Acceptability and Disalignment in Co-Animation

Sharmin Nadia Choudhury

MRES
University of Huddersfield
School of Arts & Humanities
January 2022
Abstract

This thesis examines acceptability and disalignment in the understudied interactional practice of co-animation (Cantarutti 2020) in British podcast interactions. This study builds upon the existing research by adopting the theoretical and methodological tenets of Conversation analysis, Interactional Linguistics and Multimodal analysis to determine the properties and environment which contribute to acceptability within animated sequences.

The way in which participants design (co)animated utterances to minimise miscalculation and disalignment through management of epistemic entitlement, sequential positioning and turn design is analysed in multi-party interactions. The study revealed that whilst animators worked to level their epistemic entitlement through the design of their initial turn to facilitate safe entry into the animation space, the point at which co-animators chose to enter was also integral to the way in which the co-animated utterance was received. Co-animators were also found to be using devices such as formulations (Heritage and Watson, 1979) and reformulations (Gülich & Kotschi, 1995; Caipuscio, 2003; Cuenca, 2003; Murillo, 2012) due to their paraphrastic and implicative nature.

The study also provides a basis for the analysis of multimodal appreciation devices and third-party appreciation in co-animation as well as the investigation into the affiliative nature of co-animation in primary and secondary audiences.
Table of Contents

Abstract .............................................................................................................................................. 2

Table of Contents ............................................................................................................................. 3

Table of figures and visual representations ..................................................................................... 5

Copyright Statement .......................................................................................................................... 6

1. Introduction. Interaction as a form of doing ................................................................................. 7
1.1 Interaction roles ............................................................................................................................ 7
1.1.1 Cooperation and collaboration ............................................................................................... 8

2. Literature Review. What is animation? ......................................................................................... 14
2.1 Introduction .................................................................................................................................. 14
2.2 Reported Speech .......................................................................................................................... 15
2.3 Animation: why not demonstrations, enactments, or reenactments? ........................................ 23
2.4 (co)animation ............................................................................................................................. 30

3. Methodology .................................................................................................................................. 45
3.1 Introduction .................................................................................................................................. 45
3.2 A Conversation Analytic approach ............................................................................................. 45
3.3 Data .............................................................................................................................................. 50
3.3.1 Description of the data and ethical issues ............................................................................. 50
3.3.2 Limitations of the corpus ......................................................................................................... 54

4. Co-animations: acceptability and safe entry into the animation space ........................................ 56
4.1 Introduction .................................................................................................................................. 56
4.2 Epistemic positioning .................................................................................................................... 56
4.3 Projectability ................................................................................................................................. 70
4.4 Formulations and reformulations ............................................................................................... 73

5. Co-animations: acceptability and disalignment ....................................................................... 85
5.1 Introduction .................................................................................................................................. 85
5.2 Epistemic positioning .................................................................................................................... 85
5.3 Projectability ................................................................................................................................. 88

6. Conclusion ...................................................................................................................................... 104
6.1 A summary of findings and thesis organisation................................................................. 104

7. References............................................................................................................................108
   Appendix I ..............................................................................................................................113
   Appendix II.............................................................................................................................115
Table of figures and visual representations

<table>
<thead>
<tr>
<th>Figure/Extract</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 2.1</td>
<td>Schematic representation of layering</td>
<td>23</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>Schematic representation of interaction in the here-and-now and interaction in the animation space</td>
<td>25</td>
</tr>
<tr>
<td>Figure 2.3</td>
<td>Schematic representation of the distribution of voice</td>
<td>29</td>
</tr>
<tr>
<td>Extract 4.1</td>
<td>(visual representation) “where’s my marks”</td>
<td>57</td>
</tr>
<tr>
<td>Extract 4.2</td>
<td>(visual representation) “where’s my marks”</td>
<td>58</td>
</tr>
<tr>
<td>Extract 4.3</td>
<td>(visual representation) “today I got time”</td>
<td>73</td>
</tr>
<tr>
<td>Extract 5.1</td>
<td>(visual representation) “where you going”</td>
<td>82</td>
</tr>
<tr>
<td>Extract 5.2</td>
<td>(visual representation) “where you going”</td>
<td>83</td>
</tr>
<tr>
<td>Extract 5.3</td>
<td>(visual representation) “I’ve got this”</td>
<td>91</td>
</tr>
<tr>
<td>Extract 5.4</td>
<td>(visual representation) “I’ve got this”</td>
<td>92</td>
</tr>
</tbody>
</table>
Copyright Statement

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns any copyright in it (the “Copyright”) and s/he has given The University of Huddersfield the right to use such Copyright for any administrative, promotional, educational and/or teaching.

ii. Copies of this thesis, either in full or in extracts, may be made only in accordance with the regulations of the University Library. Details of these regulations may be obtained from the Librarian. This page must form part of any such copies made.

iii. The ownership of any patents, designs, trademarks and any and all other intellectual property rights except for the Copyright (the “Intellectual Property Rights”) and any reproductions of copyright works, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property Rights and Reproductions cannot and must not be made available for use without permission of the owner(s) of the relevant Intellectual Property Rights and/or Reproductions.
1. Introduction- interaction as a form of doing

1.1 Interaction roles

The consensus on human beings is that we have an innate need to interact with one another. *Human beings are social creatures*-this phrase is often used when describing humans. Thus, to socialise humans must interact and there are a variety of ways in which we, as humans, *do* interaction. One of which is through talk. We use talk to *do* a range of things. We whine; we moan; we cry; we laugh; we plead; we shout and often we do not say anything at all. In turn we are *doing* complaining, begging, arguing, humouring and so much more.

Whilst many of us may laugh to ourselves or cry by ourselves, a large portion of human interaction is done with others. The word ‘interaction’ is composed of ‘inter’ and ‘action’. ‘Inter’ is understood as being between something. Thus, interaction is the action between humans in this case. Therefore, we recognise that there are not only speakers but there are also listeners. In order to be understood there are a multitude of ways in which we often accommodate to our listeners and our listeners display listenership. The act of talking is a joint action which requires cooperation much like other joint activities in which we must cooperate with each other such as football or hockey. Similar to the way in which roles are assigned in football and each member of the team has their part to play, there are roles in conversations also. There are speakers and listeners; questioners and answerers; advise seekers and advice givers and so on and so forth. The roles which we take in conversation
are essential to the accomplishment of activities in our daily lives and the organisation of our human interactions.

1.1.1 Cooperation and Collaboration

Central to the most basic roles of speaking and listening is the act of *turn taking*. The systematic organization of *turn taking* has been expertly analysed by Sacks, Schegloff and Jefferson (1974) in one of the papers that formed the basis of the Conversation Analytic approach (CA, hereafter). The study describes the way in which turns are allocated, ordered, and coordinated in naturally occurring speech. It establishes that generally turns are allocated to one party at a time and the transition of turns is coordinated by tools such as *transition-relevance places* (TRPs, hereafter) and *turn constructional units* (TCUs, hereafter) which seek to minimise gaps and overlaps. Thus, the number of participants in a conversation does not bear weight on the distribution of the turns but rather the organisation of the parties.

Ford, Fox, and Thompson (1996) have since analysed the notion of TCUs at a greater depth and in doing so found that:

“The Sacks et al. model employs the notion of unit in order to account for projection, a concept which is meant to capture the fact that participants can and do orient to utterances as having identifiable trajectories, that is, beginnings, middles, and ends. And the units which are considered to produce such trajectories, at least in the original 1974 model, are interpretable (and indeed have been so interpreted) as an inventory of structures. While we ob-
viously accept the existence of projection, our inquiries have led us to consider the possibility that utterances can have trajectories without being constructed out of clear units.”

(Ford, Fox, and Thompson, 1996:432)

The move away from clear cut units made in this study is further examined by Lerner (2002, 1996b). Lerner suggests that turns can be shared. Participants can share turns through anticipatory (Lerner, 1996b) and collaborative (2002) completions. With anticipatory completions, the onset of the utterance occurs at a TCU internal component completion, and not where the turn itself could be completed. The next speaker begins their anticipatory completion before the projected completion of a TCU and so the utterance takes place within the turn space of the current speaker.

Whereas collaborative completions are not utterances that pre-empt what the previous speaker was about to say but rather a co-production. “And many of them seemed designed to match the voicing of the current speaker quite closely—that is, they resulted in a fully choral rendition of the turn's completion” (Lerner 2002:227). This voice matching to the current speaker is a phenomenon that has been since termed a form of co-animation (Cantarutti, 2020). Whilst collaborative completions are not limited to reported utterances, co-animation refers to speakers co-producing in sequences that have traditionally been termed reported speech. However, Cantarutti (2020) in her study of co-animation found the implications of the term ill-fitting to the practice at hand. She asserts that the use of the term reported speech implies that it is only speech that is reported and therefore the term animation (see chapter 2), which encompasses more than just talk, would be better suited. Furthermore whilst, collaborative completions have been analysed as a practice of speakers “sharing a moment of speaking together” (Lerner 2002:249) Cantarutti (2020) builds upon this notion of speaking
However, she asserts that co-animation goes beyond “speaking together”, as doing together is a characteristic of all interactions, and rather co-animators are doing being others. For example, in an interaction taking place between A and B, co-animation is the act of B doing being A (Cantarutti 2020). There is a fusion between the identities that is managed by both participants within the realm of the interaction; B is not only speaking on behalf of A but is speaking as A.

Let us examine an example of co-animation in the following conversation: a podcast interaction (see chapter 3 for further details on the data examined in this thesis) taking place between the 3 hosts, T, VP and F, reminiscing the way marks were returned in school. T is expressing the competitive nature between the students and the impatience surrounding the wait to receive his marks.

Extract 1.1 “where’s my marks”

1. T: it was fcompetition in school lihf like rah what did you get
2. like >I wanna< I wanna ace this ;one; (. ) you’d give it in?
3. (. ) teacher’s takin ages to mark it and you’ll be hassling em
4. like where’s [my mark ]
→ 5 VP: [where’s my marks]eh ha [hah ]=
6. F: [fam::]
7. VP: =[hah: ha: ]
8. T: [coz I need to know what I got]
9. H: yep::
→ 10 F: I’m tryna out do him *fam* 
11 VP: [heh heh]
12 T: [yeheheh]

VP takes a turn during T’s turn unit in the initial animation. VP produces what Lerner (1996b) would term as a collaborative completion. He produces the utterance in overlap with T’s utterance
in chorus. The design of the utterance to match the voice of the initial speaker-T is what Cantarutti (2020) has termed co-animating in the *voice of other* (see chapter 2). The same phenomenon can also be seen in F’s production on line 10. F does not collaborate in T’s turn unit but rather collaborates to further the narrative through the production of an utterance designed to match T’s *voice*.

The appreciation of the co-animation can be seen on lines 11 and 12. A large difference between collaborating to complete an utterance being produced by the current speaker and collaborating by creating an entirely different utterance in the speaker’s voice is the interactional risk. As can be seen with VP’s utterance, T begins to say, “*my mark*” as VP comes in with “where’s my *marks*”—the ending of the utterance is comparatively more projectable than the direction of the narrative. Therefore, F runs more of a risk of getting it *wrong*. Co-animation going *wrong* can be seen in extract 1.2 below in which the participants are discussing design technology projects they undertook in school:

Extract 1.2 “oh that’s nice”

1 T: THEY MADE [me: make ]=
2 F:           [ oh my god]
3 T: =an egg holder like the fridge don’t have them
4 F: [fam:: ]
5 H: [[ah hah] hah hah ha]
6 VP: [ah huh huh huh huh] [huh huh ]
7 T:                       [mans got the] holes and everything in
8       [the fridge      ]
9 H: [but that’s the] easiest thing [to make:] 
10 T:                                 [I’m bringin] mum look what I
11      [brought an [fegg holderfahah]]
12 F:  [ehhhhh huh ]
T is narrating his mother’s reaction to the egg holder he made in his design technology class. On lines 10 and 11 T is animating the way in which he presented the egg holder to his mother. VP produces a co-animation on line 13 which is presented in the voice of T’s mother. His co-animation displays a positive attitude towards T’s school project which T then on the next proves was not the case. His mother’s response to being presented with the egg holder was in fact seemingly negative. She appears displeased with its ability to “only hold one egg”. The mis-projection resulted in the sequential deletion (see chapter 5) of VPs utterance. Projectability is one factor, amongst others, which will be analysed in this thesis to assess its contribution to the acceptability of co-animated utterance.

This study will analyse the practice of co-animation in a collection of 173 minutes of podcast talk. The examination of its production within an unelicited setting, allows for the analysis of the environment within which it is recurrent and the ways in which participants within the interaction navigate acceptability. The way acceptance or lack thereof of co-animated utterances is managed within animated sequences will be examined seeking to determine the environment in which these responses are recurrent. The data in this collection consists of multi-party interactions, during a series of podcasts, between a minimum of three participants. This allows for the analysis of sequences consisting of co-animations produced by more than one participant, as can be seen in extract 1.1,
and the assertion of how this may contribute to the design and reception of the co-animated utterances. First, however, the phenomenon of *animation* and *co-animation* must be further outlined.
2. Literature review: what is animation?

2.1 Introduction

As established in the previous chapter, the focus of this research is to determine the characteristics of the environment which contribute to the acceptability of the joint act of doing being one’s past self or other, through the recontextualization of voice. In order to be able to outline these characteristics, it is first important to define the phenomenon (co)animation itself. This chapter aims to review existing literature surrounding animation and other related practices.

The phenomenon which will be referred to as animation in this thesis has been studied from a number of different angles and under a number of different terms, one of which is reported speech. In a majority of early literature relating to reported speech there is a primary focus on written reported speech in both fiction and non-fiction text (Leech and Short, 1981; Simon-Vandenbergen, 1981; Sternberg, 1982a, 1991; Fairclough, 1992, as cited in Waugh, 1995). However, more recent studies have examined reported speech in naturally occurring talk, particularly in everyday interactions (see Holt, 1996, 2000, 2016; Drew, 1998; Couper-Kuhlen, 2007; Goodwin, 2007; Holt and Clift, 2007), but also in institutional talk (e.g., Buttny 1998).

This chapter will review the key literature and research around animation and reported speech. The following subsections will review the different terms given to the phenomena and why the term animation is best suited for the practice examined in this paper.
2.2 Reported speech

Literature surrounding reported speech traditionally distinguishes between two categories: “oratio recta (direct quotation) and oratio obliqua (indirect quotation)” (Coulmas 1986:2), also known as direct reported speech (DRS, hereafter) and indirect reported speech (IRS). The distinction between the two forms is consistent in both written and spoken interaction. DRS is considered to retain the original content of an utterance or utterances and seemingly the structure and form. Holt (1996) examines a telephone conversation in which the speaker is talking about an event she attended to her mother and uses both DRS and IRS in her retelling:

Extract 2.1 “Misses field” (in Holt, 1996:220)\(^1\)

1 Lesley: An' it wz really Dick\(\text{-}\)en\(\text{-}\)sian.
2 (0.2)
3 Mum: Oh ↑yes they really [did it,.
4 Lesley: [.t.hhhh ↑So then - ↑I looked
5 f'\underline{my} \underline{Pickwick} (. \underline{↑Pa}↑papers you know .hh[\h
6 Mum: \[Yes,\[.[Yes,
7 (0.4)
8 Lesley: An\(\underline{d} \underline{em} \underline{(0.6)} \underline{t}'have a look at- (. }how-in::
9 tho::se, (.) to compare
10 Mum: (Mm hm)=
11 Lesley: =and em I couldn't ↑find them 'n then ↑suddenly
12 we realized th't ↑I'd mi- ↑I: lent'n to Missiz
13 Field (0.2) ↑years ago.
14 Mum: Oh:.
→ 15 Lesley: .hhh An' ↑Mark as\(\underline{k}\)ed her f-back ↑for th'm this
→ 16 evening 'n she said .hh ↑Oh I thought Lesley had

\(^1\)This extract has been subject to alterations. For the original and complete extract see Holt (1996:220).
On line 15 Lesley produces IRS: “Mark asked her f-back f-or th'm this evening”. IRS does not maintain the structure or form of the original utterance, as DRS is suggested to do, but rather IRS often describes the content of the original message. Thus, although the original utterance was directed at “Missiz Field”, a third-person pronoun is used to refer to her, as opposed to the first-person pronoun which was likely used in the original context. The form has been altered but the content of the original utterance is described. In this case, the original utterance or utterances are described as doing the job of requesting the books back that Lesley had lent to “Missiz Field”. Lesley then produces DRS on line 16: “she said .hh Oh I thought Lesley had given me the:se”. In this utterance Lesley uses the first-person pronoun which would have been used in the original setting and refers to herself in third person, calling herself by name. The form and structure of the reported utterance is seemingly identical to the original utterance as well as the content.

What is traditionally termed IRS and DRS are both resources for doing animation. However, the act of doing being one’s past self or other through the use of recontextualizing voice, bears more of a resemblance to spoken DRS than IRS. The focus of this chapter, having made a brief distinction between the two most studied types of reported speech, will now turn its attention to the characteristics of DRS and why the term is ill-fitting to the phenomena at hand.

Klewitz and Couper-Kuhlen (1999) examined the way in which spoken DRS is marked by speakers through drawing parallels with its written counterpart. They explain that written DRS is explicitly
marked with a quotative verb phrase (verbum dicendi) and quotation marks. The boundaries of what is DRS and what is not are evident in the punctuation used. Moreover, the boundaries are marked identically, consistently, and exactly. Quotation marks, as is suggested by the name, mark the boundaries of quotes. Thus, by convention, the words contained within them are what is being quoted and the words outside are ascribed to the reporting speaker (Klewitz and Couper-Kuhlen, 1999). However, this is not the case with spoken DRS. The shift in speaker voice is not typographically marked. One way in which the recipients are able to identify a shift in speaker voice, from speech assigned to the speaker in the here-and-now, is through prosodic shifts. Couper-Kuhlen (1998) examines a reported speech sequence, in extract 2.2 below, which exemplifies this understanding:

Extract 2.2 “chocolate is toxic” (in Couper-Kuhlen, 1998:5)

1   J: oh and on Sunday, you can open the first door of the advent calendar.
   All: aahh oohh
   J: I never had one before.>
5   A: did you buy it?
   J: mhm,
   A: you didn’t buy the kind with chocolates in it?
   J: no.
   All: huh huh huh huh
10  J: I didn’t think I needed any < chocolate (thing).>
   → it’s a < to:xin you know.>
   A: who’re you talking to (.talking about.
   J: there-
   → 15  < chocolate is toxic.> ->
   A: yeah but you said that like somebody says that.

---

2 This extract has been subject to alterations. For the original and complete extract see Couper-Kuhlen (1998:5).
Couper-Kuhlen (1998) observes that Joy’s utterance on line 12 is understood to be reported speech. As she explains, it is not the content of the utterance which is the issue here. Although Joy appears to believe so as she rephrases her utterance on line 15 to explicitly state the subject of her previous utterance following Ann’s query. However, after Joy’s further clarification, Ann says on line 16 “yeah but you said that like somebody says that”. The delivery of the utterance is what is being queried by Ann; the way in which it was “said...like somebody says that”. Couper-Kuhlen (1998) states that both the utterances on line 12 and 15 were produced in a lower pitch than the surrounding talk and ended in a final glottal creak. It is this prosodic and paralinguistic shift in the production of these utterances which appear to be hearable to Ann as marking a change in voice.

Furthermore, Klewitz and Couper-Kuhlen (1999) explain that prosodic shifts can mark the left and right hand boundaries in spoken reported speech:

“prosodic changes can function like quotation marks in written texts by clearly delimiting left and right hand boundaries of the reported sequence. In the majority of cases, however, prosodic changes do not coincide with the boundaries of reported speech but occur nearby, functioning like a 'frame' for the interpretation of a sequence as reported or even only as a 'flag' attracting attention and inviting the listener to actively (re-)construct the corresponding boundaries” (Klewitz and Couper-Kuhlen,1999:459)
As mentioned above, typographical markers and cues are not available in spoken reported speech sequences. Punctuation, namely, quotation marks and italicisation can mark prosody and DRS in written text (Halliday, 1985, as cited in Klewitz and Couper-Kuhlen 1999). However, without these resources available to them, Klewitz and Couper-Kuhlen (1999) examine the different ways in which speakers mark reported speech boundaries with prosodic shifts. They observe that a shift in prosody does not mark the exact boundaries of a reported speech quotation but rather is used as a flagging device to signal that a change in voice may take place in the vicinity. Prosodic shifts used in this way are not exclusive to only DRS. Klewitz and Couper-Kuhlen (1999) examine instances of IRS which are prosodically flagged in much a similar way. Thus, although in written reported speech sequences, only DRS is marked with typographical conventions, this is not the case in spoken reported speech sequences. Finally, they observe that in some instances the prosodic shifts are not consistent throughout the reported utterances. In some instances, in their data the left hand boundary was clearly marked but by the end of the quotation the speaker had returned to their normal register.

Quotative verb phrases are also typically used in spoken DRS. There are variety of verbs that have been examined in the quotative system in English. Some of the more traditional forms include say, go and think. However, more recent studies have turned their attention to the be+like quotative structure. It has been observed to be growing as a quotative verb phrase in American, Canadian and British English (Tannen, 1986; Tagliamonte and Hudson, 1999).

One key distinction to make between be+like and other quotative verb phrases such as say, go, think etc. is that the utterance which follows is not necessarily recontextualised speech (Butters,
1982; Tannen, 1986; Romaine and Lange, 1991; Tagliamonte and Hudson, 1999). Tannen (1986) explains that be+like prefaced quotations can also be internal thoughts or states of emotion:

““To be” + "like" thus functions as a formulaic introducer, not by its literal meaning but simply by convention. If the literal meaning functions at all, it is to suggest that the dialogue is not being quoted but simply represents the kind of thing that character was saying or thinking.” Tannen (1986:321)

Therefore, the quotations are ambiguous, and the boundaries of speech and thought is often blurred. There is also the notion that the quotations are approximations; the “kind of thing[s]” (Tannen 1986:321) the speaker was thinking or feeling at the time. It follows then that these types of quotations do not fall securely into any of the previous categories described. Thoughts and states of emotion are abstract entities. Thus, it would not be possible to retain the form/structure as well as the content of the thought in the ways in which DRS is suggested to do. Verbatim quoting, which DRS also suggests, has been queried in numerous studies (see Goody, 1977; Sternberg, 1982; Tannen, 1986; Mayes, 1990; Romaine and Lange, 1991; Clift, 2007; Holt, 2009). The notion that the human mind has great difficulty recalling the exact intricacies of a previous spoken utterance to be able to replicate it verbatim and void of bias (Lyons, 1977; Tannen, 1986), would also suggest that the recollection of unvoiced thought would prove just as, if not more, difficult. There is also an issue of recontextualization. Sternberg (1982) describes the different facets which comprise context when reporting:
“the reporting and the reported event are (and remain) essentially distinct and independent entities, each with its own participants, spatiotemporal anchorage, sociocultural matrix, generic and linguistic conventions, and expressive goals-in short, its own context in the most inclusive sense of the term” (Sternberg 1982: 130)

Thus, it follows that taking an utterance from its original context and recontextualising it inevitably alters the utterance. Furthermore, the use of be+like in reported speech has been suggested to convey thoughts and states of emotion that speakers may, or may not, have vocalised at the time (Butters, 1982; Tannen, 1986; Romaine and Lange, 1991; Tagliamonte and Hudson, 1999). Therefore, the boundaries of thoughts, opinions, and speech can be blurred in reported speech. The term reported speech suggests voiced utterances which is not the case with internal thought and states of emotion. Thus, the term reported speech for these types of utterances is ill fitting and in some studies the authors have opted for the term reported thought instead. It goes without saying that the studies on reported speech which have been undertaken over the years do not necessarily align with the implications and biases of the term. As can be seen in the studies outlined above, internal thought and states of emotion in reporting has been considerably examined and analysed. However, the term itself, whilst has been fitting of the observations made in the past, is not a term which is fitting to carry forward for the research in this paper. Thus, despite the similarities in reported speech sequences and the sequences which will be analysed in the upcoming chapters the term animation is favoured over reported speech. This subsection has examined some of the characteristics of reported speech and the shortcomings of the term with regards to the phenomenon at hand. The
following subsection will examine terms which also bear resemblance and overlap with the term
animation albeit further expanding on the characteristics in ways which reported speech does not.

2.3 Animation: why not demonstrations, enactments, or reenactments?

Demonstrations, a term pioneered by Clark and Gerrig (1990), are suggested as an alternative term to reported speech. However, they draw a distinction between what it is to describe an interaction versus to demonstrate one. The term separates interactions into serious and nonserious subcategories, a theory which was initially proposed by Goffman (1974, as cited in Clark and Gerrig, 1990).

Clark and Gerrig (1990) explain that demonstrations are nonserious actions. They state:

“Human actions, according to Goffman 1974, divide into two broad types—serious and non-serious. When two people fight, their actions are serious. The actions ‘are said to be or actual, to be literally occurring’ (47). But when two people PLAY at fighting, as in make-believe, their actions are nonserious. The fight is ‘not literally or really or actually occurring’ although the play itself is.” Clark and Gerrig (1990:766)

Thus, the utterances which are being quoted, in the original setting may have been doing the job of requesting, offering, explaining, complaining etc. but when they are taken from the original setting and recontextualised into the current speaker setting the actions become non-serious (Clark and Gerrig 1990). And so, Clark and Gerrig (1990) suggest that quotations are non-serious actions that are contained within serious actions. For example, a speaker may be complaining about a rude customer they dealt with at work, and to do so they quote what the customer said to them. In this sce-
nario complaining is the serious action and the utterances, produced when quoting, are demonstrations. And so, it follows that they are non-serious actions. Clark and Gerrig (1990) describe DRS, IRS and free indirect reported speech (FIS) in relation to solidarity, engrossment and ineffability amongst other things. The most notable way in which demonstrations expand on previously established ideas around speech reporting, is that they do not only refer to speech and thought, but also gesture. Demonstrations in Clark and Gerrig’s (1990:766) definitions are “performed” and so they can consist of quotations of both speech, thought, changes in prosody and gestures. They are suggested to be multi modal interactions. However, they need not include quotations at all. Clark and Gerrig (1990) give the example of someone mimicking a limp. The demonstration in this example is mimicry through gesture. However, they explain that “their boundaries—their beginnings and ends—must be clear” (Clark and Gerrig, 1990:766); the demonstration must be distinguishable from the serious action surrounding it. Thus, it would not be acceptable for the person to simply start mimicking a limp. The action should be marked as a demonstration by the speaker stating in some way that they are performing a non-serious action and they themselves do not have a limp (Clark and Gerrig, 1990). This clear marking of beginnings and ends, referring back to Klewitz and Couper-Kuhlen (1999), is not always the case in reported quotation. The boundaries of what is non-serious action may not be always clear cut. As Klewitz and Couper-Kuhlen (1999) explain, the boundaries of prosodic shifts, which can be used to mark demonstrations, are not always exact and in some instances the left hand boundary of the quotation can be marked but by the end of the quotation the speaker may have returned to their normal register. Furthermore, in extract 2.2 Joy does
not clearly mark the boundary of, what Clark and Gerrig (1990) may refer to as, her demonstration which can be seen in the request for further clarification that follows.

Extract 2.2 “chocolate is toxic” (in Couper-Kuhlen, 1998:5)

10 J: I didn’t think I needed any chocolate (thing).>
   → it’s a < to:xin you know.>
   A: who’re you talking to (.) talking about.
   J: there-
   → 15 < chocolate is toxic.> ->
   A: yeah but you said that like somebody says that.

As, previously stated in sub chapter 2.0, Ann understands Joy to be shifting voice when she produces “< to:xin you know.>”. However, Joy has not clearly marked this as being a demonstration. It is after further questioning that Ann explains Joy “said that like somebody says that”. Thus, the boundary is blurred here, and it is not clear that Ann is ‘performing’ a non-serious action. That having been said, the concept of performing and the connotations which are attached to the word demonstration are ill-fitted here. Both performing and demonstrating have the implication of clear boundaries being drawn between what is being performed or demonstrated and its surrounding interaction (see subchapter 2.4) . This is often not the case, as can be seen in extract 2.2 above.

Moreover, Clark and Gerrig (1990) also explore the notion of there being levels of description. They provide an example of Kate quoting herself in a previous interaction. They then suggest that if Kate is trying to depict the lower level of description she may only be demonstrating the “loudness,
intonation, and gestures of the original speech” (Clark and Gerrig 1990). Whereas, if she were trying to depict the higher level of description she would be claiming to represent the emotions she felt at the time. They assert there is a distinction between the levels. Thus, the concept of doing being someone else, which largely embodies the phenomenon of animation, through a shift in prosody, gestures and assigning quotations, does not fit with the implications of demonstrations. Doing being and demonstrating being, although the difference is not a large one, are distinct in what they are describing. Of course, it is not possible to actually be someone else even if that person is one’s previous self in a past situation. Animators are only able to embody being another person, be it themselves, to animate a version of them through their interaction.

Furthermore, other terms which are closely related to demonstrations are enactments and reenactments. The terms enactment (Goodwin, 1990; Streeck and Knapp, 1992; Wilkinson et al., 2010) and reenactments (Sidnell, 2006; Thompson and Suzuki, 2014) seemingly refer to the same phenomena under different terms. As can be seen in the study by Thompson and Suzuki, on gaze and recipiency in reenactments, in which they state (2014:27) “Wilkinson et al. (2010), in their discussion of speakers with agrammatic aphasia re-presenting previous events, use the term ‘enactment’. We appreciate and build on their embracing of the bodily-visual aspects of ‘reported speech’ and choose to follow Sidnell in using the related term ‘reenactment’’. Both terms embrace a multimodal approach to reported speech and quotation and much like demonstrations they can include interaction that does not involve spoken utterances. The key distinguishing factor then between demonstrations, enactments, reenactments and animation is the notion of an ‘animation space’
Cantarutti (2020) distinguishes between an ‘animation space’ and ‘interaction space’. The animation space is a spatial metaphor that refers to the orientation that speakers take up when animating. However, the interaction space refers to the literal setting and current interactions taking place in the here-and-now (Cantarutti 2020).

The concept of these spaces in which one exists within the other, and interactions simultaneously taking place in both, is not dissimilar to Clark’s (1996) notion of ‘layering’. He uses a simplified sequence taken from Sacks, 1974, to illustrate the different layers which can be found in an interaction sequence:

Extract 2.3 “Grape, Britain” (Sacks, 1974, in Clark 1996:15)

1   Ken: You wanna hear—My sister told me a story last night.
2   Roger: I don't wanna hear it. But if you must. (0.7)
3   Al: What's purple and an island. Grape, Britain. That's what his
4       sister told him.
→ 5   Ken: No. To stun me she says uh, (0.8) There were these three girls
→ 6   and they just got married? [Continues joke]

Clark (1996:15) explains that the utterance produced by Ken on line 1 is an assertion made in the “actual world conversation”. Whereas the utterance “‘There were these three girls and they just got married?’ is only true in the “hypothetical world” (Clark 1996:16) in which the joke is taking place. Clark (1996) calls these different worlds layers of action. He identifies that there are two layers present in the interaction above. Layer 1 is the primary layer which is present in any interaction where the participants are speaking and addressing one another in the here-and-
now (Clark 1996). Whilst layer 2 is built on top of layer 1 and in this interaction it represents the hypothetical world. Clark notes that if Roger and Al had understood Ken’s utterance on line 1 to be in the hypothetical world of layer 2 or if they had understood the three married girls to be in the primary layer this would be a misunderstanding. Thus, participants in interactions must recognize, however imprecisely, the different layers which are present.

Figure 2.1 Schematic representation of layering (Clark, 1996:16)

Clark (1996) uses the figure above as a representation of the layers in the interaction taking place between Ken, Roger, and Al. Layer 1 is the ground layer, which is present in all forms of language use, whilst layer 2 is a temporary layer that is built on top of layer 1 (Clark 1996). Clark explains that in this interaction layer 2 is supported by layer 1. There could not be a hypothetical world without it being built upon the real-world. Layer 2 is also optional, and it is possible to have interactions which are contained with a layer 1 domain and further layers are not built upon them. Clark also explains that it is possible to build higher layers. See Clark (1996) for a breakdown of further layers and domains in both written and spoken forms of interaction. Clark (1996) explores the notion of layering in novels and dramas as well as exploring layering in storytelling, sarcasm, irony, teasing and fictionalisations. Despite the overlap between the notion of layering and the different
However, higher layers are not relevant to the analysis in this thesis. In animation sequences layer 1 are similar to what is described in the Clarkian theory⁴; they are taking place in the real-world interaction or the here-and-now. Cantarutti (2020) has termed this the ‘interaction space’. What Cantarutti (2020) has termed the ‘animation space’, is not dissimilar to layer 2. The domain in which speakers are able to open an animation space through the way they present happenings in the ‘there-and-then’ is built on top of the real word domain. This understanding can be seen when examining the animation sequence below:

Extract 2.4 “calm down” (in Cantarutti, 2020:35)

01 KAR: <<:-) >↑↑¯but ↓¯yeah- > hh°
02 we ll `see how it goes;
03 like i was talking to my ´parents,
04 my uh? (0.3) mums are `mums. = ¡´arent they,>
05 and ↑[¯she] was uh° (0.3)
06 ANG: [¯hm ]
07 KAR: getting ↑all ex`cited about the idea that °h ¯lih°ke
08 °h <<sty, br> we ll be ↑¯settling ↓¯down togethe:r>=
09 =and Ði was like?
10 ↑just ↑calm `d[0:wn:,]
11 ANG: [¯yeah;]

---

⁴ Domains of interaction in animation, the Clarkian theory, will not be further expanded upon in this paper. The theory was called upon to provide a linguistic parallel to the term ‘animation space’.
In accordance with Clark’s notion of layering, the utterances prior to line 8 are contained within layer 1 or, according to Cantarutti (2020), the interaction space. The animation sequence which follows, would be layer 2-the animation space. Cantarutti (2020) conceptualises the animation space versus the here-and-now interaction in much a similar way to Clark (1996) (see schematic representation in figure 2.2 below).

Figure 2.2 Schematic representation of interaction in the here-and-now and interaction in the animation space (Cantarutti 2020:36).

As can be seen in figure 2.2, the way Cantarutti (2020) displays the relationship between the animation space and the here-and-now bears a resemblance to figure 2.1. The ‘there-and-then’ is shown as being built upon the ‘here-and-now’. However, a key distinction between the two figures is Cantarutti (2020) makes a distinction between which of the two interlocutors is building the second layer and opening the animation space. In figure 2.2 it is Karen who is depicted as opening the animation space. Whereas Angie remains in the here-and-now and does not have a direct relationship with the animation space. The animation space does not refer to the narrative sequence itself, as
Angie can be observed to be a participant in the sequence, but rather the metaphorical domain a participant enters by animating. Thus, whilst Angie begins her narrative on line 1 she does not open the animation space until line 8. Also, note that whilst Angie is animating on both lines 8 and 10, the reportative verb phrase on line 9 is outside the animation space. This distinction is essential to the development of this paper. The way in which the animation space is navigated by the interlocutors in relation to who has access and safe entry will be further described and examined in Chapter 4.

This subchapter has examined some of the closely related terms to animation and why, despite the similarities, these terms have not been embraced when examining the phenomena in this thesis. Attention will now be turned to the practice of animation and more specifically co-animation.

2.4 (co)animation

Early usage of the term animation can be found in the work of Goffman (1981). He describes what he defines as an animator below:

“In canonical talk, one of the two participants moves his lips up and down to the accompaniment of his own facial (and sometimes bodily) gesticulations, and words can be heard issuing from the locus of his mouth. His is the sounding box in use, albeit in some actual cases he can share this physical function with a loudspeaker system or a telephone. In short, he is the talking machine, a body engaged in acoustic activity, or, if you will, an individual active in the role of utterance production. He is functioning as an “animator””

(Goffman, 1981:143)
Goffman’s (1981) definition of the terms animator and animations is much broader than the way in which they are used by Cantarutti (2020). His definition is not only applicable to interactions in which a speaker is recounting an interaction which took place at a point in the past nor fictionalisations that are abstract from the present. However, Goffman (1981) does state that in the case that a person is recounting a past interaction, there are two animators. An animator in the here-and-now, who is doing the recounting, and an embedded animator who exists within the world that is being spoken about. Thus, Goffman (1981) distinguishes between the animator in the interaction space and the animator within the animation space. Goffman (1981) also explains that animators can animate the embedded animator’s thoughts, gestures, and speech. Thus, despite the lack of expansion and development, Goffman’s (1981) term animator and animations are not dissimilar to that of Cantarutti (2020).

Furthermore, in Goffman’s (1959) early work he explores the ways in which the study of on-stage actors can be used as an analytical tool to provide an insight into everyday interaction, particularly when focusing on identity. But in Forms of Talk (1981) “Goffman’s metaphor of performance slides into a metaphor of animation, and with this change, certain analytic possibilities come to light and the analytic vocabulary changes as even live actors are treated not with a metaphor of performance but a metaphor of animation” (Manning and Gershon, 2013:110). The distinction made between demonstrating being and doing being in subchapter 2.3, appears to be what is being addressed by Manning and Gershon (2013). Despite, live actors being discussed, the term ‘performance’ seemingly does not accurately capture the essence of what these actors are doing. Goffman (1981) appears to be making a distinction about what it is to perform a character and to animate a
character. Thus, Cantarutti’s usage of the term adopts a similar perspective, and she expands on the notion of embedded animators in recollections and fictionalisations, specifically on what Mathis and Yule (1994) examined under the name of attitude echoing ‘zero quotatives’. Zero quotatives are reported speech quotations which are not prefaced with reportative verb phrases (Mathis and Yule, 1994). The subcategory of attitude echoing zero quotatives focuses on instances of reported speech interactions where a “second speaker adds her voice to the construction of a reported interaction” (Mathis and Yule, 1994:69). The example below is of the second speaker Maya, producing an attitude echoing zero quotative in response to Sara’s speech report:

Extract 2.5 “home truths” (in Mathis and Yule, 1994:69)

```
1 Sara: Kim does that a lot too—she's like: "Every little thing da da da da—
2       I hate it when you do that—I hate it"
3       I'm like: "Let me sit you down and tell you a few home truths"
→ 4 Maya: "Let me get down to the basics for you"
```

Maya on line 4 produces a zero quotative speech report. She does not attribute her utterance to a speaker through the use of a reportative verb phrase. However, Mathis and Yule (1994:69-70) describe this utterance as an “expression of attitude attributable to the first speaker (Sara)”. Cantarutti (2020) terms this phenomenon co-animation. An important distinction to make between the terms echoing attitude zero quotative and co-animation is that co-animation is considered a joint action of doing being ones past self or other. Therefore, rather than approaching it from the perspective of Maya adding her voice to the construction of the reported interaction, Cantarutti (2020) approaches
it from the perspective that Maya is doing being Sara; Maya is animating Sara’s voice in the animation space.

Furthermore, Mathis and Yule (1994) refer to Maya’s utterance as an echo or paraphrase of Sara’s previous utterance. Cantarutti (2020) also identified the paraphrasing characteristic as a facet of co-animations and split them into two categories: formulations and reformulations (see chapter 4).

Thus despite the overlap between Mathis and Yule’s (1994) examining and understanding of the phenomenon, Cantarutti (2020) approaches it from a slightly different lens. Another overlap which can be found is the concept of “merging voices to represent the voice of another speaker” (Mathis and Yule 1994:72). Mathis and Yule (1994) identify and examine instances of zero quotative interactions in which the first speaker is reporting in the voice of other as opposed to self, as well as instances where the second speaker is reporting in the voice of an absent third-party speaker. Cantarutti (2020) also identifies and examines similar instances of speakers animating in voices other than the voice of self. The figure below is representative of the range of different animated voices Cantarutti (2020) identified in her collection.

---

5 Chapter 4 of this thesis briefly outlines existing literature around formulation and reformulation and the ways in which the usage in this paper deviates and builds upon existing research. Also, see Cantarutti (2020) for her insight into the two terms and the ways in which they are used in the animation space.
As can be seen in the figure above, and as explained by Cantarutti (2020), animating the voice of self is the most frequent practice. The first animators most frequently adopt animation to produce the voice of their past self through their speech, thoughts, actions, and attitudes in a past or fictional time and place (Cantarutti, 2020). Cantarutti (2020) further explains this could be due to self-animations allowing the speaker a greater license to the animation as they have first-hand access to the setting and interaction being recontextualised in the animation space. The first animator, when animating self, not only has access to speech and actions they are recounting but also thoughts and attitudes (Cantarutti, 2020).

The second most frequently occurring practice is the animation of the voice of Another. This is the voice of an absent third party (either known to the participants involved or a fictional/constructed
person) (Cantarutti, 2020). Where in the animations of self, potential co-animationers are placed in a delicate situation due to the imbalance of epistemic rights and access (see Chapter 4), animating the voice of another poses a risk to the first animator also (Cantarutti, 2020). The first animator may be judged on the way they present the voice of the absent third party and the validity of the recollection they assert (Cantarutti, 2020). In the case of fictionalisations and constructed voice of Another the judgement is likely to vary in accordance with who the voice is assigned to (Cantarutti, 2020).

For example, fictional animated utterance which are assigned to a person are likely to be judged harsher if inaccurately presented as opposed to animations assigned to fictional characters from a television programme.

This risk factor is also present when animating the voice of other; the third most frequently occurring practice in Cantarutti’s (2020) collection. The voice of other is an animation produced in the voice of the person with which the speaker is interacting (Cantarutti 2020). Thus, should animation be taking place in a two-party interaction the voice of other would be limited to one person but in multiple party interactions the voice of other could be animation in the voice of any present participant in the interaction, as is the case with the data analysed in this paper. Cantarutti (2020) explains that due to the relational risk of speaking on someone’s behalf these interactions are often found in teases and jocular mockery. Haugh and Bousfield (2012) observe that jocular mockery, mock impoliteness, and jocular abuse are in fact solidarity enhancing devices. Thus, the relational risk taken as part of a teasing or jocular mocking sequence is minimised by the nature of the interaction. Co-animation sequences are also categorised by Cantarutti (2020) as highly affiliative due to the often
joint doing being of a participant. It can be noted also that the affiliative nature of co-animation sequences can also be extended to animation sequences where the first speaker is animating in the voice of other to jocularly mock or tease as these interactions are equated with familiarity and solidarity (Kotthoff, 1996 as cited in Haugh and Bousfield, 2012).

Cantarutti (2020) observes two types of teasing sequences in her collection and examines the way in which they are handled. One environment in which co-animated tease sequences can be found is when a participant negatively presents self. The examples in Cantarutti’s (2020) collection consist of self-deprecations and participants self-identified mistakes. The other Cantarutti (2020) examines is sequences in which the participants are discussing a transgression or offence which is treated as comical. An example in her collection is of a participant discussing being “shat” on by a bird (Cantarutti, 2020:395). She explains that in both types of teases the participants (co)animate alternative fictions as a way to “set the record straight” (Cantarutti, 2020:396). The participants create extreme or absurd alternate fictions in order to sequentially delete or override the initial transgressions or face issues presented.

The affiliative nature of co-animation sequences was also found when participants were discussing transgressions and offences in complaints and troubles-telling sequences (Cantarutti, 2020). By a speaker recounting a personal issue that took place in their ‘intimate sphere’ to their interlocutors, the issue then becomes a relational one (Drew, 1998; Emerson & Messinger, 1977, as cited in Cantarutti, 2020:309-310). The issue has been made public and therefore open to scrutiny and judgement. It is also equally open to affiliative, empathic responses (Cantarutti 2020).
Cantarutti (2020) examines co-animation sequences in indirect complaint sequences; sequences in which the subject of the complaint is not present in the interaction. Therefore, indirect complaint sequences can pose moral risks as the complainant could be held accountable for the wrongful judging of the absent third party (Edwards, 2005; Ruusuvouri et al., 2019, as cited in Cantarutti 2020). Animation sequences bring the voice of the complainee into the present interaction and thus the participants in the interaction are able to witness the evidence for themselves (Clift, 2007; Couper-Kuhlen, 2007; Drew, 1998, as cited in Cantarutti, 2020). This may further the moral risk present as not only is the behaviour of the absent third party being described but animation sequences may also include their voice and gestures.

However, Boxer (1993) concluded from a study performed to measure the theory of social distance (Wolfson 1986) that indirect complaint sequences can be a solidarity-building device. She explains that the complaint sequences invoke the recipient to engage in commiserative responses in order to demonstrate listenership and concern to maintain the social relationship or intimacy. Boxer (1993:124) states “with strangers as well as friends and acquaintances, at least one speech act sequence, that of complaint/commiseration, is frequently used not to open conversations and form temporary bonds, but sometimes even to build relationships”. Thus, co-animation as a response to indirect complaint sequences may function as a commiserative device to maintain social solidarity. They are largely found in these sequences in third position as a satisfactory affiliative response to the complaint/troubles telling (Cantarutti 2020).

The example below, found in Cantarutti’s (2020) collection, is of Karen discussing the conversation she had with her mother about her long-distance boyfriend moving to the UK to be with her:
Cantarutti (2020) observes that Karen first introduces the complainable behaviour of her mother on lines 7 and 8. The mother, it seems, is presented as getting carried away with the idea of Karen settling down with her boyfriend. Canatrutti (2020) notes that the mother's voice is animated through an utterance which contains characteristics of both DRS and IRS which can be seen in the usage of “we”. She then animates the voice of self on line 10 with a more cautious tone towards the pace of her relationship. Although, it is unclear whether Karen did produce the utterance on line 10 in the original setting or as an animated response of her reaction in the here-and-now. As Selting (2010:233, as cited in Cantarutti, 2020) notes when past incidents are recollected there is “the reporting of affects from the story world as well as the negotiation of in-situ affects in the here-and-now.”
now of the storytelling situation”. However, Angie on lines 12 and 14 can be seen to be doing being Karen and her utterances appear to be directed at Karen’s mother- addressing her lack of caution.

Angie’s utterance on line 12 can be seen as co-animating Karen’s voice not only through lexical repetition, but she also prosodically matches the rising pitch contour in Karen’s original utterance (Cantarutti, 2020). By echoing Karen’s response to her mother’s behaviour Angie can be seen to be not only empathising with Karen, but also endorsing Karen’s response as a fair judgement of her mother’s behaviour (Cantarutti, 2020). Karen finds Angie’s response to be satisfactory by producing “eXACT[y]” on line 16.

Through the examination of complaints and trouble-tellings in her collection, Cantarutti (2020) creates a list of some of the stages present in these types of sequences:

Initiation

1. A: “Adumbration” of troubles or complaints, underspecified

The anticipation of a complainable or a troubles-telling on its way is the first way of securing some form of go-ahead, not just structural, but also affective. Co-participants have a “prior warning” of the activity and actions projected.

2. B: Affiliation OR Negotiation of the terms of the complaint/trouble

Particularly when co-participants share epistemic access to the events retold, what constitutes the actual trouble or complainable may be a shared matter and may need fixing before the sequence can move forward. When this is not problematic, coparticipants may be producing tokens that act as “go ahead” because
they treat the matter as recognisable but also orient to this as not being the actual core of the activity, as more is projected.

3. A: *Topicalisation of the trouble*

Ensuring that the terms of what is troublesome or complainable are clear and negotiated by co-participants is another way of making sure there is some “safe ground” both in epistemic terms but also to secure affiliation, so at some point in the cycle the trouble or complainable is made explicit and “on the table”, as it were.

4. B: *(Minimal) Affiliation Core: Development, Escalation*

Even though the degree to which affiliation is offered at this stage differs across cases, this is a point in which co-participants offer a response that matches the stance being presented. This response generally confirms to the tellers that a matching stance has been secured, so a more empathic or stronger affiliative reaction can now be pursued more “safely”.

5. A: *Escalation, specification, enhancement, exemplification of the trouble/complainable through animation: o (Complaints only: Voice of the transgressor/offender)*

In complaint stories, at this point, the transgressor is brought to the here-and-now, as evidence and as an interactional “object” to be assessed.

   o *(Complaints AND troubles-telling: Voice of A’s affective reaction/decision)*

Before the extended turn-at-talk is brought to an end and the affiliative response is made relevant, co-participants are provided with one (more) animation, that of the teller, whose thoughts or actual responsive
speech in reaction to the trouble/transgression is made available. Sequentially, then, this animation -over-
laid with particular design features- makes the victim’s stance clear and mobilises a response in a slot that
is safe for the co-participant.

6. B: *Affiliative response involving co-animation of voice of A Closure, exit*

The co-animator instantiates their empathetic insertion and endorsement by literally “doing being” the
voice of A, by engaging themselves in the there-and-then of the animation space and keeping it open,
speaking on their co-participant’s (and their own) behalf. This is the moment of the display of a co-com-
plaint or a co-victim role.

7. A: *Confirmation OR Embedded correction OR Progressivity onto next “order of business”*

As mentioned earlier, this final stage exhibits variation in the collection. Even though no explicit rejection
of the co-animation is provided in any of the examples, no matter the epistemic differentials between the
cooparticipants, co-animations go unsanctioned, and although they may be subject to embedded correction,
they are more frequently than not explicitly accepted by the original tellers.

Figure 2.4 The overall stages of trouble-tellings and complaint sequences in the collection complied by

Cantarutti (2020) explains that these stages can differ in length and whilst some stages may be ab-
sent others may be repeated, particularly the initial stages. Taking extract 2.6 as an example of
these stages, stages 1,3,4,5,6 and 7 are present in the extract. Stages 1 and 3 are present in lines 1
through 5; Karen introduces/‘adumbrates’ the complaint sequence by telling Angie she “*was*
talking to [her] ‘Parents’” and then goes on to topicalise the trouble through her utterance
on line 4. Participant B, Angie, minimally affiliates with participant A on line 6. Karen then specifies the complaint on line 7 and moves on to exemplify it on line 8 through the animation of her mother’s voice. This is then followed by an animation in the voice of self, providing a reaction to the transgression. These animations produced by Karen are stage 5 in this complaint sequence which is then followed by stage 6: Affiliative response involving co-animation of voice of A. Angie provides an affiliative response through her co-animations on line 12 and 14. Finally stage 7 can be seen on line 16. Karen provides a confirmation of Angie’s co-animations.

Through the breakdown and analysis of the stages of complaint and trouble-tellings sequences, Cantarutti (2020:359) concludes that “the co-animation appears to be triggered not by the animation of the “wrongdoer”, but of the “victim”. It is the animation of the response to the reported trouble/complaint that mobilises a response that could but need not necessarily be a “me too”, but which should at least act as an other-legitimisation of the reaction as being valid/measured in connection to what happened”. The co-animations due to their affiliative nature and this possible “me too” factor is largely found after the victim provides their own reaction/response to the behaviour of the wrongdoer, thus, giving the co-animation the opportunity to express solidarity with the victim.

Cantarutti (2020), through the analysis of embodied resources of co-animation sequences embedded in certain social activities, established the way in which co-animation can be used as a resource for association. The study “shows how other voices and even other positions can be made relevant through animation, turned into an object of joint interactional scrutiny, and be jointly rejected, sanctioned, or praised as a team around shared moral values or concerns” (Cantarutti, 2020:437). She
analyses the way in which co-animations are prosodically and gesturally designed and the way in which first animators manage entry to the animation space.

Cantarutti (2020), observed through the examination of the co-animation sequences in her collection that a majority of the co-animations were accepted by the first animator. However, the examination and analysis of co-animation sequences which involved ‘disalignment’ or ‘miscalculations’ made by co-animators is limited. Cantarutti (2020) observes there are instances in which the first animators do not explicitly address co-animation and instead “move on to the next turn in the development of the activity in progress”. She explains that it is ambiguous as to whether the lack of explicit acknowledgement could be a form of acceptance, a possible miscalculation or an orientation to the difference in entitlement when a participant has taken the role of a co-animator (Cantarutti, 2020). This study begins to investigate this ambiguity and examine the way in which acceptance, miscalculations and differences in entitlement is handled by participants in animated sequences.

This study builds upon the analysis Cantarutti (2020) has undertaken on disaligned and miscalculated co-animation sequences and will determine the characteristics of the sequences and the environment in which they recurrently occur. Further to the research undertaken by Cantarutti (2020), who has focused on co-animation sequences in two-party interactions and the affiliative nature of co-animation within these interactions, this thesis will examine co-animation sequences which take place in multiparty settings between 3 or more participants. The study provides an initial investigation into the way participants navigate (co)animation in talk that is designed for both a primary and secondary audience (see chapter 4). Thus, the study provides further insight into the phenomenon of co-animation by examining the way in which they are received and appreciated in group settings;
how the affiliative nature of the device may be impacted by the setting in which they are produced and the environment in which miscalculated or disaligned co-animation sequences take place and how they are handled by the first animator and other participants.
3. Methodology

3.1 Introduction

Having introduced co-animation and covered the relevant literature surrounding the phenomenon in the previous chapter, this chapter focuses on the methodological processes which have formed this paper. The way in which the data was acquired, transcribed and the consent and ethical considerations are described. However, before the formation of this paper can be reviewed, it would be useful to address the approach with which the data is analysed. The following subsection of this chapter is dedicated to outlining the theoretical framework of Conversation Analysis (CA), language in interaction and Multimodality (MM).

3.2 A Conversation Analytic approach

The Conversation Analytic approach is the study of naturally occurring speech in social interactions. The foundations of CA were built upon two main ideas: Goffman’s (1983) ‘interaction order’ and Garfinkle’s concern with ordinary actions in social existence. It arose from the workings of a group of sociologists at the University of California. Among these sociologists was Harvey Sacks, a now renowned name in the field of talk in interaction. Sacks was inspired by his fellow sociologist Harold Garfinkel who pioneered the field of ethnomethodology. Ethnomethodology is primarily concerned with the way in which people comprehend and navigate day to day life. The focus is on action and interaction without any particular focus on conversation. Whilst Goffman stressed the merit of studying face-to-face interaction and a microanalytic approach to the study of conversation
Thus, Sacks began to look into spoken interaction as way of complimenting sociological studies. One of his early papers published in 1972—*An Initial Investigation of the Usability of Conversational Data for Doing Sociology*—begins to investigate the ways in which a focus on conversation can aid in “locat[ing] the collections of membership categories in terms of which the search for help for suicidalness is formulated” (Sacks 1972:31). In this paper he begins an analysis into spoken language and introduces initial investigations into laughter being used as a tool for detopicalisation when discussing suicide. Since its birth, CA has grown as a field and what is distinct about CA is its attention to detail. No detail is too small to be subject to analysis, and each detail is considered an analytical resource. This detail-oriented approach is how conversation analysts determine the formation of sequences of talk in interaction and what these actions are doing. These details, and in turn sequences, are considered the basis of an appreciation by which institutions and interpersonal relationships are formed, maintained, altered and so forth.

CA is also distinct in its focus on the social order of interactions. Often, when a person performs actions such as greeting, offering, inviting, questioning, there is an expectation that the recipient should respond accordingly. If a greeting is offered, a returned greeting is a normative expectation. These sets of actions are called *adjacency pairs* in CA. One action is followed closely by the other. This is not to say that greetings are always met with greetings and questions are always answered. Instead, it speaks to the normative expectations and frameworks of social action which are observed through the way in which these interactions are usually performed.

Similarly, another tenet of CA is a focus on interactional context. Not only through the normative pairing of interactions, but through a focus on turns surrounding an interaction. In order to uncover
what the function of a turn of talk is, and what it is doing, it is useful to examine the turn which immediately follows it. The way in which a turn is received and handled by its audience is often a place in which the answers reside. The turns which precede and follow a turn of talk are rich sources of analysis, as CA is concerned with the orderliness and sequences which underpin social interaction. One of the earliest papers in CA is a paper on turn taking by Sacks, Gail Jefferson and Emanuel Schegloff in 1974. Although Sacks was one of the forerunners of CA, he did not expand on this field alone. The paper focuses on the organisation of turn taking in conversation. For example, during an interview there is an expectation that after a question is posed by the interviewer it is the interviewee's turn and they should provide an answer. There is a specific way in which turns are taken in this setting, as opposed to the turn taking that may occur when we are catching up over coffee with a friend. The paper highlights the orderliness and the sequences behind which turns are taken.

Furthermore, CA transcripts allow for a micro-analytic approach due to the detail which is included within them. The transcription system, initially constructed by Gail Jefferson, facilitates the marking of pitch, length, stress, and volume. These prosodic features, which are audible to the analysts, can also be recorded and displayed in the transcripts. The system also allows for visible records of turns taken between speakers as well as the overlaps and pauses. Heritage and Atkinson (1984) highlight the importance of taking a micro-analytic approach to data by stating:

“[there is] an increasing awareness of the fundamental organizational importance of details that might seem, on a cursory inspection, to be random or insignificant.” (Heritage and Atkinson 1984:12)
Thus, the change in the approach to transcription was critical, especially in areas such as laughter, as it made it possible to describe structures and features of turns and sequences which would have otherwise gone unnoticed (Heritage and Atkinson 1984). Overall, the use of Jeffersonian transcription conventions altered the way in which data is presented in CA papers. And by doing this created a detailed analytical tool. However, conversation analysts by no means claim that data transcripts are interchangeable with the recorded data. The most accurate of transcripts still fall short of being able to capture the detail in recorded tapes. Thus, a transcript does not serve as a literal representation of an interaction. They are simply a useful research tool that are used alongside the recorded data to better facilitate analysis and display findings.

The use of recorded data is useful to both the researchers and readers. Conversation analysts refrain from approaching data with preconceptions of what is taking place. By approaching data looking for specific phenomena, the analyst is limited and may misinterpret the data. Therefore, recorded data not only allows for repeated and detailed examination, but it also allows the reader direct access to the data, which further eliminates the possibility of preconceptions in the analysis, as the findings are open to public scrutiny (Heritage and Atkinson, 1984).

Interactional Linguistics (IL) follows much of the methodological frameworks and ideology of CA but is an interdisciplinary, multi-perspective approach to the study of talk-in-interaction (Couper-Kuhlen and Selting, 2001) which is crucial to the analysis of the data at hand. The approach focuses on the way in which language is moulded by the context within which it used. The prior and subsequent sequential context of utterances or actions are considered essential to the analysis and under-
standing of interactions. This largely calls for a qualitative approach to data rather than a quantitative approach (Lindström, 2006). IL developed as linguistic studies moved away from primarily studying language in its isolated written form and as connections were drawn between linguistic structures and discourse function. These shifts, alongside sociologists beginning to advocate for the study of naturally occurring everyday conversation, resulted in the interdisciplinary approach of the study of language as a resource to accomplish a goal or complete a task within interaction (Couper-Kuhlen and Selting, 2001). An interactional study of sequences and organisations in natural speech, spans from audible minutiae such as attributable silence (Schegloff and Sacks 1973) as a pause within a speaker turn, which may not function as a transition-relevance place (see Sacks, Schegloff and Jefferson, 1974), to particles such as, “well” being a possible my side token or closing a topic and introducing new topics (see Heritage 2015), to visible nonaudible behaviour such as gesture.

The context allows for the differentiation of the different functions of silences and pauses and one-word tokens which otherwise would go overlooked. In regard to gesture, it was a topic of analysis in early CA papers, like Goodwin (1979:97) on gaze in “the interactive construction of…sentence[s]”. However, a large percentage of early CA recorded data were telephone calls. As such, it was difficult to be able to analyse non audible aspects of conversation because they were not accessible. As the shift to video data became more prominent, the focus on non-verbal features also became more prominent. Kress (2015) goes as far as to say:

“‘Language’, confidently assumed (in the ‘West’) as the guarantor of what is distinctively human, rational, essential for reflection, capable of expressing every aspect of human life, is being challenged in its hitherto central position by other means of making meaning, by
As Kress (2015) explains, the study around multimodality (MM) is just going on slightly over twenty years. The novelty of the term breeds a lack of consensus on the definition of what it encapsulates. However, there is somewhat of a consensus that Multimodality, not only refers to the phonological aspects of language, such as, tone and prosody, but also to visible aspects, such as, facial expressions, gaze, hand movements, body positioning and so on. This is by no means an exhaustive list and MM cannot be limited to only such practices. However, for the purposes of this paper this is the branch of multimodality that feeds into the analysis of the data at hand. By adding another dimension to the way in which language is approached, it allows for a better handle on the data and “what ‘the world in the frame’ is like” (Kress 2015:54).

However, the relation between audible and visible modes of interaction is not a simple one, as the boundaries and the temporal alignments are not always clear cut. There are studies which focus on interaction openings in public places (Mondada, 2009), as well as sequence closings (Broth et al. 2014). In these studies, the visible embodied resources to gain another person’s participation in an interaction or exit an interaction, are mobilised alongside audible resources—although they are often not simultaneous. Mondada (2009) analysed people approaching other people in the street to ask for directions. The first resource used to acquire the directions was often walking up to the other person to gain their participation. However, in some cases the person opted for a vocal attention getting device. Kendon (1980:208) suggests there is an “intimate relationship between speech and gesticulatory hand movements” which is further supported by studies done by Trevarthen (1977) and Ingram
who have observed this in children as young as three. Thus, the studies around gesture over
the years have observed co-speech gesture, as well as gesture which precedes speech and gesture
which follows speech. Goodwin (2018) states:

“talk and gesture mutually elaborate each other within...a larger sequence of action
and...an embodied participation framework constituted through mutual orientation be-
tween speaker and addressee” (Goodwin 2018:180)

This idea, of talk and gesture mutually elaborating each other, is one of the aspects MM studies al-
lows the investigation of. Mondada (2019) stresses that there is no hierarchy of resources that binds
social interaction. All resources have the potential to carry equal weight in the organisation of ac-
tions.

The main concerns of integrating MM study and IL into a CA approach are defining how these ver-
bal and non-verbal actions and interactions, which are taking place in parallel temporalities, con-
tribute to the organisation of social interaction. By integrating nonverbal resources, many aspects of
analysis, such as turn taking, become more complex. However, in the study of animation and co-
animation it is an integral factor. The main aspect of nonverbal communication that this paper will
focus on is gaze. Due to the nature and design of the data, which the next section will cover, the fo-
cus on gesture, aside from gaze, will not be as prominent.
3.3 Data

3.3.1 Description of the data and ethical issues

The data for this thesis was taken from the 90s Baby Show podcast. The weekly podcast is hosted by 3 British men of African descent who are often accompanied by a guest to discuss topical world issues and the way in which they navigate today’s society. This podcast was selected in particular due to the variety of English present within the podcast interactions. The podcast hosts can be observed to use *man* as a pronoun (see extracts 1.2, 5.1, 5.2 and 5.3). This is a prevalent feature of Multicultural London English (MLE)(Cheshire, 2013). MLE is a variation of English that has developed in the city of London, particularly amongst youth who come from diverse cultural backgrounds and acquire English as their second language (Cheshire, Kerswill, Fox, & Torgersen, 2011). The lack of a target model of English from the parents in ‘non-Anglo’ households contributes to this formation of a new variation of English (Cheshire, Kerswill, Fox, & Torgersen, 2011). Whilst MLE can be found in the ‘Anglo’ population of London also, it is used mainly by ‘non-Anglo’ speakers and ‘Anglo’ speakers with multi-ethnic friendship groups (Cheshire et al., 2011). Although, MLE has been studied through a sociolinguistic lens, comparatively, there appears to be a lack of representation of variations of English that are used by people of colour in CA studies. Therefore, whilst this paper does not focus on the features of MLE, the use of this podcast serves to bring diversity to the variations of English studied through a conversation analytic and interactional lens. The podcast is available in both audio and video formats and the collection in this thesis con-
sists of five video podcasts which are 173 mins of talk in total. The extracts are transcribed in accordance with Jeffersonian (2004) transcription conventions and multimodal transcription is done in accordance with Goodwin (1981) (see appendix I for a copy of the conventions). Furthermore, the podcasts are not scripted. However, there is an element of organisation in terms of the set up.

Each podcast is recorded in a studio and the participants are all sitting on a sofa and facing the camera which is positioned at the front of the room. The camera is set up on a tripod but there is a cameraman (who is referred to as VP in the data extracts that follow) who is manning the camera. Due to his role, he is not in front of the camera and is participating in the conversations audibly but not visibly. There are three to four participants in each podcast including the hosts themselves. The video and audio versions of these podcasts are available on multiple streaming platforms and were accessed using YouTube (see appendix II for the weblinks to the podcasts examined in this thesis).

However, when using data in the public domain there is still some consideration to be given to the speaker’s intentions with which they made the data public. With social media data in particular, tweets etc., the poster is likely to only have accounted for their followers seeing their posts. Therefore, despite it being in the public domain, the use of such data in academic work could still raise ethical issues. In terms of the podcast, the currency of YouTube is views; the more views a video receives the more revenue is accumulated for the video poster and therefore there is not an expectation of set viewership or audience in the same way there is with a personal Twitter account for example. However, the speakers in the podcast likely would not have accounted for their podcast to be subject to linguistic analysis. Therefore, the researcher contacted each host and guest in the se-
lection of podcasts used and obtained their consent. The participants were informed their conversations would be subject to linguistic analysis and they were able to ask any questions about the study at this point.

Using data that is available on YouTube aids in ensuring higher ecological validity compared to media available on television. Content on YouTube is not subject to the censoring which is present in television media of ‘indecent’ or ‘profane’ content amongst other restrictions because it is an online streaming platform. Another advantage to using YouTube podcasts is that the channel is run by the participants of the podcast themselves and so they are not subject to any contractual requirements. The emergent nature of the conversation in the podcast can be observed in the way in which the hosts and participants refer to what is taking place around them. The participants can be seen to make comments about the food or drinking being consumed, the clothing being worn, members of staff or family entering and leaving the studio and so on.

3.3.2 Limitations of the corpus

The presence of cameras and the awareness of the conversation being recorded can trigger observer’s paradox (Labov 1972). Two of the issues Labov tried to eliminate was the amount of attention speakers paid to their speech and the presence of the field researcher. The podcasts were not recorded with the awareness that they would be subject to linguistic analysis but rather for entertainment purposes. Therefore, the likelihood of the phenomenon being influenced or altered by the presence of the camera is low. However, despite how comfortable the speakers may be with the camera and cameraman, there is still an issue of some form of censorship. As previously stated, the
censorship is not as strict as it may be with content shown on television, but the speakers are still likely to be conscious of the reach of the podcast. Had the conversations they were having, not been recorded it is likely it would differ to the ones which are uploaded on YouTube. However, for the purposes of this thesis, censorship is not a factor that influences the analysis in this paper.

One obvious limitation is the layout of the studio. Due to the layout, the conversations are taking place with all participants facing towards the camera, which is positioned at the front of the room. Therefore, the participants are sitting side by side which could affect gestural aspects such as gaze. However, seating positions when analysing gaze has been taken into consideration in a multitude of settings (see Muirhead and Goldman, 1979; Thompson and Suzuki, 2014; Holler and Kendrick, 2015; Auer and Zima, 2021); the examination of gaze in this setting can be as analytically rich as gaze in other settings. There is also the issue of the cameraman (VP) not being visible. Whilst VP is a host of the podcast he does not appear on screen. This further hinders the analysis of gaze, as it difficult to ascertain his exact location. However, although his exact location is not evident, his general location can be approximated in the way participants and the other hosts gaze at what appears to be a single location beyond the camera. The approximation of VP’s location has been explicitly stated in the analysis and the possible unreliability of the observation has been made.

Nonetheless, it is seldom found that data is void of limitations. Despite the shortcomings, the podcasts still prove useful visible instances of interactions that are analytically rich. The lack of prior knowledge that the podcast would be subject to linguistic analysis, as well as the lack of censorship and topical suggestions all contribute to a high degree of naturally occurring speech.
4. Co-animations: acceptability and safe entry into the animation space

4.1 Introduction

In previous chapters, studies defining animation and co-animation have been reviewed and the properties which constitute these phenomena have been described (see chapter 2). This chapter focuses on examining instances of co/animations that were accepted by the first animator. Examples of the most frequent appreciation devices of co-animation sequences found by Cantarutti (2020) are laughter and agreement tokens such as “yeah”. By examining interactionally successful instances of co-animation, this chapter hopes to determine the factors which contribute to acceptable co-animated contributions.

4.2 Epistemic positioning

Co-animation, as previously defined, is the joint production of a voice in an animation space opened by a first speaker. This first speaker/animator, therefore, in a majority of cases, has more epistemic access to the animation space. As Heritage (2012) explains, epistemic status is on a gradient:
“we can consider relative epistemic access to a domain or territory of information as stratified between interactants such that they occupy different positions on an epistemic gradient (more knowledgeable \(K^+\) or less knowledgeable \(K^−\))” (Heritage, 2012:4)

However, epistemic access does not only refer to the knowledge state of a speaker in regard to knowing and unknowing. Issues such as degree of certainty, provenance, and directness of the knowledge are also taken into consideration (Stivers, Mondada, and Steensig, 2011). Directness refers to the way in which the knowledge was acquired—does the person have first-hand or second-hand knowledge. The directness of knowledge is also closely related to how certain a person may be of said knowledge (Stivers, Mondada, and Steensig, 2011). Kärkkäinen (2003, as cited in Stivers, Mondada, and Steensig, 2011:15) shows that people can flag their level of certainty of knowledge with lexical markers such as “probably”, “maybe”, “I think” and so on. Thus, speakers animating an event and producing utterances from a setting from which they do not have first-hand access, may place themselves in a position of low epistemic access on the gradient. Co-animators, in these instances, are often even lower on the gradient as they would be receiving second-hand knowledge from the first animator. In the case that co-animation is taking place in a group setting, as is the case of the extracts in this collection, the positioning of first animators and co-animators on the gradient becomes further complex. This is because the potential co-animators may not share the same position of epistemic access. For example, in the case that the first animator is recalling an event at which one of the participants in the group setting was also present, it would then follow that, that participant would be in higher position of epistemic access in comparison to the other po-
tential co-animators. Furthermore, the co-animator/s is often doing maintaining a voice, in an animation space that the first speaker has presented to them. The first animator, in a majority of cases found by Cantarutti (2020), is animating the voice of self (see figure 2.3). The speaker is animating their own speech, thoughts, actions, or views in a past or fictional time and place.

In the following example (4.1), T is discussing, during a podcast episode, an experience at school surrounding exams.

Extract 4.1 “where’s my marks”

→ 1  T: it was £competition in school lih£ like rah⁶ what did you get
→ 2  like >I wanna< I wanna ace this ↓one↓ (.) you’d give it in?
→ 3  (.) teacher’s takin ages to mark it and you’ll be hassling em
→ 4  like where’s [my marks *fam*? ]
→ 5  VP: [where’s my marks] yeh ha [hah ]=
→ 6  F: [fam::]
→ 7  VP: =[hah: ha: ]
→ 8  T: [coz I need to know what I got]

T is describing the competitive nature of exams at school and the urgency he and the other students felt about receiving their exam results back. T opens the animation space on line 1 with a first animation in the voice of self + Another. The voice of Another in this case, appears to be the collective voice of his school cohort.

---

⁶ ‘rah’ is a slang term which is used when something is unexpected, unbelievable, or excitable. This usage of rah can be found in Multi-cultural London English. See Cheshire et al. (2011) for more information on the features and emergence of MLE.

⁷ The word ‘fam’ is derived from the word ‘family’ and was originally used to refer to close ones who are not related by blood. The term has since been extended beyond close ones and friends and can be used to refer to any person and is often used to mark emphasis.
Lerner (1996a) states:

“In conversation, participants maintain special rights to speak about certain things, such as their own experiences and opinions. Sacks (1975) pointed out, "One is responsible for knowing some things on one's own behalf" (p. 72)...This distribution of knowledge rights is consequential for the organization of conversation.” (Lerner 1996a:317)

This distribution of knowledge rights is also consequential to the organisation of co-animation.

However, epistemic rights go beyond just knowledge rights. It is not only a case of rights to knowledge but also the rights to narrate, describe or assess the events taken place. Gill (1998, as cited in Heritage and Raymond, 2006) in a study of doctor patient interactions found that patients are often reluctant to diagnose the cause of their symptoms when speaking with a physician, despite having more knowledge of their specific symptoms and what they are experiencing. Thus, similarly it is important to note that the first animator, in this case T, has both the epistemic and knowledge rights to speak on his own behalf. The same right is not automatically awarded to the co-animators, VP and F, when they enter the animation space. These grounds to claiming knowledge are referred to as epistemic authority (Heritage and Raymond, 2005). Referring back to the study on patients and physician interactions, the patient may have more knowledge on their experience, and therefore more epistemic access, but the professional has the epistemic authority to claim the knowledge and diagnose. However, in this case, T is discussing a general school experience, it is not a setting or an experience which T himself or his peers has exclusive access to.
Furthermore, Cantarutti (2020) asserts that the first animator is able to do a range of different things when opening the animation space to create a safer environment for co-animation production. In the case of extract 4.1, T prior to opening the animation space on line 1, states “it was competition in school”. Noticeably, the statement is lacking detail; he does not specify which school he is referring to. “school” could refer to primary or secondary school with students ranging from age 5 to 16. He also does not give the name of the particular school he attended. Therefore, it appears T is anticipating shared knowledge between himself and the other participants. Although T is the first animator, the other participants have equal epistemic access to this experience as they appear to have shared knowledge. T also can be seen to levelling his epistemic rights and authority by presenting the experience in a general sense as opposed to talking about his own specific personal experience at a particular school at a particular given time. T also uses the pronoun “you” on lines 2 and 3 as opposed to I or we. This could be another way in which T is positioning himself as epistemically equal to his recipients; giving in the paper was not only something he was doing but also something they were doing.

However, another aspect that should be taken into consideration, when examining T’s use of pronouns, is the setting of the talk. The relational stance, which the use of “you” is indicative of, may extend further than the primary setting in which it was produced. Burger (as cited in Bubel, 2008) presented a model in 1984 that described the way talk on television and radio is structured. He argued that participants, in these settings, are not only communicating with the people present, but also with the audience behind the screen. He described two communication circles: an inner circle where the talk is physically taking place and an outer circle where the talk is being broadcasted to
the audience members. An ethnographic paper by Duranti (1986) argues that the audience co-constructs meaning. He talks about how meaning and interpretation lie with the addressee and not the speaker. A study by Bubel (2008) applies both of these theories directly to film audiences who would be considered in the outer circle. The inner circle in film productions would be the directors, producers, co-actors, and other production staff present. Thus, the inner circle would be concerned with delivery of the dialogue and script and not interpretation. However, films are meticulously scripted with the audience in mind—the outer circle. For example, it is not often there is overlap or interruption taking place in film dialogue—unless it serves a specific purpose in the film. The use of these linguistic features, which are found in naturally occurring speech, is carefully scripted. However, television, radio, film, and podcasts all differ in the level of scripting. Films and television programmes are, majority of the time, very strictly scripted. The podcast the extract (1.1) above was taken from is not a scripted podcast. Therefore, there was not any foreplanned structure of talk designed for the audience. This is not to say that there is not an awareness of the audience. Podcasts, similar to other forms of media are timed, recorded and edited. The audience is likely to be taken into consideration when deciding on the length of the podcast and whether or not to edit out content that could be deemed as offensive etc. Therefore, the pronoun ‘you’ may not only be extended to the inner circle audience, but also the outer circle audience. The podcast must maintain some form of entertainment and relatability in order to keep viewers. By speaking about topics to which the outer circle has epistemic access, such as, experiences at school, it gives the audience the ability to relate. T could be extending the access to this experience to, not only, the participants in the primary setting, but to the secondary audience behind the screen through his choice of pronoun.
Furthermore, the appreciation for the co-animation in a group setting like extract 4.1 is not limited to first animator response. The visual representations, extract 4.1 and 4.2, below observe the ways in which the participants are showing appreciation for the co-animation produced by VP.

Extract 4.1 (visual representation) “where’s my marks”

The image displayed behind the participants is the result of a green screen.

From left to right we have F, H and T. VP it appears is sat to the left of the camera (the right of F, H and T).

It is not certain that the participants are gazing towards VP as his exact location cannot be confirmed.

---

8 The image displayed behind the participants is the result of a green screen.

9 From left to right we have F, H and T. VP it appears is sat to the left of the camera (the right of F, H and T).

10 It is not certain that the participants are gazing towards VP as his exact location cannot be confirmed.
Kendon (1967) and Duncan (1975) have argued that gaze can function as a way to apportion the interactional floor. Goodwin (1981) also observes that gaze could be one of the means available to a recipient to display their listenership. In extract 4.1 above both F and H are gazing at T who has the interactional floor. The brief moment of gaze aversion taking place during T’s utterance on line 2, first initiated by H and shortly followed by F, was not towards another recipient but downwards.

This could be a home position for the eyes which is often observed to be downwards, towards hands or legs (Rossano, Brown & Levinson, 2009). Home positions for the eyes are observed as unrelated to the apportionment of the floor and managing turn-taking (Rossano, Brown & Levinson, 2009).

Both gazes return to T who still has the floor and continues his narrative and re-enters the animation space on line 4. However, T’s gaze is during the entirety of the interaction is oriented towards the camera, further indicating his awareness of the outer circle audience.

As VP comes in with a pre-completion co-animation, F begins to orient his gaze towards VP.

Extract 4.2 (visual representation) “where’s my marks”

Refer to footnote 10
As can be seen in extract 4.2 above, all 3 participants are gazing towards where VP is assumed to be sitting during his utterance on lines 5. Whilst the camera is directly in front of the participants, they can be observed to be looking past the camera to the left where VP, the cameraman, is assumed to be sat. Bolden (2013) in a study analysing repair sequences found gaze can be used to validate, accept or reject repair solutions visibly. A similar occurrence can be observed in extract 4.2. F, H, and T all orient their gaze towards VP as he is co-animating thus possibly showing appreciation for his co-animation. Appreciation can be seen in not only the apportionment of the interactional floor but also the smiles, during VP’s laughter, from both H and F which can be seen above. Although, the appreciation through smiling does not come from T, the first animator, himself-only the other participants. Also, notably F does not orient his gaze towards VP until the first animator, T, has completed his utterance on line 4. Rather than an epistemic issue, which is unlikely to be the case, because as stated previously, T frames the animation to level the epistemic positioning of the co-participants and himself, it appears that F is unsure of who will be the prevailing speaker in the overlapped utterance. Zima, Weiß and Brone (2019) observed in a study on gaze and overlap resolution in triadic interaction, that the third participant often orients their gaze towards the withdrawing speaker as opposed to the prevailing speaker until the withdrawing speaker has withdrawn. This

12 Conclusions drawn in the analysis are based on the approximate location of VP and whilst the gaze of the participants is highly indicative of his location it is still deemed unreliable.
could be the case with F. VP could be perceived as the prevailing speaker in this overlap sequence as T does not reclaim the floor and VP produces laughter after his overlapped utterance therefore keeping the floor. H can be observed to be orienting towards VP after T finishes the word “marks”. The word “fan” which follows is produced quieter than the surrounding talk. It may be the case that H recognises T to be the withdrawing speaker at this point and therefore begins to orient her gaze towards VP and show appreciation for his co-animation with her gaze and smile. Therefore although, it appears gaze could be one way in which participants are able display acceptance of co-animated utterances this is not readily apparent from the instances of co-animation examined in this chapter. Whilst this paper provides an initial analysis, the relationship between gaze and the acceptance of co-animations, requires further exploration.

A further instance of the switching of pronouns and an instance of appreciation from the first animator can be seen in extract 4.2 found in Cantarutti (2020). Cantarutti’s (2020) analysis focuses on the structuring of the stages of the complaint sequence through animation (see chapter 2), whilst the following analysis establishes the way in which pronoun selection can be used to facilitate safe entry into the animation space. Jesse and Fiona are discussing problems Jesse is facing with the paper he is writing.

Extract 4.2 “page limit” (in Cantarutti, 2020:342-343)\(^\text{13}\)

09 JES (cause) one of the ‘rerally annoying things:=
10 =with the paper I’m writing ‘now, {{sniff}} (0.2) {{click}}
11 i:s (0.6) it makes ‘so little ‘sense,=

\(^{13}\) This extract has been subject to alterations. For the original and complete extract see Cantarutti (2020:342-345).
Jesse is describing to Fiona the difficulty he is having providing adequate background in his paper because of the limited word count. On line 15, Jesse uses the pronoun “you”. However, Jesse has been speaking about problems with his own paper up until this point. And so, it is unlikely that this “you” is directed at Fiona and that he is saying she is “purely wasting words”. He continues the use of the pronoun “you” on line 18, saying that he would be “re'treading old ground” in reexplaining things written in his first paper. When orienting to his first paper on line 12 Jesse also uses the pronoun “you”. Again, it is unlikely that Jesse is orienting to Fiona but rather the readers of his paper. Fiona responds by backchannelling (Yngve, 1970), providing minimal responses such as “m::m” and “hm” on lines 13, 16 and 19. This further reinforces that Jesse’s utterances are not orienting towards Fiona as she does not treat them in this way. Thus, the use of “you” here could be seen to be doing relational work, giving Fiona epistemic access and shared authority to his experiences and that of the readers. This levelling of rights could help to create safer entry to the animation space for the co-animator. After the basic explanation of his problem, Jesse goes into a specific criticism he got on his paper and begins an animation sequence.
Extract 4.2 “page limit” (continued)\textsuperscript{14} (in Cantarutti, 2020:344-345)

33 JES ((lip smack))that was one of the `criticisms that i `got. (0.7)
34 (through/from) `adam (who was just) `s:aying (0.7)
→ 35 `you do that too often
→ 36 `everything needs to be? = like self con`ta:ined; °hh (0.9)
37 and `then at the sa:m? `same time;
→ 38 `add all this: ; = but `don’t take anything `out? (0.2)
39 FIO m:`hm:m
→ 40 JES huh `what do you `want me to `do (0.7)
41 FIO [m:m]
42 JES ?huh ?huh ?huh ((sniff)) (1.63)
→ 43 FIO do `yu under`stand how `page limits `work?
44 JES nhh° nhh°

Jesse opens up the animation space on line 35 with a first animation in the voice of another (see chapter 2). He carries on the animation sequence to line 38. The animated voice up until line 38 is the voice of “adam”. At this point, the subject of the complaint has been made evident. The complaint is about an absent third party and therefore an indirect complaint (Edwards, 2005). Thus, Fiona entering the animation space at line 43 could be due to it being safe not only positionally, but morally. As recipients of complaint sequences are presented with information to be able to perceive the rightness or wrongness of the events or actions taking place; they are able to make moral evaluations (Drew 1998). Thus, evaluating or commenting on an event before Jesse has provided the utterances of Adam and his own stance towards them, could put Fiona at risk of making a moral transgression.

\textsuperscript{14} This is not a direct continuation from extract 4.2-lines 20-32 have been omitted as they are nonessential to the analysis at hand. For the original and complete extract see Cantarutti (2020:342-345).
In “The Territories of the Self” Goffman (1971) discusses a range of different territorial preserves people have and the way in which people defend or protect them, and also the ways in which other people can violate them. One of these territorial preserves Goffman calls the information preserve.

He defines this as:

“The set of facts about himself to which an individual expects to control access while in the presence of others. There are several varieties of information preserve...There is the content of the claimant's mind, control over which is threatened when queries are made that he sees as intrusive, nosy, untactful” (Goffman 1971:38-39).

Although, Goffman did not include knowledge in his definition a “control over rights to information is evidently the object of linguistic interactional management (Kamio, 1997; Maynard and Zimmerman, 1984)” (Heritage and Raymond, 2005:34). Thus, the co-animator/s is faced with issues of epistemic and territorial access when entering the animation space. Goffman (1971) also describes territorial offences as the act of intrusion into a territory which the person has no right to access. Thus, once Jesse produces his own displaced voice on line 40, it is then that Fiona enters the animation space and co-animates in the voice of Jesse. If Fiona were to enter the animation space earlier and produce the voice of Adam, because she lacks epistemic access to the conversation Jesse is describing, she could also risk making a moral transgression if she were to make an incorrect inference. Animating the voice of an absent third party, in a setting where the animator lacks epistemic access to the events being described, could be an interactional risk. However, appreciation
from Jesse can be found on line 44 where he produces laughter particles. Thus, it seems the interac-
tional risk was avoided. As stated above, this could be largely due to where in the complaint se-
quence Fiona entered the animation space. This is further discussed in the following subchapter.

This section has examined the role of epistemic positioning in safe entry into the animation space.

It can be seen that first animators design their animations in a range of different ways to level the
epistemic rights, access, and authority to their animated content. One way in which first animators
are observed to be doing so, is by animating content with which the other participants have shared
knowledge. Designing their animations to refer to general experiences and framing them using non-
specific information, first animators are able to level epistemic access, rights and authority to their
animations. Another way in which they are able to do so is by taking a relational stance with their
choice of pronouns when presenting animated content. By using pronouns such as “you”, the first
animator is able to extend the experience they are animating beyond the participants present in the
initial interaction and provide epistemic rights and authority to their recipients. This section has
also made initial investigations into the relationship between the design of animations for a second-
ary audience and the relationship between gaze and appreciation. These areas whilst addressed in
this analysis, require further exploration in order to be able to achieve a deeper understanding.
4.3 Projectability

A large contributory factor to safe entry into the animation space could also be projectability. Duncan (1974:179) suggests that there is a “between-unit auditor back channel” which “indicates that the auditor is following the speaker’s message as it’s developing”. Jesse’s description of his problem is granular in the way it is constructed, and so Fiona could be producing backchannels that suggest she is following Jesse’s complaint as it is unfolding. Moreover, in line with Cantarutti’s (2020:356-358) stages of troubles-tellings and complaint sequences (see chapter 2), the backchannel responses from Fiona could be categorised as stage 2. The backchannels could be perceived “as tokens that act as “go ahead” because they treat the matter as recognisable but also orient to this as not being the actual core of the activity, as more is projected” (Cantarutti, 2020:356). Likewise, Schegloff (1982) expands on the notion of back channelling by suggesting that not only is the listener following the message, but they are taking the stance that the speaker should extend their turn and keep the floor, through their back channels. He suggests this is evident through the listeners passing on the opportunity of producing a full turn of talk. Therefore, this indicates the speaker should continue and when tokens such as *uh huh*, *mhm* etc. are found in this position they are termed *continuers*. This dual function of both understanding and stance taking is evident in extracts 4.2 from Cantarutti below:

Extract 4.2 “page limit” (in Cantarutti, 2020:342-343)

(1)  11 i:s (0.6) it makes `so little `sense,=
     12 if you `haven’t read the ´first one?
→ 13 FIO ↑"m::m (0.3)
By line 43, Jesse has presented his problem with the word count and provided both the stance of Adam, and his own thoughts through his animation on line 40-stage 5 in Cantarutti’s (2020) model. Until this point, the continuers produced by Fiona illustrate the understanding that the problem has not been explained in its entirety yet and she is taking the stance that he should continue. Once the problem and stance have been outlined by Jesse, she is able to safely enter the animation space at stage 6 and “display…a co-complaint or a co-victim role” (Cantarutti, 2020:358). The risk she faces of getting it wrong has been minimised when she produces a co-animation on line 43.

This is also evident in extract 4.1 where VP on line 5 comes in with a pre-completion co-animation.

Extract 4.1 “where’s my marks”

3 (. ) teacher’s takin ages to mark it and you’ll be hassling em
→ 4 like where’s [my marks *fam* ]
His production comes in before the projection of a transition relevance place. Pre-completion co-animations are found in pre-emptive positions, either in overlap as a form of choral production or in anticipatory completion (Cantarutti, 2020). Producing a co-animation before the first animator has completed their utterance runs the risk of getting it wrong. There is a risk of misjudging stance or miscalculating the direction of the talk. Therefore, they could be deemed interactionally unsafe. Moreover, Lerner (2004a), in his study on collaborative turn sequences, identifies, what he calls, *opportunity spaces*. These are breaks in the progressivity of the turn where the opportunity arises to produce a collaborative completion. However, as mentioned above, VP does not enter the animation space at the pause on line 2, or the pause on line 3. He also does not come in after ‘like’ but instead comes in after ‘where’s’. As has been highlighted by previous studies on anticipatory/collaborative completions (see Hopper and Thompson, 2008; Liddicoat, 2004; Lerner 2002, 1996b; Ford, Fox, and Thompson, 1996), projectability is one of the biggest contributors to successful collaborative turn completions. Thus, line 5, after the T produces ‘where’s’ and is coming to the end of his turn, is arguably a safer place to gage projectability and co-animate. The point of entry into the animation space may contribute to minimising the risk of getting it wrong and therefore producing a more acceptable co-animation.
4.4 Formulations and reformulations

The previous subsection has outlined how a co-animator can choose the point of entry into the animation space to minimise the risk of getting it wrong. This next section focuses on the way co-animators can design their utterances to allow for more acceptable co-animations. One way in which co-animators can be seen doing so in the data set, is through formulations.

In extract 4.2 below Fiona uses formulation in response to Jessie’s complaint sequence:

Extract 4.2 “page limit” (continued) (in Cantarutti, 2020:344-345)

33 JES ((lip smack))that was one of the `criticisms that i `got. (0.7)
34 (through/from) `adam (who was just) ~:saying (0.7)
35 `you do that too often
36 `everything needs to be? = like self con`ta:ined; °hh (0.9)
37 and `then at the sa:m` `same time;
38 `add all this: ; = but `don’t take anything `out? (0.2)
39 FIO m:hm:m
40 JES huh `what do you `want me to [`do ?h° `ad?h°am?]
41 FIO [m:m]
42 JES ?huh ?huh ?huh ((sniff)) (1.63)
→ 43 FIO do `you under`stand how `page limits `work?
44 JES nhh° nhh°

The origins of the term formulation can be found in an observation made by Garfinkel and Sacks (1970). They suggest that participants can explicitly outline where they are in the conversation by formulating the talk so far. Formulating is the act of making a meta-pragmatic observation on the talk (Antaki 2008). However, the term formulation was further developed by Heritage and Watson (1979, 1980) to become a so what you are saying is observation rather than a this is where we are
in the talk observation (Antaki 2008). This can be seen in the formulation on line 43. Fiona is outlining what is being implied by Jesse in his previous utterances, but she is also confirming her understanding that the person overseeing Jesse’s paper is expecting more from Jesse than the page limit allows. Furthermore, Heritage and Watson (1979:124-125) observed that formulations possess different characteristics depending on whether they were produced by a “news deliverer” or “news recipient”. The subset of formulations the study focused on was specifically formulations produced by the “news recipient”, in this case Fiona. ‘News’ in their definition is understood as being complaints, stories, reports, and announcements, the news in this case is a complaint sequence. They also identify three features of formulations: deletion, preservation, and transformation. The formulations deleted some aspects of the prior talk whilst preserving and reproducing other parts to transform the talk into an utterance, or sequence of utterances, which surmise what has been said so far or identify an apparent implication. Fiona can be observed to be deleting the details of the criticism Jessie presents; preserving his questioning of Adam and transforming his question to directly address the implication that Adam is not taking into consideration the constraints of the page limit. A majority of studies on formulations in CA have adopted this definition of formulation (Drew, 2003; Hutchby, 2005; Antaki, 2008; Bolden, 2010; Almeida and Drew, 2020).

However, there are slight differences in the ways in which the formulations examined by Heritage and Watson (1979) and the formulations found in the data at hand function in the talk. The examples of formulations examined by Heritage and Watson (1979) come from a range of different interactional settings. These include radio interviews, crisis intervention call centres, and group therapy.
sessions. Other studies examining formulation have studied interactions in psychotherapy (see Antaki, Barnes and Leudar, 2005; Antaki, 2008), medical consultations (Ostermann and da Silva, 2009, as cited in Almeida and Drew, 2020), phone in radio programmes (Hutchby, 1996), police investigation interviews (Almeida and Drew, 2020) and Animations (Cantarutti, 2020). Cantarutti (2020) examines the function of formulation in animation in regard to teasing and complaint sequences. As a result of the diverse usage of the term ‘formulation’, the instances in the data at hand and the instances examined in some of the institutional settings above are not identical. This is largely because what formulating the talk is doing interactionally is contrasting. In a police interview or medical consultation, the formulations function to question the validity of what the professional has understood from the talk and often request confirmation (Almeida and Drew, 2020; Ostermann and da Silva, 2009). However, the formulations in the animation sequences that will be examined in this section do not share this characteristic. Police officials, doctors and psychotherapists possess epistemic rights and authority over their interlocutor due to the nature of their position and role. Thus, the nature of requesting confirmation, that characterises formulation in the settings in which the previous research was undertaken, is not identical to the nature of requesting in the animation sequences in this collection.

For instance, in extract 4.2 Fiona’s formulation can be seen as performing a dual function. Brown and Levison (1987:226 as cited in Clift, 1999:524) when discussing irony state that “the mechanism by which irony works is that the utterance, if taken literally, is obviously grossly inappropriate to the situation”. Fiona’s utterance on line 43 if taken in a literal sense would be grossly inappropriate in the situation Jesse was recollecting. Whilst the person overseeing Jesse’s paper, it seems is not
understanding of the constraints of the page limit in regard to Jesse’s work, it would be logical to assume that a person placed in charge of supervising papers would have a firm understanding of page limits. Thus, whilst Fiona’s formulation is drawing upon the information Jesse has made available to her through the talk so far it is not doing so in the way in which formulations are observed to be doing in the institutional settings discussed above. However, despite the differences in the usage of the term, the definition in its most basic form is suited for the purposes of this paper and so this section will examine formulations provide implications, abstractions or generalisations produced by the news recipient, or in this case the co-animator, and the ways in which they provide safe entry into the animation space.

The act of co-animating can be potentially risky in terms of alignment or, as previously outlined, epistemic entitlement. Bolden (2010) found and-prefaced formulations to be articulating implications in the previous utterances. These formulations were structured to receive confirmation or disconfirmation from the participant. In other words, they were dually showing understanding of the talk so far, whilst ensuring they have understood correctly and requesting further information or confirmation. The use of formulations when co-animating can be a way of managing the lack of access to the original setting which is being described in the animation space. Thus, formulations reduce the potential risk of getting it wrong. An instance in which the implicative nature of formulation in co-animations can be seen in the continued sequence of extract 4.1 below:

Extract 4.1 “where’s my marks” (continued)

1 T: it was £competition in school lihf like rah what did you get
2 like >I wanna< I wanna ace this †one† (.) you’d give it in?
3 (. ) teacher’s takin ages to mark it and you’ll be hassling em
4 like where’s [my mark ]
5 VP: [where’s my marks] yeh ha [hah ]=
6 F: [fam::]
7 VP: =[hah: ha: ]
8 T: [coz I need to know what I got]
9 H: yep::
10 F: I’m tryna out do him °fam°
11 VP: [heh heh]
12 T: [yeheheh]

The formulation on line 10 is articulating a possible implication made by T. On line 1 T states that ‘it was fcompetition in school’ when taking exams, followed by the urgency he felt to receive his mark. F formulates the implication that the T’s urgency for receiving his mark is fuelled by this competition and, more specifically, the want to outdo his classmates through the co-animation on line 10. This is confirmed by T in line 12 with the agreement token ‘yeh’ and the laughter particles which follow.

Furthermore, extract 4.3 below, contains an instance of formulation, similar to the examples seen above, but also of reformulation in the animation space. M is discussing a conversation he had regarding the controversy around an item of Gucci clothing having racially prejudice undertones.

Extract 4.3 “today I got time”

1 M: so (. ) I was speakin to someone the other day and they said ah (. )
2 hhh >you’ve seen what gucci is doing (. ) blah bla blah blah< would
3 you wear gucci? I was like >↑yeah↑ if I go into a top< if I go
4 into a shop and I see like ah (. ) coz I want I wanted a
5 T: th-th the [side bag]
6 M: [the side] thing
7 VP: [yeah yeah ]
8 M: [>I think that’s wh<] I want a side bag >I was like yeah id get it
9 she’s like oh you’re a coon<
10 T: Ha ha ha [she called you a [coon ]]
11 VP: [nah nah nah ]
12 M: [>so I] was like< <okay cool> I wasn’t gonna speak about this but now I’m gonna speak about it
13 VP: >mmm mmm mmm mmm< now lemme talk [about it ]
14 M: [so I was like] T: today I got time heh hehh
15 VP: and tod[ay heh heh heh]
16 M: [so I was like ]
17 F: [uh huh huh [huh huhhh ]]
18 T: [uh huh huh [huh huh huh]]
19 M: [okay so (.)]when you buy your food is it all fairtrade?

T enters the animation space on line 16 with a formulation. M on line 13 explains that he was not going to address the issue of racism in the fashion industry in regard to an event which took place with a designer brand. However, he follows this by saying that he was now going to address it after the statement made by his interlocutor. T formulates the implication that M has the time to address the issue and also that it would be time consuming. The issue itself is a delicate and controversial one, as conversations around race tend to require handling with sensitivity in most cases. Although, T is not the first participant to enter the animation space. VP enters on line 14 with a reformulation.

---

15 The term ‘coon’ stems from the oppression and racial discrimination of black people during slavery. The term was originally thought to have been used for black slaves who were forced to entertain their slave masters. The word has evolved and when used in this context it is referring to a black person who is ignorant to the racial discrimination the black community face and willingly partake in practices which reflect self-hatred or ‘selling out’ behaviour.
Reformulation has been the focus of, mainly, discourse semantic studies with a particular interest in how they are used in institutional talk (Gülich & Kotschi, 1995; Caipuscio, 2003; Cuenca, 2003; Murillo, 2012). Reformulations are suggested to be paraphrastic in order to reword specialised terminology so that it may be understood by laypeople (Mortureux 1982 as cited in Cuenca and Bach, 2007). Thus, this definition of reformulation suggests that the reformulated utterance does not abstract from the original utterance. They are proposed to have a “single coherence relation” (Blake-more 2007:312) with the previous utterance. It is this definition of reformulation which will be taken forward in this paper. Whilst formulations provide implications, abstractions or generalisations, reformulations provide a paraphrased or reworded version of the prior utterance.

Furthermore, reformulations have a degree of repetitiveness to them. Repetition is suggested to function as a receipt of the talk (Greer et al., 2009). The repetition of sections of the prior utterances can mark recipiency and understanding. Through reformulating the previous utterance, or utterances, and producing a paraphrased version, the co-animators are able to display listening and understanding. Moreover, similar to the way in which repetition can function as a marker of agreement and alignment (Pomerantz, 1984; Schegloff, 1996a), reformulations can be used by co-animators in similar ways. However, an identical repetition could index epistemic primacy (Clift, 2006, as cited in Cantarutti, 2020). Thus, a paraphrased version of the prior utterance is more suited to the act of co-animation. VP on line 14 produces a reformulation which does not abstract from the original utterance. Both “but now I’m gonna speak about it” and “now lemme talk [about it]” have a single coherence relation. Thus, although the implicative nature of formulations, which can be seen in the instances analysed so far, provide a level of safety when entering the animation
space, reformulations are comparatively safer. Formulations do refer back to the previous talk and so the co-animators are able to draw upon the knowledge they have received so far when entering the animation space. However, implications are, as is given away by the name, not made explicit in the talk. Therefore, there still may be a risk of getting it wrong. Whereas reformulations are using explicit knowledge that can be taken from the talk and so are interactionally safer.

The appreciation for the formulated and reformulated co-animated utterances can be seen through the laughter exhibited by T, F and VP. Whilst the first animator, M, does not laugh he produces a half smile, and the gaze direction is oriented towards M throughout the ongoing laughter. This can be seen in extract 4.3 below:

Extract 4.3 (visual representation) “today I got time”

12  M: [>so I] was like< <okay cool> I wasn’t
     F: M__________________________
     T: M__________________________

13  gonna speak about this but now I’m gonna speak about it
     F: M__________________________
     T: M__________________________

14  VP: >mmm mmm mmm mmm< now lemme talk [about it ]
     F: M__________________________
During M’s animation on lines 12 and 13 M receives steady gaze from both the participants in front of the camera. T begins to orient his gaze towards, where VP is assumed to be sat, midway through his reformulation and continues throughout his formulation\textsuperscript{17}. However, F’s gaze remains on M, the first animator, throughout this interaction and T is quick to return his gaze to M once he has completed his formulation. VP also does not complete his utterance on line 17 which could be due to M, the first animator, also attempting to take the floor. The gaze of both F and T are oriented towards M during this time, apportioning the interactional floor to him.

A further example of reformulation can be found, in extract 4.4, when M is discussing an encounter his friend had whilst shopping with his mother:

Extract 4.4 “mum we need salt”

1  M: he said that um in the next week he was in lidl wiv his mum
2       texting the girl

\textsuperscript{16} Refer to footnote 10
\textsuperscript{17} Refer to footnote 12
M on lines 10 through 13 produces an animation in what appears to be the voice of another-his friend. M’s friend says to his mother “mum... we need salt” as a way of navigating her towards the aisle where the girl that he is texting is. F reformulates M’s animation on line 17 by saying “oh mum did we get salt yet”. He paraphrases M’s animation without abstracting from the original utterance. The format of the utterance is shifted from a declarative statement to an interrogative, but the single coherence relation remains. As previously stated, co-animationing the voice of an absent third party, with limited access to the original setting in which the interaction took place, poses the

---

18 Refer to footnote 1
risk of making a moral or interactional transgression. However, reformulating an utterance produced by the first animator minimises this risk. As is evident on line 18 where the reformulation is appreciated by the first animator using laughter.

This chapter has examined instances of co-animation that were accepted and appreciated—both by first animators and by the other participants present in the interaction. Appreciation devices such as agreement tokens, smiling, laughter and gaze have been explored as ways in which the participants accept and appreciate co-animated utterances. The environment in which they are produced has been determined through the examination of epistemic positioning and projectability, combined with the paraphrastic or implicative nature of certain co-animation structures, such as formulation and reformulation, and how they contribute to safe entry into the animation space. First animators have been observed to be levelling the epistemic positioning between themselves and the co-animated through the design of their first animated turn. Whilst the co-animate are able to ensure safe entry into the animation space through their point of entry. Co-animators are also able to increase the likelihood of the acceptance of their co-animated utterances through paraphrasing the animated content the first animator has provided using a reformulation. The use of previously presented information may minimise the likelihood of a miscalculation or putting words into the first animators’ mouth. Formulations are suggested to be interactionally riskier than reformulations due to the nature of participants suggesting implied information. However, the statistical use of reformulations versus that of formulations would be beneficial in providing an insight into the way in which each of these structures are used and if there is a preference for one over the other. This is an area that requires further investigation and development.
The upcoming chapter will examine instances of co-animation which were not accepted by the first animators and other participants present in the interaction. The characteristics of these co-animated utterances will be examined in contrast to the characteristics of the acceptable co-animations examined in this chapter thus far. Similarities and differences will be drawn between the co-animated utterances to determine which factors contribute to acceptability.
5. Co-animations: acceptability and disalignment

5.1 Introduction

As established in the previous chapter, the epistemic positioning of co-animators, the projectability of the animation sequence and lexical framing of co-animated utterances are integral to ensuring safe entry into the animation space. This chapter focuses on examining instances of disalignment in co-animated sequences and the environment in which they are recurrent. Factors observed in chapter 4, that contribute to acceptability, will be drawn upon in this chapter also to determine the characteristics of the environment in which miscalculations and disalignments are present.

5.2 Epistemic positioning

As discussed in chapter 4, Heritage (2012) proposes a gradient for epistemic status, positioning participants by their access to knowledge. Although, epistemic access also refers to the directness of knowledge, certainty, and provenance (Stivers, Mondada, and Steensig, 2011). The participant who opens the animation space, the first animator, more often than not, is the participant with the most epistemic access, right and authority over the animated content. However, it is possible for a first animator to open an animation space recollecting an event in which the other participants were also present and therefore have equal access to the event. If, for example, three friends were engaged in a conversation about a party two of them had attended and the specific event being recollected was of one of the friends having a wardrobe malfunction, theoretically the participant who was not at the event would be at the lowest end of the epistemic gradient. Whilst the two participants who
were at the event would have somewhat equal access, the participant who experienced the wardrobe malfunction may have more direct knowledge and epistemic rights and authority, as it is their first-hand experience. The voice in which animated utterances are produced also bears strongly on epistemic positioning of the participants. As covered in chapter 4, first animators animating in the voice of self, places them higher on the gradient of epistemic access, as the utterance being animated is coming directly from themselves and they also possess the epistemic rights and authority to reproduce the utterance. In the following example, the first animator animates in the voices of both self+another (see figure 2.3). T is discussing a project he carried out in design technology class at school:

Extract 5.1 “oh that’s nice”

1 T: THEY MADE [me: make ]=
2 F: [ oh my god]
3 T: =an egg holder like the fridge don’t have them
4 F: [fam:: ]
5 H: [[ah hah] hah hah ha]
6 VP: [ah huh huh huh huh huh huh]
7 T: [man’s got the] holes and everything in
8 [the fridge ]
9 H: [but that’s the] easiest thing [to make: ]
10 T: [I’m bringin] mum look what I
11 [brought an [£egg holder£ahah]
12 F: [ehhhhh huh ]
13 VP: [ahahahah ] ^oh that’s nice^
14 T: >that can only hold one egg?< There’s
15 [five of us in this house]
16 VP: [eh hah haha hhhhh ]
T opens up the animation space on line 10 with an animation in the voice of *self*. The utterance in its original context was directed at T’s mother, as can be seen from the term of address. VP on line 13 produces a co-animated utterance which appears to be in the voice of T’s mother, praising the egg holder. VP appears to be depicting feigned praise that may be offered to a child by their parent regardless of the outcome of the project itself. However, this was not the case as can be seen through the animation in the voice of T’s mother on line 14, expressing her criticism of the egg holder T had made. She is presented as critiquing the lack of use for an egg holder “that can only hold one egg”. A critique which is foreshadowed in T’s own criticism on line’s 1 and 3. He states he was made to make egg holders when the fridge already has them and thus it is implicative of the redundancy of his school project. T’s animation sequentially deletes VP’s utterance, leaving the co-animation unaddressed.

The co-animation is produced in a higher pitch to present a female voice and fitted in response to T’s first animation but, despite the utterance being a possible response from T’s mother, it is not the one she gave. This miscalculation could be due to VP animating a voice which T has yet to bring to the animation space and therefore has not provided his participants with epistemic access to. VP is in a K-position (Heritage 2012)-the less knowing position. Thus, to produce a co-animation in a voice which has not yet been animated by the first animator-the person with the highest epistemic access to his own experience-is likely to increase the chance of a miscalculation and getting it wrong.

This can be seen in extract 4.2 (see chapter 4) where Fiona despite, having access to both the voices of *self* and another, co-animates in the voice of *self* presented. This is likely because the epistemic
access to the voice of another is not equal amongst the participants, despite them having been brought into the animation space. The first animator, in this case Jesse, is in a position of having more epistemic access and authority as it is an interaction he participated in and a complaint sequence he initiated. Thus, although the epistemic positioning is somewhat levelled by Jesse bringing both the voices of self and another into the animation space, the first animator and co-animator are still on unequal standing. That having been said, had Fiona animated in the voice of another the acceptability of her co-animation possibly would have been more likely than VP’s due to Fiona being given epistemic access by Jesse. The positioning of her entry into the animation space could have been deemed interactionally safer.

Sequential deletion of disaligned co-animated utterances in regard to entry into the animation space is further analysed in the following subchapter.

5.3 Projectability

As outlined in chapter 4, projectability could be a large contributory factor of safe entry in an animation space or the lack of safe entry. At which point a co-animator enters the animation space could contribute to how interactionally safe his utterance may be. Extract 5.2 below exemplifies an instance of disalignment due to a miscalculation in the direction of the talk:

Extract 5.2 “where you going”

1 T: I almost felt like white people don’t wanna talk to me we-we
2 haven’t got anything in common
3 A: mmm ye[ah: ]=
4 VP: [yeah]
T is discussing a fictional scenario in which he is addressing the differences in the way he responds to a joke versus the way that his ‘white’ friend may respond. He opens the animation space on line 10 with the voice of self, directed towards his fictional friend. He then goes on to animate in the voice of another. On line 16 T imitates laughter that could be interpreted to be that of his fictional white friend who is just giving him “shoulders”. VP also animates the voice of another with a pre-completion co-animation (Cantarutti, 2020). Both T and VP produce laughter in the voice of another in chorus, which is then followed by a co-animation from F, also in the voice of another. Unlike extract 5.1, the first animator has brought the voice of another into the animation space. The animation sequence in this case is also fictional and therefore the constraints of epistemic access...
differ. The participants are animating from a point of more symmetrical epistemic access, as the animation space is being maintained through joint imagination (Stukenbrock 2014, as cited in Canta-rutti, 2020). However, despite the animations occurring in a joint fictional animation space, T is still in the position of first animator and is the one who presented the initial fictional scene. Thus, when F’s co-animation is sequentially deleted and T returns to animate in the voice of self, on line 18, it could be interpreted that F miscalculated the direction of the talk. It becomes more apparent from T’s utterance on lines 18-21 that his issue with the fictional interaction is the disproportionate response to his joke and not a possible lack of understanding, like F suggests through his co-anima-
tion.

The production of a co-animation when the first animator has yet to establish the direction of the animation sequence could therefore be seen as likely to increase the chance of a miscalculation and getting it wrong. Whilst T’s animations on lines 10 and 12 are met with laughter doing apprecia-
tion, from VP on line 11, and smiley voice from F on line 13, F’s animation appears to go unacknowledged. This can be further seen in extract 5.1 below:
Extract 5.1 (visual representation) “where you going”

10  T: if you bust a joke I might run away and laugh
   F: T____________________________________________
   A:*****gaze elsewhere*****,,,,,............T

11 VP: mhuhmuhmuhmuh
   F: T____________________
   A: T________,.............

12  T: if I bust [a joke you’re here ]=
   F: T____________________
   A:,,........T____________________

As T opens the animation space on line 10 it can be seen that both F and A are displaying listener-ship through their gaze. The interactional floor is being held by T and both speakers are orienting their gaze towards him. Initially, A’s gaze is oriented downwards. As previously stated, this could be a home position (Rossano, Brown & Levinson, 2009) and she is observed to begin to orient her gaze towards T as he continues his utterance. Lines 10 through 12 it can be seen that the interactional floor remains apportioned to T. However, floor apportionment during the production of F’s co-animation is not handled in the same manner. This can be seen in extract 5.2 below:
Extract 5.2 (visual representation) “where you going”

As F enters the animation space he is in overlap with T-the first animator. T can be seen to be gazing towards what is assumed to be VP, behind the camera. He does not orient his gaze towards F to apportion him the interactional floor at any point during his utterance. F’s gaze is withdrawn from T as F produces his co-animation, possibly indicating the withdrawal of the interactional floor from T and apportioning it to himself. A can also be observed to withdraw her gaze from T. She does not however orient her gaze towards F. F’s co-animation is both not tended to audibly, nor is he apportioned the floor through gaze.

The sequential deletion of co-animated utterance due to a mis-projection can further be seen in extract 5.3 below:
Extract 5.3 “Step your shit up”

1  T:  and I was like i remember when i was young someone used to
→ 2  make me promises of like (.) if you do that i’ll get you
3  trainers or if you do that i’ll give you money >blah blah
4  blah i was like< man needs to start doing that now
5  VP:  mmmmmmm
→ 6  T:  so (>i was like you know what<)=
7  F:  [that’s important ]
→ 8  T:  what’s your gcse’s saying she’s like >yeah=
9  F:  >word<
→ 10 T:  =im this year she’s this year blah blah blah< i was like
11  (. ) what you predicted n she said im predicted >this this
12  that< its in numbers now
→ 13  [not even >A’s and B’s she said im predicted<]=
14  VP:  [yeah yeah it is its crazy ]
→ 15 T:  =she said im predicted a 5 i was ooo >what’s that out of six;=
16  she’s like nah nah that’s outta 10 i was like ahh
→ 17 F: you need to step your [shit up boy]
18  VP:  [ah shit ]
→ 19 T:  i was like ah (>but she said >=
→ 20 F:  [not my family:]
21  T:  but she said< it’s like a B: or something like that like its
22  mad the way it tran[slates in (. ) ]=
23  VP:  [()i don’t under]stand it
24  T:  like(.)=
25  VP:  yeah
26  T:=5 to suttink is like a B >blah blah blah<

In the extract T is recounting an interaction with his younger cousin about her GCSE marks. T
opens the animation space on line 2 and continues to describe the interaction through to line 17,
where F enters the animation space with a co-animation. Much like extract 4.2, T is very granular in
the way he constructs his account of the interaction. Both VP and F display active listenersh, through VP’s backchannel on line 5 and F’s use of “word”\textsuperscript{19} on line 9. However, at line 17, despite T’s granular account, F seems to enter the animation space with a co-animation that is disaligned which results in a sequential deletion from T. On line 16 T provides minimal insight into his reaction towards his cousins GCSE mark. He states that he “was like ahh” and does not further expand on this reaction as he continues his account because it appears he misunderstood the conversion of her numerical mark to the system he is familiar with. However, before he can explain the misunderstanding, F enters the animation space to produce a co-animation which appears to be in the voice of T directed towards his cousin, stating that she needs to step her “shit up”. T sequentially deletes F’s utterance by repeating his initial reaction and then attempts to go on to explain his misunderstanding. However, he is overlapped by F on line 20 who produces a second co-animation furthering his initial understanding of T’s reaction. T again sequentially deletes F’s co-animation by repeating his utterance at the end of line 19.

As can be seen through T’s sequential deletion of F’s co-animations, the point of entry once again can be observed to play an integral part in the interactional risk of co-animation. Another instance of mis-projection resulting in the sequential deletion of a co-animated utterance can be found in extract 5.4 below:

\textsuperscript{19} ‘word’ can be used in a multitude of ways. It can be used to display understanding, agreement and can also be used as a greeting. The usage in extract 5.3, appears to be that of displaying understanding.
Extract 5.4 “one of you won it”

→ 1  T: someone said (.)<one of err> work colleagues said <ah how can they talk about representation one coloured person just [won ]

4 VP: [hohh]

5 £He said coloured?£

→ 6  T: one she said one coloured person just won the award

7 VP: AHAHHhhh

8 T: I WASsss uh [it was theres so much layers to what she said]=

9 A: [yeah: them >them< yeah: mhm: ]

10 T: =DAT privilege is there like (.)

11 A: *coloured*

→ 12 T: ONE [you said coloured]=

13 VP: [one yeah: one ]

→ 14 T: =and two you said one person so [basically be grateful ]

15 VP: [( ) representation yeah] 

→ 16 A: [ :aww awwwwh:: yay! ]

→ 17 T: be grateful n*gger

18 VP: yeah: yeah

→ 19 A: I’m so grateful

→ 20 T: one of you won it(.)=

21 VP: [yehh]

22 A: [mmm ]

→ 23 T: how many do you want now

24 VP: yeahah hah [hah HAH ]

→ 25 A: [how dare you]

26 VP: that’s [the thing they’ll be like]=

27 T: [dyu know how mad that is ]

→ 28 VP: =^they just wanna take ova^ 

29 T: YE[AH::]

30 A: [yeah] >yeah yeah yeah<
There are two instances of co-animation taking place in this extract. VP’s is accepted by the first animator with a loud, emphasised, and prolonged appreciation token “YE[AH::]”. Whereas A’s is sequentially deleted by the first animator and unacknowledged by the other participants present.

The point of entry into the animation space is one substantial difference between the two co-animated utterances. The animation space is opened at Stage 1 of the complaint sequence (see chapter 2), taking place on lines 1 through 3. The animated voice is that of T’s co-worker who T is presenting as having made a racially insensitive comment about an award show victory. On line 5, the participants can be seen to be moving into Stage 2 of the complaint sequence. VP is negotiating the terms of the complaint by querying if he heard T correctly. T then elaborates his initial animation further before moving on to stage 3-topicalisation of the trouble. T tropicalises the trouble through animations in the voice of self. These animations then receive affiliative responses from A and VP on lines 9, 13 and 15. These minimal affiliative responses are stage 4 of the complaint sequence.

The next stage is stage 5- “escalation, specification, enhancement, exemplification of the trouble/complainable through animation” (Cantarutti, 2020:358). A enters the animation space at stage 5. On lines 16 and 19, A produces a co-animation in the voice of T. However, on line 17, T produces an animation in the voice of the transgressor; both of the animations which sequentially delete both of A’s co-animations are presented by bringing the voice of the transgressor to the here-and-now. Thus, it appears T is still in stage 5 and is further escalating the complaint whilst A is doing being T and moving into stage 6-doing affiliation and moving towards an exit. The entry into the animation space prematurely also resulted in a misprojection of the complainable. T’s animations in stage 5 focus on the inferences being made by his co-worker through her initial statement,
whilst A’s co-animations focus on T’s possible attitude towards the inferences. Thus, when VP enters the animation space on line 28, not only is he entering at stage 6, after T’s request for an affiliative response on line 27, but he also produces a co-animation in the voice of the transgressor-furthering the inferences T suggested in his previous animations. These two factors contribute to the acceptability of VP’s co-animated utterance and were possible due to his point of entry into the animation space.

Another animation sequence which exemplifies sequential deletion due to a misprojection in the direction of the talk and untimely entry into the animation space can be found in the conversation taking place between F, T, VP and H below about a conversation between F and his cousin:

Extract 5.5 “I’ve got this”

1  F: obviously they had the-uh wh-uh lockdown school innit
2  H: mmmmm «at home»
3  T: on zoom yeah?
4 VP: ive been seeing some mad[ness ]
5  F: [on zo]om on the phone app so she’s like → 6  oh* she’s got homework for?y questions I was like ah(kisses teeth)[«cool» bring the]
→ 7  T: [I’ve got this ]
→ 8  F: bring the paper [let’s]
10 T: [HA HA]ha ha ha::
→ 11 F: let’s see innih
12 T: £you know it’s a challenge for you too£
13 VP: £uh yeah£
→ 14 T: £your thinkin£ lets go
15 F: she’s all bringing out iphone I’m thinkin
16 (0.1)
F opens the animation space on line 6 in the voice of self. However, the action of kissing his teeth which he performs directly before the production of his animation creates a slight pause in his utterance. It appears the lack of immediate continuation of the utterance to drive forward the narrative at this point is perceived as an invitation for collaborative completion by T. This results in T entering the animation space on line 8 by doing being F. However, F sequentially deletes T’s co-animation and attempts to complete his original animation by repeating “bring the” and makes it as far as “let’s” before he is overlapped by T’s laughter. The laughter does not illicit laughter from the other participants present and his co-animation continues to go unacknowledged. F’s original animation is then completed on line 11. However, it is followed by what appears to be a justification by T for his co-animated utterance. Notably, whilst doing so, T switches from using the pronoun “I” to “you”. He begins his utterance on line 12 with “you know” to possibly invoke relatability and reformulates (see chapter 4) his previous co-animation on line 9 by stating “your thinkin”. This switch in pronoun could be to invoke relatability, not only to the inner circle audience, but also to the outer circle audience (Burger, 1984)(see chapter 4). It is difficult to determine the target audience of T’s reformulation however, the reformulated co-animation once again goes unacknowledged by the first animator and participants in the inner circle. F then continues his narrative and
VP enters the animation space with a pre-completion co-animation on line 17. VP’s co-animation is met with laughter doing appreciation from both H and T before F goes on to confirm through further animation that VP was correct. It becomes apparent on line 23 that F’s narrative revolves around confusion surrounding the platform the homework in question is on. His clarification on line 5 about the “phone app” animation on line 9, where he asks for the “paper” foreshadows the later event of the “iPhone” being brought instead. However, due to T’s premature entry into the animation space, he mis-projected the direction of the talk and produced his co-animations around the difficulty or number of homework questions rather than the platform on which the homework was given. It is likely due to this miscalculation that despite T’s numerous attempts his co-animations go largely unacknowledged. Whilst his first co-animation is not acknowledged in the talk or through gaze, his second co-animation is somewhat acknowledged through gaze as can be seen in extracts 5.3 and 5.4 below:

Extract 5.3 (visual representation) “I’ve got this”

5 F: [on zoom on the phone app so she’s like

H: ****gaze elsewhere****,...F________
T: ******gaze elsewhere******,,.........

6 °oh° she’s got homework for?y questions I was like ah(kisses

H: F ________________________________

T: ..F_________________**gaze averted**,..F________________

7 teeth)(°cool° bring the)

H: F ________________________________

T: F ________________________________

8 T: [I’ve got this ]

H: F ________________________________

F: **gaze elsewhere**

9 F: bring the paper [let’s]

H: F ________________________________

T: F ________________________________

As can be seen in extract 5.3 above, the floor is apportioned to F by both T and H throughout the onset of his narrative. T briefly averts his gaze downwards to return to home position (Rossano, Brown & Levinson, 2009) before returning his gaze to F, the first animator. T’s gaze remains on F throughout both F’s animation, and his own co-animation. However, F does not meet his gaze and apportion him the interactional floor. Instead, F continues with his original animation and H’s gaze remains on F. However, in the extract 5.4 below of T’s second co-animation, it can be seen that both F and H do somewhat acknowledge his utterance through their gaze.
5.4 (visual representation) “I’ve got this”

As can be seen above, whilst T’s co-animations are sequentially deleted and unacknowledged in the talk, F does orient his gaze towards T when he is producing his second co-animation. H can also be observed to be smiling as T begins producing his utterances in smiley voice. However, the duration of the gaze during each of T’s utterances are brief. F’s gaze is withdrawn from T before he can complete his co-animation and H does not orient her gaze towards T at all during his co-animation.
Thus, whilst there is acknowledgement taking place it appears T is still not apportioned the interactional floor entirely. F reclaims the interactional floor from T on line 15, where he presents the issue at hand—the confusion at his cousin bringing out her iPhone.

This chapter has examined instances of co-animation that were sequentially deleted or went largely unacknowledged by the first animators and/or the participants present in the interaction. The use of sequential deletion and floor apportionment through the use of gaze (where relevant) have been examined as ways in which the first animators and participants deal with disalignment in the animation space. The environment in which they are recurrent has been determined through the examination of epistemic positioning and projectability, in contrast with that of co-animations that were accepted by first animators and participants.

Whilst the previous chapter established the role of epistemic access and authority in the acceptability of co-animated utterances, this chapter has served to provide further insight into the relationship between epistemic positioning and the probability of disalignment. The instances of disaligned co-animation sequences examined in this chapter provide a further understanding of how animation sequences are navigated by both the first animator and co-animator(s) in relation to the voices that are brought into the animation space. The gradient of epistemic access in regard to animations produced in the voice of self versus that of another are analysed and the importance of the voice in which co-animated utterances are produced is established.

In exploring the relationship between the way voices are animated and co-animated it was established that the chances of disalignment in animated sequences are likely to increase when the co-
animator is animating in the voice of another due to the lack of epistemic access to the original interaction. The data analysed in this chapter indicates that co-animators when given access to both the voices of self and another may be likelier to co-animate in the voice of self due to the perceived lack of epistemic access and rights. It was further found that co-animations produced in a voice which the first animator did not provide epistemic access to, are more likely to be sequentially deleted. The data set revealed that this is often due to, not only the lack of epistemic access but also the lack of projectability of the trajectory of the animated sequence.

Moreover, the analysis of projectability in relation to disalignment revealed that despite co-animators animating from a K+ position (Heritage 2012), of having epistemic access, their point of entry into the animation space contributed significantly to the probability of miscalculation and thus possible sequential deletion. Although the co-animator, in several of the cases examined, produced co-animations in the voice previously presented in the animation space, their premature entry into the animation space contributed to the reception of their animation; first animators, in a range of sequences, complaints, misunderstandings and joint fictionalisations, sequentially deleted co-animations that did not align with the progressivity of the animation sequence. Other participants were also observed to be doing sequential deletion, through their entry into the animation space, lack of forthcoming appreciation and lack of floor apportionment.
6. Conclusion

6.1 A summary of findings and thesis organisation

This thesis has been centred on the analysis of the practice of *co-animation*. The instances of co-animated utterance examined in this paper have sought to establish the environments and properties which contribute to the acceptability of co-animated utterances and the environment in which the joint act of *doing being* ones past self or other, through the recontextualization of voice, is dis-aligned. The role which the first animators, co-animators and other participants play in co-animated sequences have been examined to assess the varying degrees of appreciation and responses to the phenomenon in multi-party interactions.

Chapter 2 reviewed previous literature surrounding *reported speech, demonstrations, enactments, re-enactments* and *layering* in order to establish the key features of *animation* as a practice which allows animators to bring voices from previous and or fictional interactions into the here-and-now. The act of animating allows speakers to separate themselves from the re-enacted actions of the speaker’s voice they are bringing into the here-and-now whilst simultaneously *doing being* the speaker. Thus, the ways in which the previous literature have contributed to the understanding of animation as a complex, yet distinct, phenomenon is outlined through the examination of each term.

Chapter 3 reviewed the methodological approach and decisions undertaken in order to examine the cases of animation and co-animation found in this study. The theoretical and ethical considerations
are outlined before assessing the limitations of the corpus and the way in which these aspects are handled within the thesis.

Chapter 4 determined the factors which contributed to safe entry into the animation space. The way in which both animated and co-animated utterances are designed by participants and the interactional relevance of the turn design is outlined. It was established that the epistemic position of potential co-animation was an integral factor to safe entry into the animation space. Cantarutti (2020) suggested that the affiliative nature of co-animation trumps the epistemic entitlement between participants. The findings in this study support this claim to some extent. This can be observed in the way first animators navigate epistemic positioning when entering the animation space. The first animators are observed to be neutralising their epistemic access, rights, and authority to animated content through the way animated experiences are framed. The presentation of experiences in a general manner is one way in which first animators were observed to be doing this. In the case that an experience was specific to the first animator, the use of pronouns to present a relational stance was observed as a method which the first animator used to level the epistemic positioning between themselves and that of their recipients. These factors were observed to aid in allowing co-animation to safe entry into the animation space. Therefore, whilst the affiliative nature of co-animation could be suggested to trump considerations of epistemic positioning, the first animators can be observed to be taking into consideration their epistemic position when designing their animations. Co-animation are also observed to be designing their co-animated utterances to ensure safe entry into the animation space through the use of formulations and reformulations. The implicative and paraphrastic na-
ture of these devices could be suggested to minimise the possibility of a miscalculation. Furthermore, the point at which the co-animators entered the animation space was also a key factor when ensuring safe entry into the animation space. Co-animations which were produced at a point where the stance of the first animator or the direction of the talk was made apparent were more likely to be accepted by the first animator.

The analytical conclusions reached in chapter 4 are drawn upon to examine the instances of disaligned co-animations in the final chapter of this thesis—chapter 5. The investigation into disaligned instances of co-animation further revealed that the voice in which animations, and in a majority of cases, the consequent co-animations were produced bore weight in the way the turns were designed and received. As previously established first animators can be observed to provide first animations in the voice of self, another, self + another and so forth. In cases where the first animator brought more than one voice to the animation space, co-animations which were produced before the first animator provided epistemic access to the secondary voice were observed to have an increased chance of disalignment. The premature entry into the animation space resulted in co-animators lacking knowledge on the direction of the talk or miscalculating the stance taken by the owner of the animated voice.

To conclude, this paper forms the basis of an initial investigation into the disalignment of co-animated utterances and the environments in which they are recurrent. It also begins to investigate co-animations in multiparty interactions and interactions with both primary and secondary audiences. Whilst these aspects are drawn upon in this study, due to the limitations of time and length in this
thesis, they are areas which could be further researched at a later date. The way in which (co)anima-
mations are designed and how the affiliative nature is affected by the presence of a secondary audi-
ence, such as podcast listeners, requires deeper investigation. Furthermore, the relationship between
gaze and appreciation or lack thereof could be expanded further to the initial investigations made in
this paper. However, this thesis serves as a foundation for further analysis and provides insight into
a complex phenomenon to further the understanding of co-animation and the way in which humans
use doing being to associate to their listeners and associate with their fellow humans.
7. References


Appendix I: Transcription conventions

Transcription conventions (Jefferson 2004) and adaptations:

[ ] overlap and simultaneous talk

= latching

°word° the utterance is softer and or quieter than surrounding talk

( . ) a brief pause (approximately a tenth of second) within or between utterances

( 1. 0 ) measured pause (to the tenth of a second)

___ Stress in pitch and or amplitude

: prolongation of the prior sound

.,?? the ‘usual’ intonation

> < the utterance is produced faster than the surrounding talk

< > the utterance is produced slower than the surrounding talk

£ smiley voice/supressed laughter

( ) unintelligible talk

^^ accented voice
Transcription conventions (Goodwin 1981) and adaptations:

Name: the person whose gaze is being recorded

,... withdrawal of gaze

... incoming gaze

Name______ Point at which gaze has reached the participant named

** inserts accounting for gaze when elsewhere
Appendix II: Access to data presented in thesis

The YouTube weblinks to the podcasts used are provided below with an individual time stamp for each data extract:

https://www.youtube.com/watch?v=j8qvXIXavrc

Extract 1.1: (22:18 – 22:27)

Extract 1.2 and 5.1: (29:15 – 29:27)

Extract 4.1: (22:18 – 22:29)


Extract 5.5: (20:36 – 21:00)

https://www.youtube.com/watch?v=EVTdMqgW7IA

Extract 4.3: (11:43 – 12:14)

https://www.youtube.com/watch?v=80cyVMWE1zk

Extract 4.4: (10:23 – 10:46)

https://www.youtube.com/watch?v=5ohC11_Z-eA&t=852s

Extract 5.2: (23:53 – 24:10)

Extract 5.4: (15:10 – 15:45)

https://www.youtube.com/watch?v=8kIFuWeHRjM

Extract 5.3: (11:54 – 12:27)