimodal approaches to Interactional Linguistics (Couper-Kuhlen & Selting, 2001; Mondada, 2018) and Conversation Analysis (Sacks et al., 1974). Transcriptions follow GAT-2 (Selting et al, 2011) and Mondada's (2018) Multimodal Transcription conventions. Verbal animations are marked in **bold**, and multimodal detail is transcribed in more granularity in the focal sections of the extracts. An overall parametric approach was adopted for the study of the different resources making up these co-animated multimodal gestalts, analysing the temporality and alignment of syntax, interactional-semantic processes (Deppermann, 2011), phonetic features, and gestural articulators and trajectories. Phonetic detail was recorded using an initial impressionistic parametric approach (Local & Walker, 2005) and validated through instrumental acoustic techniques employing Praat (Boersma & Weenink, 2016). Speakers' reference values for prosodic parameters were collected from a representative sample of speech (Walker, 2017), and normalisation for cross-comparison of speakers' parameters was done through logarithmic scales (Hz log and semitones), which stand closer to speakers' own perception (Nolan, 2003). The characterisation of voice quality was made via relativistic listening (Local & Kelly, 1989) to samples from Laver (1980). Rhythmic isochrony (where present) was treated as a perceptual gestalt (Auer et al., 1999) and measured following considerations in Ogden & Hawkins (2015). Gestural, postural and gaze detail in terms of trajectories, articulators and gestural phases (Kendon, 2004; McNeill, 1992) was annotated in ELAN (Brugman et al., 2004) with the video on mute, and then integrated with the observations in Praat to determine speech-gesture alignments.

3. Participant problems around the deployment of (co-)animation

We established earlier that in the here-and-now of interaction, animation temporarily foregrounds an event or behaviour transposed from a different place and time (real, hypothetical or fictional). Furthermore, we demonstrated that animation can also be an option in the response slot, culminating in the joint animation of what is interactionally treated by co-participants as the same figure, in other words, resulting in two 'doing being one' in adjacent positions.

The first animation is deployed for the needs of the ongoing activity, and in our collection, it occurs followed by transition-relevance places, which implies that it contributes to making a particular kind of response relevant immediately after its

deployment. Co-animated responses, in turn, orient to the design and sequential concerns of prior turns in particular ways and are subsequently treated as appropriate by first animators.

Therefore, part of the complexity of animation and co-animation lies in this *simultaneous management* of the interactional spaces of the *here-and-now* and the animated *there-and-then*, their participation frameworks, and their projected stances. Because of positional differences and the existence of these concurrent planes of interaction, *first* and *responsive animators* face different kinds of interactional exigencies in making their actions intelligible to each other.
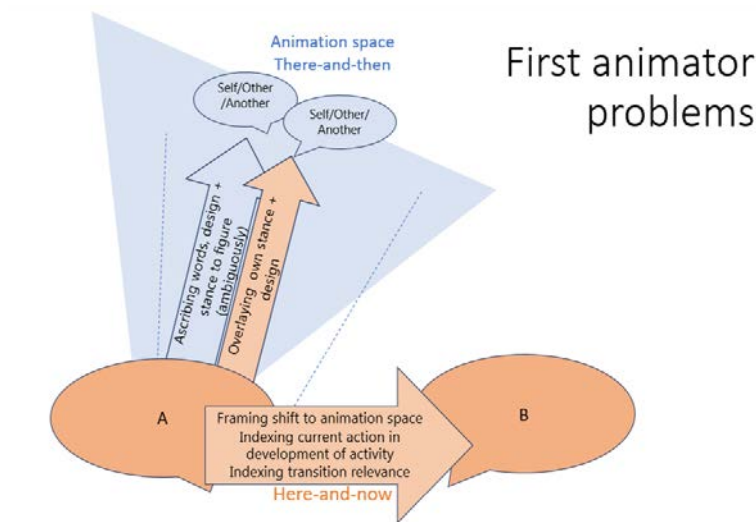
3.1 Participant problems for first animators

One of the challenges for first animators lies in syntagmatically marking the shift from the here-and-now into the parallel there-and-then of the animation space to make recognisable the shift from 'doing being' oneself with the current co-participant, to 'doing being' Self or Other in a different time or space; a figure who is ascribed a particular action and stance. In other words, first animators should contextualise this *disjunction* from the current interactional spaces and participation frameworks especially towards the left boundary of the animated content.

Simultaneously, first animators deploy the animation to contribute to an ongoing course of action at a particular point in the activity. That is, animation is a resource in itself that provides, for example, evidence for a described state of affairs, or contributes to the assessment of a particular behaviour (e.g. Clift, 2006; Couper-Kuhlen, 2007; Niemelä, 2011) in, for example, a storytelling activity.

Therefore, as figure 4 shows, first animators need to frame their shifts into animation to their co-participants, also contextualising in some way that certain words and stances are ascribed to a figure in the there-and-then, while simultaneously contextualising the stance taken regarding the animated content and indexing the action that it makes relevant from their co-participants in the here-and-now, and when it is due.

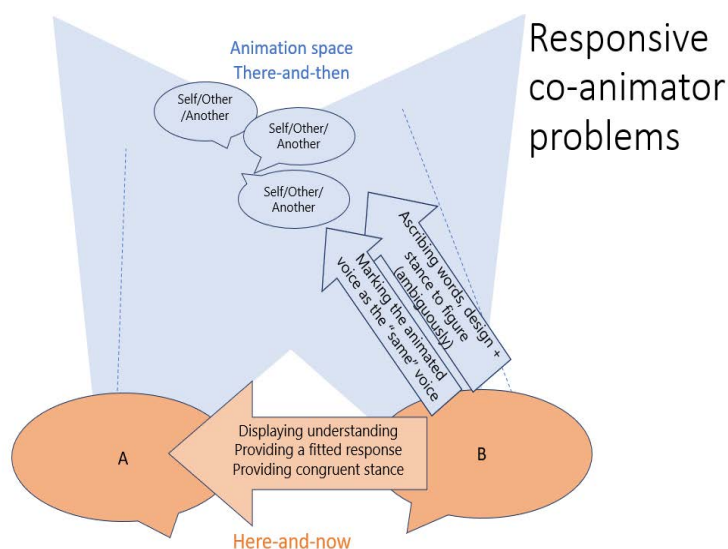**Figure 4. Participant problems for first animators.**

3.2 Participant problems for responsive co-animators

Once a next responsive action is projected as relevant, co-participants may respond to a first animation by engaging in co-animation. On these occasions, they need to make their contributions designedly hearable as continuations of an animation of the same figure by format-tying to prior animations. Therefore, co-animators should establish *coherence* and some form of *similarity* between their animated contributions and those of first animators.

Concurrently, in the here-and-now, co-participants' contributions are expected to be designed as congruent responses informed by the conditional relevance of prior turns and the stances thereby displayed, and in this sense, co-animated responses also need to build on prior cues as to what an aligning stance and action as a response would be. Responsive co-animations need to be hearable as supporting the ongoing course of action, although co-animation could also be employed to implement other independent projects. Figure 5 summarises the concurrent exigencies for co-animators in responsive position.

Figure 5: Participant problems for responsive co-animators.

Responsive co-animator problems

Animation space
There-and-then

Self/Other/Another

Self/Other/Another

Self/Other/Another

Ascribing words, design + stance to figure (ambiguously)

Marking the animated voice as the "same" voice

Displaying understanding
Providing a fitted response
Providing congruent stance

A

B

Here-and-now

These interactional problems are generally seamlessly managed, and as the next sections will reveal, this is done through the deployment of clusters of multimodal resources.

4. Co-animator problems and multimodal solutions: contextualising disjunction/coherence, and stance

This section will offer an analysis of two of the main concerns for co-animators regarding the relationship of animated material to the here-and-now, and of animated turns with each other: the (partial) marking of *disjunction* from the here-and-now and change of roles and participation frameworks (for first animators), and the marking of *coherence* with first animations (for responsive co-animators).

In our collection, animators exercise great variability in their forms of animation contextualisation, from lexico-grammatical markers such as quotative prefaces and tense and deictic displacement from the here-and-now into a past or imagined there-and-then, to more complex multimodal gestalts that involve particular deviations from modal voice and changes in gesture, gaze or posture. So whereas some resources may be more frequent than others, there is not necessarily an ever-present marker of animation in our collection with the exception of deictic displacement, which supports our concern with describing processes of disjunction and coherence as participant concerns rather predicting the use of specific multimodal resources.

In turn, there is not a specific resource deployed to do *co*-animation either, but rather, a set of processes of orientation, including matching, integration and transformation (Auer, 1996; Deppermann, 2011; M. H. Goodwin, 1990; Szczepek Reed, 2006) that format-tie to design features in prior animations. Therefore,

explaining how co-animation is recognisable as such relies on defining levels of *sameness and integration* between the two animations, going from productions that are remarkably multimodally similar across speakers, to those where coherence hinges only on lexico-grammar.

This variability across the collection as to what resources hearably/visibly 'do (co-) animation' depends on local contingencies, such as participants' F-formation (Kendon, 1990). Moreover, there is some temporal variability in terms of the alignment of different resources, differing in their onsets, offsets and overall duration relative to each other. This has been described at length in prior work (for an interactional discussion, see e.g. Bolden, 2004; Couper-Kuhlen, 1998; Klewitz & Couper-Kuhlen, 1999; Mandel & Ehmer, 2019). However, no prior work has discussed *co*-animated contributions in this way, which is why our contribution lies on determining the recurrent resources that are facilitative of co-animation, and on showing which of these co-participants themselves make relevant in their responsive animations.

4.1 Quotative prefaces, gaze and body torque

This section provides an overview of the way in which co-participants mark a change in participation frameworks and role shifts (for first animators) and the maintenance of this new configuration (for co-animators). Forms of contextualisation involve lexico-grammatical forms of displacement (e.g. pronouns and demonstratives anchored in a different place and time) and quotative prefaces, as well as body posture and gaze shifts.

Quotative prefaces (*say, think, be like* etc; see Buchstaller, 2013; Lampert, 2013, 2018; Robles, 2015 for a review), have been widely studied as the typical verbal means of projecting an animation. However, in this collection, quotative prefaces appear on 49% of first animations (72% of those being *be like*)*,* meaning the other 51% of cases bear *zero quotatives* (Mathis & Yule, 1994). Beyond any sociolinguistic trends, the prevalence of *be like* seems to be facilitative of co-animation, as this preface is said to create a sense of approximation to an experience rather than a claim to verbatimness (Fox & Robles, 2010; Romaine & Lange, 1991). Its deployment may then ease any initial epistemic differentials between the speakers, allowing co-participants to focus on the affiliative content of their response and not on its accuracy.

Responsive co-animators, on the other hand, have to build coherence based on at least one of the configurations in first animations. The fact that 98% of responsive co-animations in the collection lack a quotative preface[4] means that the first quotative in the first animation (whenever present) is retrospectively

---

[4] The cases where quotative prefaces do happen involve cases of disalignment around epistemic entitlement where co-participants provide a new version on record of the animation offered by the speaker of the prior turn.

treated as extending its domain to include the responsive co-animation. The normative lack of a new quotative preface shows how co-animators embed their responsive animations in a way that privileges some degree of contiguity between both animations[5].

Our next example illustrates animators' lack of concern with verbatimness, and the relative contiguity between co-animations. Friends Evie and Sam are discussing the Divine Right of Kings and how kings and subjects alike at the time took this as the truth. As part of her evaluative argument, Evie animates a generic figure of a king at the time upholding the dogma (lines 43-44, 46-48), and then animates the collective subdued response of the subjects with present-day forms of expression (lines 51-52, 'yeah, yeah, okay, sure'). After a short gap and without a new quotative preface, Sam co-animates the figure of the subjects with a reformulated and equally modern version of the same acquiescing action ('whatever you say', line 54), which Evie unproblematically ascribes as addressed to the animated figure of the King and not herself.

Excerpt 2. MCY20MUG "Whatever you say"

```
42 EVI:   like the ´KING thing? = is a gOod `SELL.
43        like (0.4) <<hyperartic> gOd has ˇCHOSEN me, =
44        =↓to ˇLEAD you, > (0.6)
45        ´THATS? = a hArd thing to `SELL.
46        but (.) <<hyperartic, h, mod> ʔI ʔAm a ‾GOD =
47        = wAlking on ˇEARTH, (0.3)
48        ˇW::ORship me?> haha
49        [haha   ]
50 SAM:   [nh° nh°]
51 EVI:   <<all> and then Everyone went>(0.5) ↑YE:AH.(0.2)↓`YE:AH.
52        <<pp> O`KAY = ↓`SURE. = ‾yEah ↓`YEAH.>
53        (0.5)
54 SAM:   <<p> whatEver you ^SAY.>
55        (0.2)
56 EVI:   like `I can sEe like the ˇKING (.) thing,
```

Strict contiguity between the two animations is not always the case, but any mediating elements tend to be more canonical forms of recipient appreciation, such as 'yeah' tokens and laughter. That is, participants orient to their recipient duties before (and/or after) co-animating. This was the case in our earlier 'Desperate' example (reproduced below) where Fiona first appreciates Jesse's playful self-animation through laughter (line 40) and then co-animates the figure, maintaining the pronominal reference of 'I' anchored in the figure of Past Jesse, and then producing a post-positioned canonical recipient response ('yeah', line 43).

---

[5] This matches the configurations seen in other forms of collaborative completions (Lerner, 1996a, 2004) or Other-turn-continuations (Couper-Kuhlen & Ono, 2010; Sidnell, 2012).

Excerpt 3: MCY04SANDWICH 'Desperate'

```
38 JES: as oˇPPOSED to:; (.)
39      <<f,h>`PLE::ASE = c(h°)ome on my `SH(h°)O:W> [°hh ] ha
40 FIO:                                            [ʔah°] ʔah° hahaha
41 JES: °hh hh° °hh
42 FIO: <<mod, raised lrx, velarised> im `DESperate.(.)
43      ↑ˇ`PLE:::ASE,> (0.5)
44      ↓`YE:AH.
43      (1.7)
```

After verbal quotatives, when present, other frequently deployed markers of disjunction and displacement from the here-and-now comprise **gaze aversion**[6] - looking away to the side, looking up or down, or momentarily closing the eyes – and in some cases, body torque (Schegloff, 1998) away from the co-participant. The relationship between temporary gaze aversion and enactments has been widely described (e.g Sidnell, 2006; Stec et al., 2015; Thompson & Suzuki, 2014) as a soft delimitation of the boundaries between content presented as animated and the rest of the talk. Moreover, the fixation of gaze onto another point away from the here-and-now of the interaction is also indexical, in that the attention is placed on an alternative space away from the current interactional space, as if the scene or the absent parties addressed were located in that imagined or invoked space which is for both co-participants to 'see', thus inviting joint attention (Streeck, 1993; Stukenbrock, 2014). In particular, in our collection, gaze aversion and/or body torque index first that the co-participant is not the addressed recipient within that particular stretch of animated content and that the response relevant in the next slot is not a reply to the animated content per se, but to the action supported through the deployment of animation, which typically involves some sort of appreciation or assessment of the animated 'object'.

Temporally, these aversive moves of the gaze or body tend to precede the verbal material of the animation; that is, they do something akin to forward-gesturing (Deppermann & Streeck, 2018; Streeck, 2009). These configurations may extend and be held over the whole of the verbal material, or they may only happen as a head flick that flags the left bracket (Klewitz & Couper-Kuhlen, 1999) of an upcoming momentary change of addressivity.

Some of these configurations appear in the next example, where Ivan and Lila are talking about their areas of literary research. Ivan mispronounces the word 'niche' and is called out by Lila, which leads to shared laughter and to Ivan producing an imprecation (line 9). Ivan turns away from Lila to playfully animate

---

[6] It is acknowledged that in many cases the exact gaze direction is difficult to establish. The observations in this section only apply to the cases where shifts can be more clearly determined (50% of the collection).

a request to the absent researcher, using 'legalese', that this section be removed from the recording (line 11).

Excerpt 4: MCY11LAPTOP 'Niche' (delimitation of gestures: *◊Ivan, + Lila)

```
02 IVA: not `E:verybody is as ¯NICHE ((ni:tʃ))
03      a:s [ʔu:h]
05 LIL:     [↑¯N ]ICHE ((ni:tʃ))
06      nh°e hehe (0.89)
07 IVA: what `IS it.
08 LIL: ^NI:CHE. ((ni:ʃ))
```


#1   #2   #3   #4

```
09 IVA: ´NI#*CHE? ([ni:ʃ]) =´DA:Mmit, [°hh haha  ]*◊
    iva     *leans forward…………-----------,,,,,,,,*
                                           ◊looks away
    fig.  #1
10 LIL:                              [hahahahaha] ha◊
    iva                                      ◊looks right & back->
11 IVA: ◊#*°h <<all>`STR*#IKE thAt◊ *from the> <<p, cr>↓`REcord.>#°hh
    iva  *torque....--,,*        *leans forward…………………………………----...>
        ◊looks tw front,,,,,,,,,,,◊looks down, closes eyes--------------->
    fig. #2           #3                                    #4
12 LIL:  #hahaha*haha◊ °hh
    iva        ->*
              ->◊opens eyes
    fig. #5
```


#5   #6   #7   #8   #9   #10
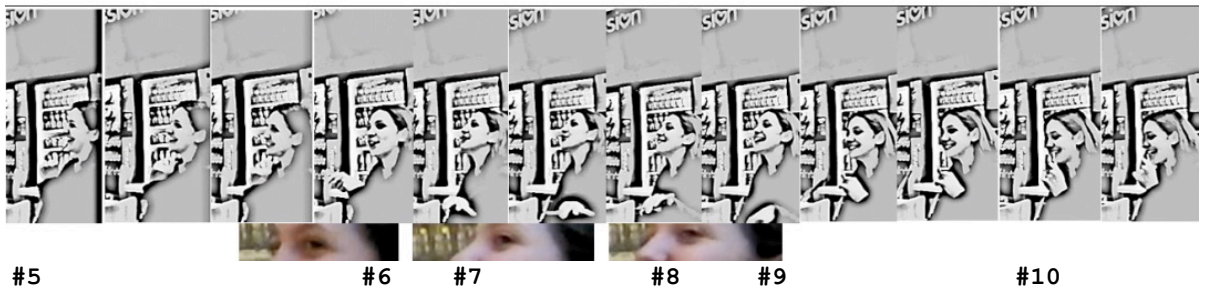
```
13      +°h<<l, cr> thats #`NOT #how i tAlk.>#
    lil +looks to L@cam-------------------->
    fig.                   #6    #7        #8
14 IVA: [hahaha                    ]
13 LIL: [<<l, cr>↓(e#+very      +`DAY.)>]= #+o´KAY?   +
    lil          >+closes eyes+ looks L   +looks down+
    fig.          #9                    #10
14      i +was just <cr> prEtty `TI:+RED.>
    lil    +looks up---------------+looks down at notebook->>
```

The torquing of Ivan's body and his subsequent return to home position with closed eyes are completed before the verbal content of the animation is finished, when Ivan opens his eyes again. After shared laughter and mutual gaze, Lila co-animates by producing two accounts on behalf of the Ivan figure ('that's not how I talk every day, okay?', lines 13, 15 and 'I was just pretty tired', line 16), also

initially directing her gaze at the same point in space. Lila then moves her head to the side, closing her eyes before finally looking down towards a notebook she has accidentally hit with her pen, keeping her gaze away from Ivan throughout her whole co-animated contribution.

Concurrently, gaze and body shifts bear a role in turn-management. In this sense, the right bracket of the animation content is also important in that these animations happen just before speaker transition, which means that the early fading away of features before the end of the animation (e.g. Bolden, 2004) is also functionally relevant. In our collection, as the animation is projectably complete or is being completed, a returned gaze towards a co-participant activates the imminent conditional relevance of a responsive action and indexes the upcoming transition space (as noted in e.g. Streeck, 2014) for co-participant involvement in the here-and-now, restoring the current participation frameworks. This is also the case for co-animators who, after a moment of mutual gaze at turn transition, often avert their gaze away from their participants until they approximate the point where the appreciation of their co-animated response is a relevant next.

This is illustrated in our next example (see full analysis in section 4.2), where a mother (Cassie) and her daughter (Leonie) complain about Cassie's partner. Cassie displays indignation about her partner's constant moaning and animates a telephone conversation she had with him earlier. The animation of the partner ('oh, I'm ill', line 36) is prefaced with a click and an iconic telephone gesture held throughout (for a discussion of iconic gestures, see section 4.2). This gesture is released as Cassie moves on to self-animate, voicing an annoyed response at the absent partner ('I spoke to you ten minutes ago, you were ill, yes, I get it', lines 37-39), and returns her gaze towards Leonie before completing the verbal material of this self-animation.

Excerpt 5.1: MCY09PIN 'Ill' (gesture delimitation: +≠≈Cassie, left ; *Leonie, right)

```
33 CAS: i dont ↑`KNOW. it s just ¯LIKE (0.2) ((sniff))
   Cas >> looking away from Leo and front ----------->
34      <<downstepped> you `jUst dont Always go ˆON about it;> (0.3)
```



```
#1              #2              #3  #4          #5      #6
```

```
35 LEO: nO you +ˇDO::[NT,       ]
   cas         +...RH up Y-handshape->
36 CAS:         ≈+[((click))] #2oh im#`ILL; (0.3)+
   cas          >+..............................-----------,,,,,,,,,+
                ≈looks down & left--------------->
   fig.                       #1      #2
37      <<all> i#≈spOke to you `TEN minutes ago;> =
   cas          ≈looks front and down---------------->
   fig.        #3
38      = you #≈were `ILL; =<<artic tension>#`YES:.
                ≈looks left & down-------------------->
             #4                          #5
39      i #*≈`G[ET it2.>]
   cas    ->≈looks @ Leo
   leo      *looks down R--->
   fig.    #6
```

Leonie, who was looking at Cassie at that moment, then withdraws her gaze and engages in a co-animated complaint ('you told me you were ill yesterday, you told me you were ill the day before', lines 41-42), returning her gaze towards Cassie at the end of her contribution, which is not reciprocated at this point. In fact, Cassie continues the animation (line 44, 46) and gaze will not be mutual until the end of the extended co-animated segment (line 46).

Excerpt 5.2: MCY09PIN "Ill" (gesture delimitation: +≠≈Cassie, left ; *Leonie, right)

| #7 | #8 | #9 | #10 |

```
40 LEO:    [ ̄I `KNO::]W,=
41       =<<all> you #`tOld me you were *Ill> `YESterda:y;
   leo                              ->*looks down front-->
   fig.              #7
42       ↓<<all, p> you `tOld me you were ill the day> be#*  ̄FO::RE-
   leo                                            ->*looks down right->
                                                    #8
43       #huhuh°#*(0.3)((sniff))
   leo        ->*looks at Cas
   fig. #9      #10
```


| #11 | #12 | #13 | #14 | #15 | #16 |

```
44 CAS: #*(we re at) the        *`CLInic in `YO::#+RK;# (0.3) +
   leo   *looks down R to front*
   cas                                         +looks @Leo  +
   fig. #11                                    #12  #13
45       °huhuhu
46       #+(g(h°)oing h° round a) `ROUNDab#+out `TWI:#*:CE.          *h°
   cas  +looks down                  +moves head L to look at Leo->>
   leo                                       *looks R & up @Cas*nodsx2
   fig.#14                           #15          #16
47 LEO: bu?°hh he dIdnt give a `SHIT
48       <<acc>↓when he was on the ´PHONE,> (0.4)
```

The most frequent configurations of gaze aversion during co-animation in our collection are schematised below (Fig. 6). Four points can be made around the anticipatory and contextualising work of gaze and body shifts: a) first animators tend to avert gaze or torque their bodies away from their co-participants slightly before the lexico-grammatical material of the animation starts; b) first animators tend to return to mutual gaze slightly before the verbal material of the animation is completed; c) a moment of mutual gaze tends to happen during turn transition; d) co-animators offer the same withdrawal + return pattern.

**Figure 6. Alignment of gaze shifts around animation.**
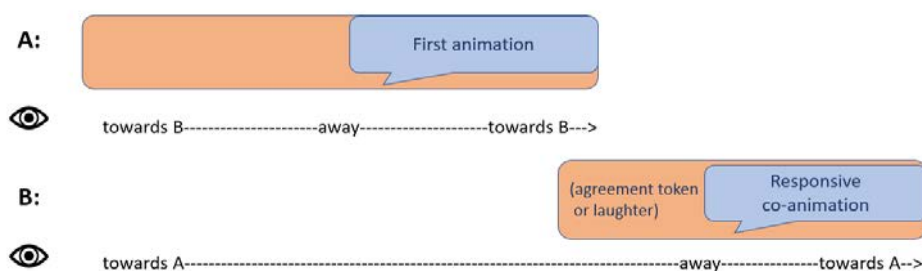
In this section, we observed the creation of contiguity of co-animated contributions through unprefaced responsive co-animations. We also discussed the role of gaze aversion and body torque for both first and co-animators in concurrently indexing changes in participation frameworks, turning the animation into an object of joint attention in the here-and-now, and for turn-management.

4.2 Prosodic and gestural shifts and orientation processes

The hearability and/or visibility of some disjunction away from a speaker's modal configuration in voice, posture or gesture that is sustained for enough time to be perceived is another functional way of marking an embodied action as not your own, or at least not your own in the here-and-now. As prior research has also found, these gestural and prosodic shifts vary in terms of when they are deployed relative to the lexico-grammatical content of the animation (where present), with some having their onset early on and others more closely aligned with the beginning and whole duration of the animated lexico-grammatical material (e.g. Couper-Kuhlen, 1998; Klewitz & Couper-Kuhlen, 1999; Mandel & Ehmer, 2019), with a tendency to wane away before the lexico-grammatical content of the (co-) animation is complete. This section will supplement those studies by focusing on co-animation, and providing evidence of what design configurations are taken up by co-animators, and in what way.

Figure 7 anticipates our findings in relation to the alignment of prosodic and/or gestural resources in animated material. Upward arrows show points of possible onsets, upward-downward arrows show only quickly fleeting changes, whereas downward arrows show offsets and fading away of these contextualising features.

Figure 7. Alignment of gestural and prosodic configurations for co-animation.
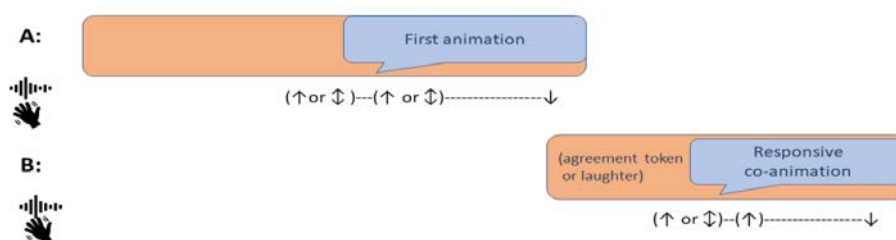
Prosodic shifts in first animations may be of different kinds: changes from modal voice into other articulatory settings; use of a noticeably wider, lower, or higher pitch or loudness span; deployment of stylised contours (Ladd, 1978); temporary turns to perceptually isochronous rhythmic patterns (Günthner, 1997; Müller, 1991). The examples below will show how some of these shifts interact with other multimodal resources and how they are taken up in co-animation, while acknowledging these configurations also have an important role to play in the marking of stance.

In turn, responsive co-animators in our collection frequently orient prosodically (Szczepek-Reed, 2006) to these shifts by engaging in forms of prosodic integration (Auer, 1996), by which they adjust their pitch configurations to offer prosodic (re)completions or continuations of the configurations in first animations. That is, in ways resembling self-increments (Walker, 2004), participants may start their co-animations at a comparable or lower pitch height in absolute terms (i.e. in comparison to the pitch height at the end of first animation) or in relative terms (that is, low with respect to the co-animator's own range). Apart from prosodically re-completing the first animations into closure by continuing the line of declination towards lower pitch and normally creakier points in their ranges, prosodic integration can be done by coming in on beat and continuing any existing rhythmic patterns (Ogden & Hawkins, 2015). Alternatively, or in addition, co-animators may prosodically match (Couper-Kuhlen, 1996; Gorisch, Wells, & Brown, 2012; Szczepek-Reed, 2006) and upgrade (Ogden, 2006) features of prior animations, thus re-doing or enhancing features of the prosodic design of first animations by e.g. repeating and/or expanding the pitch contours used, engaging in similar tempo changes or adopting similar or more exaggerated articulatory settings as those deployed in the first animation.

Co-animated contributions may vary in their degrees of prosodic similarity. Our next example shows a very close matching of prosodic parameters and gestural activity. Friends Laura and Becky are discussing bird droppings in big quantities in the city and at this point, Becky remembers the topic featured in the latest episode of 'Planet Earth'. Laura displays recognition (line 6) and launches a telling that involves overlaying onto a real event (her watching Planet Earth with her friends) a counterfactual scenario and version of Self. She proceeds to animate something she felt tempted to and presumably did not do: boast of

having seen first-hand the huge amount of bird droppings in Rome featured in the show ('oh, I was there and I saw it' lines 14-15). This is presented and oriented to as a laughable (lines 16-19), as a form of mockery of her attempted self-aggrandisement in front of her friends, which will eventually become a joint co-animated fictionalisation, with Becky expanding this first animation playfully (lines 20, 22).

Excerpt 6.1 MCY03SPIDER 'Planet Earth' (lines 1-15; Gesture delimitation: *†‡ Laura)


```
01 BEC: <<all> all ´RIGHT? = cause ↑i think
02      ↑`THAT was in like the lAst plAnet `EARTH. (0.6)
03      (when they) were ↓`jOking about like? the amOunt of `SH:IT. =
04 LAU: `YE::↓¯AH =`YEAH.
05 BEC: =that was like << f> `DR(h°)O[PPED; in one ˆDA:(?h°):(?h°):Y. >]
06 LAU:                       [<<f>↑↑YE:(?h°)E:(h°)E:S:>        ]
07      it was ˆRO::ME.
08      =↓as [ˆWE:LL.] ?h?h>
09 BEC:      [´WAS it?]
10 LAU: <<all>*yeah i was> ˆWATCHing i?;=
   lau        *looks away
11      =and i was +trYing really *`HARD;=
   bec             +looks @Lau--------------->>
   lau                           *looks down->
12      =<<all, l> cause i was just with *`FRIENDS;= *
   lau                           ->*closes eyes*looks up->

13      = <<all,l>‡↓`NOT *to (just `gO;)>;
   lau               ->*opens eyes------>
                   ‡straightens up -------->
14 LAU: <<mod, low lrx, rounded lips> †*Oh †↑I was      †ˆTHERE;=
   lau                                 *looks up
                                     †straightens neck
                                       †swings head&shoulders x3->
                                     (†1 twdsR      †2 twdsR->
15      =and i [†ˇSAW †it,*]>
   lau                 *looks @Bec
                ->†
            (†3 twds L)
```


Laura's animation is foregrounded by gaze aversion and a change in posture that involves straightening up and lowering her chin before the production of the quotative preface (Fig. 8). During the self-animation, an initial step-up in pitch and a sustained change in voice quality (a low-larynx, slightly faucalised quality; Fig. 10) with rounded lips is produced. Even though she is self-animating, her voice quality is markedly different from her own modal voice, something that is common to forms of self-mockery in the collection. Gesturally, she sways her shoulders and head slightly up and down and left and right.

Figure 8. Laura's self-animation (lines 11-14)

After shared laughter, Becky co-animates the figure put forward by Laura (lines 16-19) matching the low-larynx voice quality[7] and expanding the animation of Laura's self-aggrandised figure by first extending the temporal domain of the verb (from 'saw', line 15, to 'I've seen', line 20) and then unpacking Laura's 'it' into an extreme case formulation (Pomerantz, 1986): 'one of the great nature spectacles on Planet Earth', followed by an anticlimactic definition of what it really was ('shitting starlings', line 21), formulated as an increment. What is more, Becky matches some postural elements in Laura's animation, straightening her neck, and re-cycling Laura's left-to-right movement first with her head, looking in different directions, and then turning to swinging movements, adding a bouncing movement of her upper body on the punchline of her co-animation (Fig. 9).

---

[7]It was suggested during a data session that this voice quality is reminiscent of the Planet Earth programme's host, Sir David Attenborough.

Excerpt 6.2 MCY03SPIDER "Planet Earth" (lines 16-26; gesture delimitation: +·◊ Becky)

```
16 BEC:          [?huh hi ↑hi]
17 LAU  ?huh°?huh°?huh°?huh°?huh° °hh
18 BEC  °s::
19 LAU: ?i°hi?hi ?hi ?hi
20 BEC: <<mod,low lrx>+ive sEen [•One of    ]the ◊`great +nAture =
   bec             ->+opens mouth, smiles
                              •leans twds R    ◊sits straighter-->
                                                   +looks L->
21 LAU:                         [?h°i?h°i°hh]
22 BEC: =+´SPECtacles?= on [+•plAnet •´EARTH?>]
        +looks R           +looks L
                            •swings •L-R(x2)
23 LAU:                   [?h°i ?h°I         ]
24      ´YE:?h°:A?Hh°
25 BEC  <<l, p, low lrx>◊•`shItting ◊´S•TARlin•gs;>+=
   bec                  ◊bounces(x2)◊bouncesx2     ◊
                            •swings L-R   •swings•L-R
26 LAU  ?h°eh ?h°eh ?h°eh °hh
```
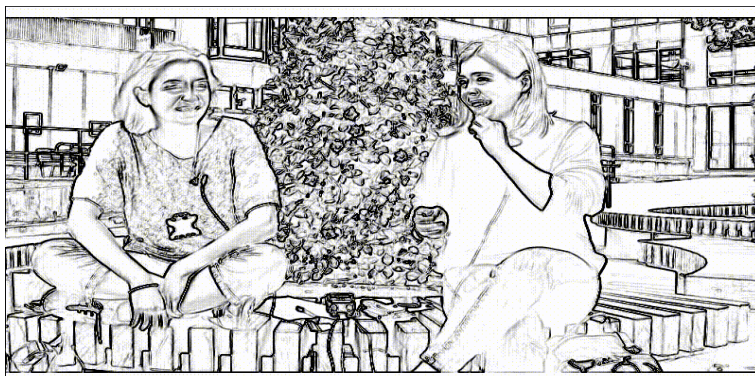
Figure 9. Becky's co-animation (lines 20, 22, 25)

This example is representative of our self-mockery collection where marked prosodic shifts contextualise the figure as a non-serious detached version from the real person (see also excerpts 1: 'Desperate' and 8: 'Bolo'), iconically marking a form of Self-Othering. Prosodic matching by co-participants is a way of indexing not only coherence but also a congruent stance in joint play.

Apart from prosodic matching, the speakers in our example also show mutual orientation to their gestural activity. **Gestural shifts** for first animators involve postural changes as well as the introduction of manual gestures (Kendon, 2004), in particular, iconic gestures (representing characteristics of their referents), beat gestures (quick gestures aligned rhythmically with speech), and metaphoric gestures (turning abstract concepts into concrete objects represented gesturally). For co-animators, the previously-described orientation processes of matching continuing, or upgrading for prosodic design are also relevant in terms of gestural

configurations (see e.g. Hayashi, 2005; Yasui, 2013, for gestural repetition and gestural complementarity across speakers), especially for metaphorical and beat gestures, and not so much for iconic gestures, which – if introduced by co-animators – are not tied to those in first animations. The following examples will demonstrate further different combinations of prosodic and gestural configurations in the co-animations in the collection and the role they hold in creating disjunction and coherence.

Iconic gestures visually represent features of their referents, and they tend to index role shifts (Stec et al., 2016) from speaker to figure, or from one animated figure to the next, and may be used to represent animated behaviour from either character or observer viewpoints (McNeill, 1992). These gestures may be fleeting or sustained throughout the animation. Our next example shows the role of gesture in shifts of animated figures, and with it, the concurrent overlaying of stance.

Cassie and her daughter Leonie are engaged in a joint complaint about Cassie's partner's constant moaning. As shown earlier, Cassie introduces her partner's voice with a Y-shaped telephone gesture, which is held until her role shift (Fig. 11). Her simultaneous frowning expression is continued throughout her next animation, where she enacts herself offering an annoyed response to her absent partner. Part of this animated reproachful response is also produced with an acceleration of tempo (line 37) that leads to the crux of her indignation display with tense articulation ('yes, I get it'). This display of indignation matches others in the collection in that they also involve brief gestalts of rhythmic scansion (i.e. the accumulation of strong accents separated in quasi-regular intervals of time; see Müller, 1991 and Uhmann, 1992) with beat gestures as 'batons' (Loehr, 2007). In Cassie's case, this is produced as three closely accented syllables ('ill', 'yes' and 'get') and two palm-up beat gestures (Fig. 12) aligned to the last two accented syllables, which are also uttered with greater articulatory tension and volume.

Excerpt 7.1 MCY09PIN 'Ill' (lines 35-39; gesture delimitation: +≠ΔCassie; ◊* Leo)

```
33 CAS: i dont ↑`KNOW. it s just ¯LIKE (0.2) ((sniff))
   cas  ->> looking away from Leo and tw front ------>
34      <<downstepped> you `jUst dont Always go ^ON about it;> (0.3)
35 LEO: nO you +ˇDO::[NT,     ]
   cas          +...RH up Y-handshape->
36 CAS:           ≈+[((click))] ?oh ≠im `ILL; (0.3)+
   cas           >+…………………………….-----------,,,,,,,,,+
                               ≠frowns-------->
          ≈looks down & left--------------->
37      <<all> i≈spOke to you `TEN minutes a+go;> =
   cas                                    +lowers sides of lips->
          ≈looks front and down---------------->
38      = you ≈were Δ`ILL; =<<artic tension>≠`YES:.Δ    ≠
                                    ->≠eyebrows up≠
          ≈looks left & down-------------------->
                Δhands separating…………………---,,,,Δrest->
39      Δi *≈`G[ΔET it2.>+  Δ]
   cas  ---------------+>
          ->≈looks @ Leo
      ->Δhands  Δup and downΔ
   leo     *looks down R--->
```

Figure 11. Cassie's embodied contextualisation of animated dialogue.
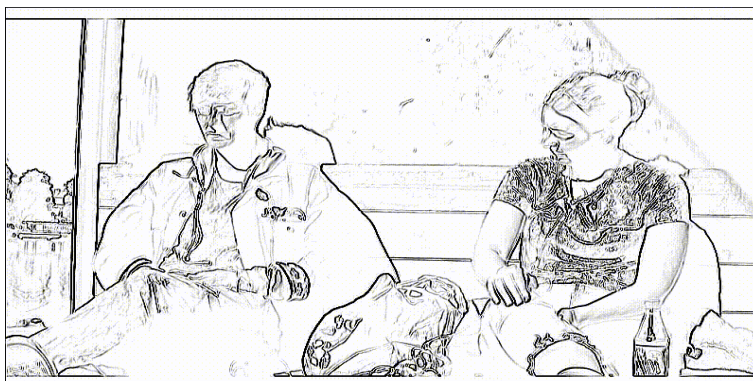
Leonie's co-animation comes in on beat in overlap (line 40), with a functionally ambiguous 'I know' that could be synonymous with Cassie's 'I get it', or mean 'tell me about it', indexing her own access to the complainable. Although Leonie's co-animation does not take up any of Cassie's gestural elements except for frowning (Fig. 13), her production is marked as coherent in terms of both design and stance. Firstly, this is done through the lexico-grammatical choices of pronouns 'I' and 'you', matching the deictic anchoring of Cassie's version and in ways that semantically expand the temporal domain of Cassie's animated complaint: 'spoke to you' (line 37) > 'told me', (41-42), 'ten minutes ago' (37) > 'yesterday', 'the day before' (41-42). Moreover, Leonie's 'you told me you were ill' (lines 41-42) matches Cassie's overall increase in tempo (10.34 syllables per second, an 85% increase to the rate in her previous turn), tying back to Cassie's fast-paced animation of Self in 'I spoke to you ten minutes ago, you were ill', (9.82 syll/sec,

a 93% increase from her prior turn). Beyond this matching of parameters, Leonie introduces another independent design element that contextualises the habituality and predictability of the behaviour complained about: the use of stylised contours (pitch contours made of sequences of levels, sounding like sung notes), which are said to mark things as routine and as 'no news' (Ladd, 1978; Ogden et. al, 2004; Szczepek-Reed, 2006). In this case, Leonie's own design elements contribute to the expression of a shared stance of annoyance, with her stylised contours distributed in a narrow pitch span, and final vowel lengthening on the temporal items, as visible in Fig. 14.
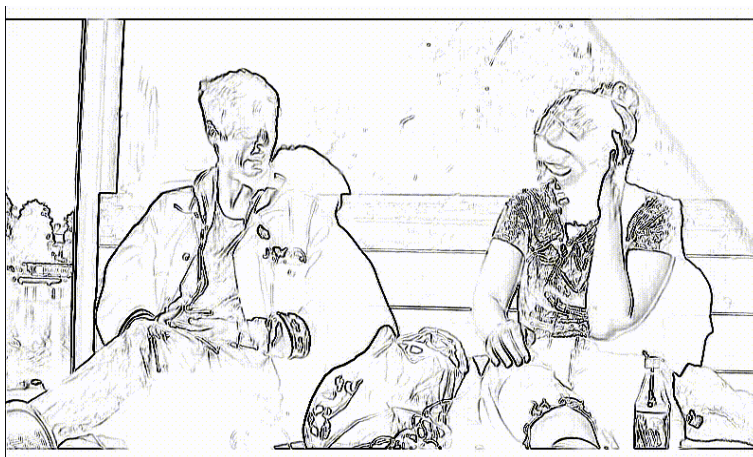
Excerpt 7.2: MCY09PIN 'Ill' (lines 40-43; gesture delimitation: +≠≈Cassie, left ; *Leonie, right)

```
40 LEO:    [¯I ˇKNO::]W,=
41       =<<all> you `tOld me you were *Ill> `YESterda:y;
   leo                              ->*looks down front-->
42       ↓<<all, p> you ◊`tOld me you were ill the day> be*¯FO:◊:RE-
   leo                                              ->*looks down right->
                    ◊frowns-------------------------------◊
43       huhuh°*(0.3)((sniff))
   leo      ->*looks at Cas
```

Figure 13. Leonie's co-animated response (lines 40-42)

The prior example showed the role of different types of prosody and gestures in the marking of role shift as well as contextualising stance, demonstrating the complexity and richness of expression available to speakers through the simultaneity of different strands of vocal and visual activity. Our next example will show how participants can not only foreground animation through prosody and gesture, but also engage in demonstration (Clark & Gerrig, 1990) through gestural animation while being verbally engaged in a description. Charlotte and Liz have been discussing their fear of waterfowl. Charlotte launches a telling involving a familiar swan that once chased her (lines 1-4, 7), which is appreciated

through shared laughter (lines 8-11). This disclosure of her victim status turns to an account (line 12) as Charlotte moves on to a more granular description of the menace that swans pose (lines 12-14). This verbal description of the swan as a threat is simultaneously embodied with a non-vocal animation of the swan's wings, with Charlotte leaning her neck and head forward representing the bird's head and beak (fig. 15), before verbally animating the swan with a vocalisation (line 14), where vocal and non-vocal resources join forces to animate the bird. This first part of the animation is responded to by Liz with high-pitched and loud recognition and agreement tokens ('mm' and 'yeah', line 15-16). As described for prior cases, Charlotte's gaze is averted from Liz during the bodily animation and reciprocation is resumed at points of transition (Fig. 7).

Excerpt 8.1: RCE01CIGARETTE 'Steady' (gesture delimitation: +‡Charlotte)

```
01 CHA: they `SCA:RE me; = when they put their `WINGS up. (1.6)
02      you know that blAck `SWAN =
03      = that's Always over ˘THERE,
04      [with the] rEd: (.) ˘BEAK,
05 LIZ: [`Mhm;    ]
06      `YEAH; (0.8)
07 CHA: <<l, cr> ↓rEally `CHASED me ↓↓lAst `YEAR> [h°uh° ]
08 LIZ:                                            [nh°uh°]
09 CHA: ?huh° °hh
10 LIZ: ?hu?hu°
11      [°hha                                                    ]
12 CHA: [+↑thEyre +b (.) theyre +rEally  +↑↑˘BIG though?]
   cha   +looks down           +looks R +looks @Liz---->
13 CHA: <<all> and (‡thEn when) they +put> their +`WINGS up.
   cha                 ‡raises L&RA backwards (iconic: bird wings)->>
                                   >+looks down +looks @ Liz
14      they[re li+ke +ˉK]WGR::: ((qʷħ:::))                 +hh°
   cha          +looks away                             +looks @Liz
                   +extends neck forward and shakes head+
```

Figure 15. Charlotte's (left) animation of the swan (lines 13-14)

Charlotte next enacts her response to the swan's attack through a prosodically upgraded response cry, 'woah', produced with a lengthened vowel, high volume and very high in pitch (11 semitones above her estimated midline; see Fig. 17). Moreover, she embodies a step back in the imaginary space as if being threatened live (Fig. 16). She once again reactivates transition relevance by looking towards Liz, who co-animates the figure of the victim only verbally, by presenting a synonymous formulation, the interjection 'steady' (line 19), also with high volume and longer vowel duration (Fig. 17) and followed by laughter. She also momentarily averts gaze as she produces this co-animation.

Excerpt 8.2: RCE01CIGARETTE 'Steady' (lines 14-21; gesture delimitation: +‡ Charlotte, left; * Liz, right)

```
14      they[re li+ke +¯K]WGR::: ((qʷ ħ:::))                    +hh°
  cha         +looks away                                   +looks @Liz
                +extends neck forward and shakes head+
15 LIZ:    [↑↑ˆM::m.    ]
16      [↑↑ˆYE::AH                             ]
17 CHA: [<<all, l>*(+you know +`what/like)>]
  cha           +looks@Liz +lowers head->
  liz           >*looks @Cha-------------->
18      *+youre ‡like ‡↑ˆWO:::AH    +[h°‡ ]
  cha ->+looks down raises eyebrows+looks @Liz
             ‡raises L leg            ‡
             ‡steps back             ‡
  liz *looks away
19 LIZ:                              [u?  ]he*he *ˇSTE:ADY::? e?*heheh°*
  liz                                         *IF away from mouth  *
                                              *looks away---*@Cha

20      e?heheh°
21 CHA:↑ˇYE:A:H?
```

Figure 16. Charlotte's self-animation as victim (line 18) and Liz's (right) co-animation.

Our prior example illustrated the role of iconic gestures in first animations and the possibility of 'distribution of labour' between description and demonstration. Our final example will show how gestures also enable forms of 'distributed animation' across participants. Whereas co-animators may not take up the gestures of first animators, they can offer the gestural representations themselves, resulting in the distribution between participants of the vocal and the gestural parts of an enactment. In the following excerpt, Ivan and Lila are discussing possible pseudonyms they can take for the recording. Ivan suggests a pseudonym for Lila (line 2), proposing 'Bolo' and securing recognition of the meaning of the term (lines 4-8) to then categorise Lila as a bolo-tie-wearing hipster (lines 12-14):

Excerpt 9.1: MCY11LAPTOP 'Bolo' (lines 1-16; gesture delimitation: + Ivan)

```
01 LIL: [`WHY didnt i (even go)] for ↑`SCHMEVin.
02 IVA: [yOure ˇB:Olo,          ]
03 LIL: (0.5) °h <<p>> ¯BO:lo: > = [↑E:↓¯BO:la::        ]
04 IVA:                           [¯LIKE a b = ¯LIKE a b]
05      like a ´BO:lo tIe? (0.6) <<p>¯YEAH>
06 LIL: `Oh ´YEAH,
07 IVA: you `knOw what that ´IS,
08 LIL: ↑yEah ´YEAH,
09 IVA: hahahah (0.5)
10 LIL: `thAt could ´WORK? °hh (0.58)
11      i dont `kno:w; = ↓what ¯Other (0.45)
12 IVA: you [you +`DRESS+ kind        ] of like+ a +`HIPster      +you-
   iva          + lifts RH, IF pointing with circular movements(x2)+
                     (+circle 1 large          +   +circle2 small  +)
13 LIL:      [my ↑nAmes quite `HARD to:;]
14 IVA: (then/i` thInk) youd be `wEaring a ´BOLo tie
15      [hahahahahah]
16 LIL: [hihihihihi ]
```

Lila accepts the pseudonym and initiates a sequence of joint fictionalisation with a first animation (lines 17-19) that adopts the character created for her, embodied verbally through a velarised voice quality and non-verbally by touching the collar of her shirt (Fig. 18).

Excerpt 9.2: MCY11LAPTOP 'Bolo' (lines 17-21; gesture delimitation: * ∞ Lila)

```
17 LIL  *∞°hhh oh   *`YEAH; =<<all, br, velarised> i just Actually
   lil  *raises LH  *LH hits chest ----------------------------->
         ∞fixed gaze on IVA------------------------------------>
18       had to *take it>´OFF earlier,
   lil        >*LH Th&IF fiddle w shirt->
19       <<cr> becAuse it ´BRO∞KE?>*           ∞
   lil  ----------------------,,,,*
        --------------------∞looks down & up @Iva∞
20 IVA: hahahaha
21 LIL: °ah °h
```

Figure 18: Lila's (right) first self-animation as 'Bolo' (lines 17-19)

After a moment of mutual laughter (lines 20-21), Ivan continues the fantasy in response ('I thought you were protecting it', line 22) and animates the 'Bolo' figure in a particular scenario where she had to protect the bolo tie ('I've taken off my bolo tie, I have to treasure my bolo tie', lines 24-25). It is during this co-animated contribution by Ivan that Lila simultaneously holds her collar and fiddles with it, that is, resulting in Ivan 'doing being' Bolo verbally, while Lila does so gesturally (Fig. 19).

Excerpt 9.3: MCY11LAPTOP 'Bolo' (lines 22-27; gesture delimitation * ∞ Lila; + Ivan)

```
22 IVA: i thOught you (were) pro[ˆTECting] it;
23 LIL:                         [?a    ]
        *°hh
    lil *raises LH towards collar-------------->
24 IVA:  *(ive +tAken off) my +`BOLo tie;
    iva        +looks up once +looks at L->
    lil ->*L&RH touching her collar, raised shoulders->
25      *i (have to) +*`TREAsure +my ´BOlo tie hah+
    iva            >+looks up   +looks @ Lil     +
    lil>*looks @collar*holds collar with one hand->
26 LIL: *haha
    lil  *rearranges collar with T&IF->
27 IVA: hihi ?a hh°
```

Figure 19: Ivan's co-animation of Bolo with Lila's collar-holding gesture (lines 22-25)

Ivan adds another element to increase the absurdity, which ties back to the beginning of the episode and their discussion of pseudonyms ('also, my name is Bolo', line 28), with Lila still fiddling with her collar. Lila closes the extended sequence of joint play by upgrading Ivan's formulation, extending the domain of 'Bolo' to her whole self ('everything about me is Bolo', line 31, Fig. 20), and the gesture is then released.

Excerpt 9.4: MCY11LAPTOP 'Bolo' (lines 28-32; gesture delimitation * ∞ Lila; + Ivan)

```
28 IVA:+↑Al[so my]   +nAme is ´BOlo,
      +looks up once+
29 LIL:   [(ey) ]
30    hahahaha
31    *`Everything a[bout *`ME is      *(.)<<cr> ´BOlo>]*
     >*hands down on collar--------------------------*release
                      *closes eyes   *
32 IVA:             [hahah haha ha                 ]
```

Figure 20: Ivan and Lila's final co-animation of Bolo (lines 28, 31).
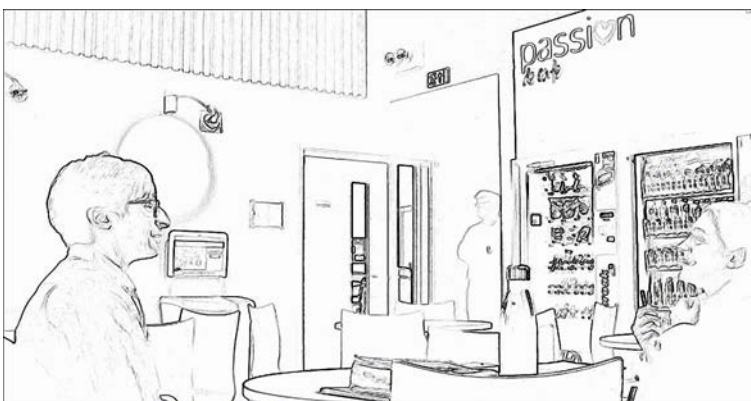
Lila's final contribution is both lexically and prosodically upgraded in that it starts at a higher pitch level relative to Ivan's prior production as well as higher in her own range, while still matching the intonational contour of Ivan's prior animation (Fig. 21). Lila's version of the name 'Bolo' is also prosodically upgraded, in that it is produced with a lengthened initial plosive – the closure period for [b] is longer – and she uses a wider pitch range, starting 'Bolo' in a low area within her range and with creakier voice.

In prosodic terms, our 'Bolo' example above displays forms of orientation through the co-participant's upgrading of some parameters (duration, pitch range) while matching others (e.g. the intonation contour). Moreover, the example demonstrated that co-animators may match some parameters for the sake of coherence display while also introducing their own gestural and prosodic renderings in a more independent way, as a form of co-shaping in their own terms what is now a joint interactional project.

4.3 Summary

Our central section has focused on animation as the entry into an alternative interactional space that involves a change in participation frameworks and roles. Quotative prefaces, gaze aversion and disjunctive prosodic and gestural-postural shifts make for noticeable embodied forms of foregrounding and contextualising a first animation and an imminent shift in figures, as well as upcoming transition-relevance places. Co-animators have a restricted set of resources to mark coherence with prior animations so that their contributions are hearable as (re-) completions or continuations of prior animations. They embed their contributions normally adjacently to first animations without a quotative preface, maintain the referents in their use of pronouns, tenses and demonstratives, and offer forms of prosodic integration, matching or even upgrading of these parameters to index coherence and different levels of similarity between their contributions and those of first animators. Gestural activity was not found to be matched as frequently as other parameters, but it was shown how the vocal and visual aspects of an animation can be distributed between participants.

Our excerpts also provided evidence that multimodal resources contribute to concurrently displaying a particular evaluative stance towards the animated content or figure, and it is these multimodal cues that provide to the co-participant a view as to what a matching (dis)affiliative response should involve, and what the animation is doing in the context of the wider activity in progress. Even though prosodic and gestural configurations are not the only ways in which stance is overlaid – as stance may be anticipated by e.g. prior assessments – they have an important role to play in stance-contextualisation.

It is not easy for the analyst to establish in each case which resources contribute solely to the contextualization of the figure, or to the overlaid stance in the here-and-now. Participants seem to successfully and gestaltically ascribe each

concurrent interactional function irrespective of any specific differentiation of resources, and as mentioned earlier, many of these resources simultaneously orient to these dual purposes. However, specific clusters of parameters in our collection have been found to appear repeatedly as multimodal gestalts during very specific social activities (namely, troubles-talk on the one hand, and self-mockery and teasing on the other), where particular affective and evaluative stances are relevant, and these were briefly previewed in this section and will be expanded on in further work.

5. Concluding remarks

This article set out to introduce the practice of co-animation and detail the clusters of multimodal resources that co-participants deploy to manage the contextualisation problems that co-animation entails. Our representative examples have shown an array of configurations that co-participants jointly construct in their co-creation of these animated figures in different activities.

Co-animation involves several levels of lamination where the here-and-now and the there-and-then are concurrently managed with multiple and complex temporalities and simultaneities (Mondada, 2018). Multimodal gestalts of varying lexico-grammatical, prosodic and gestural detail contribute to contextualising the content as animated, with the left and right boundaries fuzzily framed and blended into the here-and-now through gaze and body torque patterns. These configurations simultaneously contribute to the marking of both displacement/disjunction and coherence on the one hand, and stance displays on the other, with the syntagmatic fading of resources foregrounding imminent transition relevance and the activation of conditional relevance. In this respect, we have shown how instead of attempting to determine individual and specific resources that may be deployed when it comes to collaborative practices, it makes better sense to identify multimodal gestalts and processes of orientation.

First animators need to deal with the interactional demands of marking disjunction from the here-and-now in terms of shifts of roles and participation frameworks, as well as of contextualising 'why that now'. Figure 22 schematises our findings regarding the 'multimodal solutions' they deploy:

Figure 22. Multimodal solutions to first animator problems
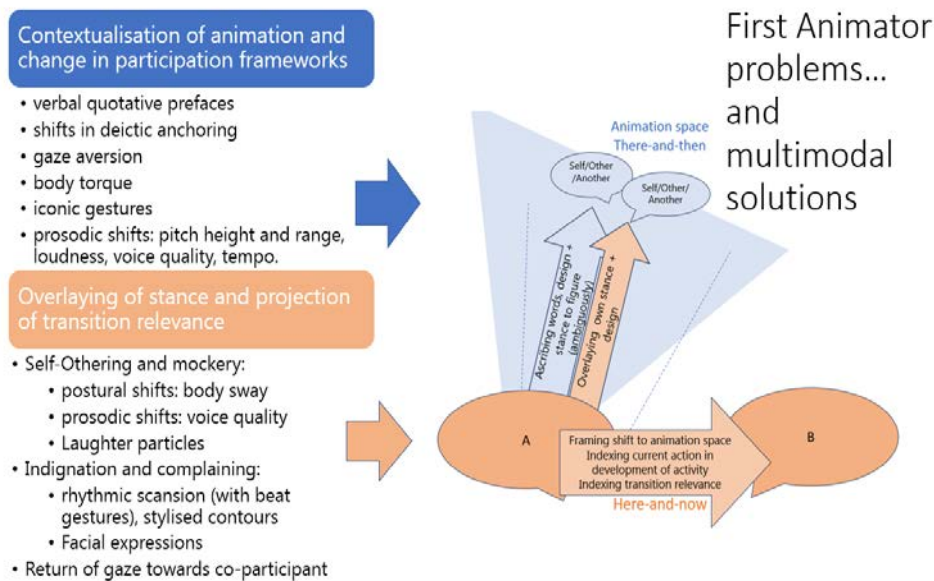
Co-animators, in turn, need to ensure some similarity with first animations, while producing a response hearable as coherent with the ongoing course of action. Figure 23 summarises their 'multimodal solutions':

Figure 23. Multimodal solutions to responsive co-animator problems
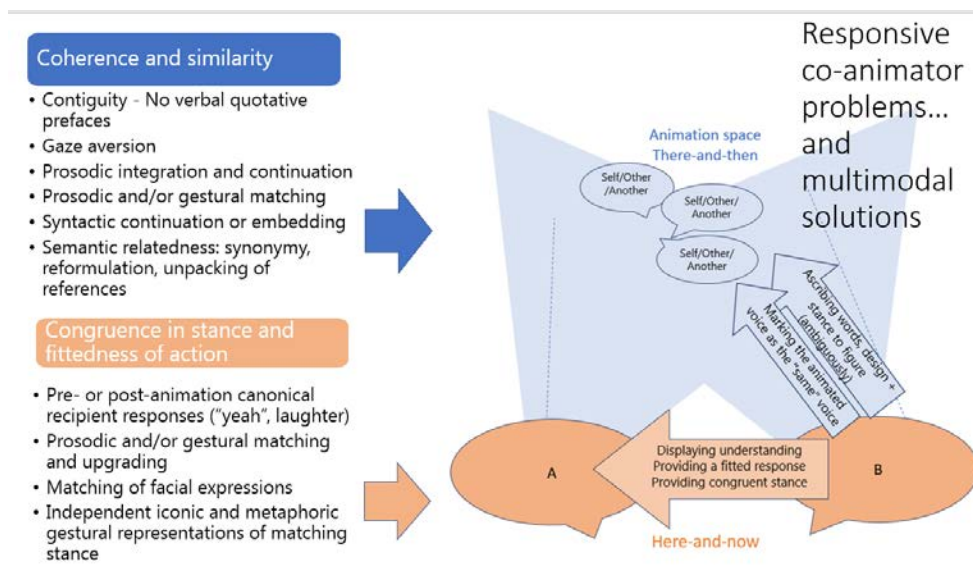
Further work will duly address what this paper, focused on design features, has selectively underplayed: the social-relational consequences of co-animation (for example, see Cantarutti, 2020, in press) and the role that these clusters of resources may bear in each situated activity where they are deployed.

Notwithstanding this, this article has demonstrated with a small sample how in such a complex and laminated practice as co-animation, co-participants successfully and smoothly manage concurrent planes of stance ascription and action multimodally when jointly and creatively 'doing being' others.

References

Auer, P. (1996). On the prosody and syntax of turn-continuations. *Prosody in Conversation*, 57–100.

Auer, P., Couper-Kuhlen, E., & Müller, F. (1999). *Language in Time: The Rhythm and Tempo of Spoken Interaction*. Oxford University Press.

Blackwell, N. L., Perlman, M., & Fox Tree, J. E. (2015). Quotation as a multimodal construction. *Journal of Pragmatics*, *81*(Supplement C), 1–7.

Boersma, P., & Weenink, D. (2016). *Praat: doing phonetics by computer [Computer program], Version 6.0. 14*.

Bolden, G. (2004). The quote and beyond: defining boundaries of reported speech in conversational Russian. *Journal of Pragmatics*, *36*(6), 1071–1118.

Brugman, H., Russel, A., (2004). Annotating Multi-media/Multi-modal Resources with ELAN. In: Proceedings of LREC 2004*,* Fourth International Conference on Language Resources and Evaluation.

Buchstaller, I. (2013). *Quotatives: New Trends and Sociolinguistic Implications*. John Wiley & Sons.

Cantarutti, M. N. (2018). *MCY Corpus of English Interaction* [Dataset]. University of York.

Cantarutti, M. N. (2020). *The Multimodal and Sequential Design of Co-Animation as a Practice for Association in English Interaction* [PhD, University of York]. http://etheses.whiterose.ac.uk/id/eprint/27344

Cantarutti, M.N. (in press). Co-animation in Troubles Talk. *Research on Language and Social Interaction*. Manuscript accepted.

Clark, H. H., & Gerrig, R. J. (1990). Quotations as Demonstrations. *Language*, *66*(4), 764–805.

Clift, R. (2006). Indexing stance: Reported speech as an interactional evidential1. *Journal of Sociolinguistics*, *10*(5), 569–595.

Couper-Kuhlen, E. (1996). The prosody of repetition: On quoting and mimicry. In E. Couper-Kuhlen & M. Selting (Eds.), *Prosody in conversation: Interactional studies* (pp. 366105). Cambridge University Press.

Couper-Kuhlen, E. (1998), Coherent Voicing. On Prosody in Conversational Reported Speech. *InLiSt - Interaction and Linguistic Structures*, vol. 1

Couper-Kuhlen, E. (2007). Assessing and accounting. In E. Holt & R. Clift (Eds.), *Reporting Talk: Reported Speech in Interaction* (Vol. 24, p. 81).

Couper-Kuhlen, E., & Ono, T. (2010). "Incrementing" in conversation. A comparison of practices in English, German, and Japanese. *Pragmatics*, 17(4), 513–552

Couper-Kuhlen, E., & Selting, M. (2001). Introducing interactional linguistics. *Studies in Interactional Linguistics, 122*.

Deppermann, A. (2011). The Study of Formulations as a Key to an Interactional Semantics. *Human Studies*, *34*(2), 115–128.

Deppermann, A., & Streeck, J. (2018). *Time in Embodied Interaction: Synchronicity and sequentiality of multimodal resources*. John Benjamins Publishing Company.

Drew, P. (1987). Po-faced receipts of teases. *Linguistics and Philosophy*, *25*(1).

Fox, B. A., & Robles, J. (2010). It's like mmm: Enactments with it's like. *Discourse Studies*, *12*(6), 715–738.

Goffman, E. (1981). *Forms of Talk*. University of Pennsylvania Press.

Goodwin, C. (2007). Interactive footing. In Holt, Elizabeth, & Clift, R. Reporting Talk: Reported Speech in Interaction. (pp.16-46). Cambridge University Press

Goodwin, M. H. (1990). *He-said-she-said: Talk as Social Organization Among Black Children*. Indiana University Press.

Gorisch, J., Wells, B., Brown, G. (2012). Pitch contour matching and interactional alignment across turns: An acoustic investigation. *Language and Speech* 55.1, S. 57-76

Guardiola, M., & Bertrand, R. (2013). Interactional convergence in conversational storytelling: when reported speech is a cue of alignment and/or affiliation. *Frontiers in Psychology*, *4*, 705.

Günthner, S. (1997). The contextualization of affect in reported dialogues. In S. Niemeier & R. Dirven (Eds.), *The Language of Emotions:*

Conceptualization, expression, and theoretical foundation. John Benjamins Publishing.

Günthner, S. (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*(5), 685–708.

Hayashi, M. (2005). Joint turn construction through language and the body: Notes on embodiment in coordinated participation in situated activities. *Semiotica*, *2005*(156). https://doi.org/10.1515/semi.2005.2005.156.21

Kendon, A. (1990). *Conducting Interaction: Patterns of Behavior in Focused Encounters*. CUP Archive.

Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge University Press.

Klewitz, G., & Couper-Kuhlen, E. (1999). Quote--unquote? The role of prosody in the contextualization of reported speech sequences. *Pragmatics. Quarterly Publication of the International Pragmatics Association (IPrA)*, *9*(4), 459–485.

Ladd, R. (1978). Stylized intonation. *Language*, *54*(3), 517–540.

Laver, J. (1980). *The Phonetic Description of Voice Quality*. Cambridge University Press.

Lampert, M. (2013). Say, be like, quote (unquote), and the air-quotes: interactive quotatives and their multimodal implications: The "new"quotatives remind us of the vocal, verbal, and gestural dimensions of speech. *English Today*, *29*(4), 45–56.

Lampert, M. (2018). *"Speaking" Quotation Marks: Toward a Multimodal Analysis of Quoting Verbatim in English*. Peter Lang.

Lerner, G. H. (1996a). On the" semi-permeable" character of grammatical units in conversation: Conditional entry into the turn space of another speaker. *Studies in Interactional Sociolinguistics*, *13*, 238–276.

Lerner, G. H. (1996b). Finding "face" in the preference structures of talk-in-interaction. *Social Psychology Quarterly*, *59*(4), 303.

Lerner, G. H. (2004). Collaborative turn sequences. In G. H. Lerner (Ed.), *Conversation Analysis: Studies from the First Generation* (pp. 225–256). John Benjamins Publishing Company.

Levinson, S. C. (1988). *Putting linguistics on a proper footing: Explorations in Goffman's participation framework*. Polity Press.

Local, J., & Kelly, J. (1989). *Doing Phonology: observing, recording, interpreting*. Manchester: Manchester University Press.

Local, J., & Walker, G. (2005). Methodological imperatives for investigating the phonetic organization and phonological structures of spontaneous speech. *Phonetica*, *62*(2-4), 120–130.

Loehr, D. (2007). Aspects of rhythm in gesture and speech. *Gesture*, *7*(2), 179–214.

Mandel, D., & Ehmer, O. (2019). *Fuzzy boundaries in quotations*. [Conference Presentation]. IIEMCA 2019: International Institute for Ethnomethodology and Conversation Analysis Conference, Mannheim.

Mathis, T., & Yule, G. (1994). Zero quotatives. *Discourse Processes*, *18*(1), 63–76.

McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. University of Chicago Press.

Mondada, L. (2018). Multiple Temporalities of Language and Body in Interaction: Challenges for Transcribing Multimodality. *Research on Language and Social Interaction*, *51*(1), 85–106.

Müller, F. E. (1991). *Metrical emphasis: Rhythmic scansions in Italian conversation*. Arbeitspapier des KONTRI-Projektes, Universität Konstanz, Konstanz (1991)

Niemelä, M. (2010). The reporting space in conversational storytelling: Orchestrating all semiotic channels for taking a stance. *Journal of Pragmatics*, *42*(12), 3258–3270.

Niemelä, M. (2011). *Resonance in storytelling: Verbal, prosodic and embodied practices of stance taking*. University of Oulu.

Nolan, F. (2003). *Intonational equivalence: an experimental evaluation of pitch scales*.

Ogden, R., Hakulinen, A., & Tainio, L. (2004). Indexing "no news" with stylization in Finnish. In *Typological Studies in Language* (pp. 299–334). John Benjamins Publishing Company.

Ogden, R., & Hawkins, S. (2015, August 10). Entrainment as a basis for co-ordinated actions in speech. *International Congress of the Phonetic Sciences*. International Congress of the Phonetic Sciences.

Pomerantz, A. (1986). Extreme Case Formulations: A Way of Legitimizing Claims. *Human Studies*, *9*(2/3), 219–229.

Reber, E. (2020). Visuo-material performances: "Literalized" quotations in prime minister's questions. *AILA Review*, *33*(1), 176–203.

Robles, J. S. (2015). Quotatives. In *The International Encyclopedia of Language and Social Interaction*. John Wiley & Sons, Inc.

Romaine, S., & Lange, D. (1991). The Use of like as a Marker of Reported Speech and Thought: A Case of Grammaticalization in Progress. *American Speech*, *66*(3), 227–279.

Rossi, G. (2011). *The MPI/Rossi Corpus of English* [Data set]. The Max Planck Institute.

Sacks, H. (1992). *Lectures on Conversation* (G. Jefferson (ed.); Vols. 1 & 2). Basil Blackwell.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*(4), 696–735.

Schegloff, E. A. (1998). Body Torque. *Social Research*, *65*(3), 535–596.

Selting, M., Auer, P., Barth-Weingarten, D., Bergmann, J., Bergmann, P., Birkner, K., Couper-Kuhlen, E., Deppermann, A., Gilles, P., Günthner, S., Hartung, M., Kern, F., Mertzlufft, C., Meyer, C., Morek, M., Oberzaucher, F., Peters, J., Quasthoff, U., Schütte, W., Stukenbrock, A., Uhmann, S., (2011), "A system for transcribing talk-in-interaction: GAT 2", Gesprächsforschung: Online-Zeitschrift zur verbalen Interaktion, vol. 12, pp. 1–51

Sidnell, J. (2006). Coordinating Gesture, Talk, and Gaze in Reenactments. *Research on Language and Social Interaction*, *39*(4), 377–409.

Sidnell, J. (2012). Turn-Continuation by Self and by Other. *Discourse Processes*, *49*(3-4), 314–337.

Stec, K., Huiskes, M., & Redeker, G. (2015). Multimodal analysis of quotation in oral narratives. *Open Linguistics*, *1*(1), 291.

Stec, K., Huiskes, M., & Redeker, G. (2016). Multimodal quotation: Role shift practices in spoken narratives. *Journal of Pragmatics*, *104*(Supplement C), 1–17.

Streeck, J. (1993). Gesture as communication I: Its coordination with gaze and speech. *Communication Monographs, 60*(4), 275–299.

Streeck, J. (2009). Forward-Gesturing. *Discourse Processes, 46*(2-3), 161–179.

Streeck, J. (2014). Mutual gaze and recognition. In *From Gesture in Conversation to Visible Action as Utterance* (pp. 35–56). John Benjamins Publishing Company.

Stukenbrock, A. (2014). Pointing to an "empty" space: Deixis am Phantasma in face-to-face interaction. In *Journal of Pragmatics*, 74, 70–93.

Szczepek-Reed, B. (2006). *Prosodic Orientation in English Conversation*. Springer.

Tannen, D. (2007). "Oh talking voice that is so sweet": constructing dialogue in conversation. In *Talking Voices: Repetition, Dialogue, and Imagery in Conversational Discourse* (pp. 102–132). Cambridge University Press.

Thompson, S. A., & Suzuki, R. (2014). Reenactments in conversation: Gaze and recipiency. *Discourse Studies*, *16*(6), 816–846.

Uhmann, S. (1992). Contextualizing relevance: On some forms and functions of speech rate changes in everyday conversation. *The Contextualization of Language*, 297–336.

Yasui, E. (2013). Collaborative idea construction: Repetition of gestures and talk in joint brainstorming. *Journal of Pragmatics*, *46*(1), 157–172.

Walker, G. (2004). On some interactional and phonetic properties of increments to turns in talk-in-interaction. In *Typological Studies in Language* (pp. 147–169). John Benjamins Publishing Company.

Walker, G. (2017). Visual Representations of Acoustic Data: A Survey and Suggestions. *Research on Language and Social Interaction*, 1–25.

Acknowledgements