Pursuits in Collision

Affiliation, Disaffiliation, and Multimodality in Persian Interaction

Reihaneh Afshari Saleh

PhD

Abstract

This study is on pursuing an interactional outcome in the face of a co-interactant's resistance. Despite at least a forty-year history of research on pursuits in social interaction (Jefferson, 1981; Pomerantz, 1984b), there is still much to explore about this ubiquitous social phenomenon. This research employs a multimodal conversation analytic methodology to address some less-explored questions on pursuits: what practices does an interactant use to further their course of action against their co-interactant's resistance? Do the details of these practices have implications for the trajectory of the interaction towards escalation or de-escalation? What do these practices tell us about the agentive stance adopted by the pursuing party? And how can interactants heading towards an escalated pursuit manage disaffiliation? Two different types of pursuit sequences are introduced: *persisting* in furthering one's course of action and *gradually desisting* from a course of action. The findings show a novel phenomenon called *multimodal gradation*: a temporally coordinated up- or downgrading of a multitude of resources that are simultaneously used in formatting a social action. Borrowing Mondada's terms (2014), a whole "multimodal Gestalt" by which a turn at talk is delivered is up- or downgraded. Multimodal upgrading of a pursuit turn projects further expansions to the pursuit sequence and it can escalate an initial clash. On the other hand, multimodal downgrading of a pursuit turn projects a contingent sequence closure and de-escalation. Also, upgrading the multimodal Gestalt of a pursuit turn displays the pursuing party's stronger agentive stance compared to downgrading the turn. The project introduces another multimodal phenomenon termed *mock aggression*. Used between intimate interactants, mock aggression offers opportunities for affiliation despite its aggressive appearance. The findings have implications for our understanding of sequence and preference organization in CA, multimodality, agency, and conflict management. Data are in Persian and collected in Iran.

Contents

List of Non-Transcript Figures

List of Tables

Acknowledgements

I would like to express my heartfelt appreciation and affection to my participants, some of whom saw me begin my project but are not here to see it finished. Wherever they are now, my most sincere thanks go to them for their enormous care, for going the extra mile to make the recording sessions successful, for generously offering me time the night before their important exams or the day after their long journeys, for sharing with me their affective moments of sadness and happiness, and for allowing me to have a glimpse of their daily lives as authentic as can be. Time has passed since the recordings were made. Those who were about to start their higher education are now graduated and those who were about to get married are now parents, but for me time stands still when I work with their data. And no experience or achievement can even come close to a sense of 'timelessness'. I am indebted to them for giving me a treasure to use for the rest of my academic career.

I am extremely grateful to my supervisors *Richard Ogden* and *Kobin Kendrick* for the pearls of wisdom they have offered me and for their enthusiasm about this project. Many thanks to Richard who showed me the beauty of exploring social interaction when I had come to the UK to become an experimental phonetician. Six years ago, when I started my master's degree on 'Phonetics and Phonology' at York, it was Richard's course on 'the Phonetics of Talk in Interaction' that made me fond of analyzing social interaction. When the whole cohort was using pre-collected English data, Richard trusted me with collecting and working on Persian data, and it was his trust and encouragement that made me feel confident enough to step on my PhD path. My warmest thanks go to him for his patience and support in each step of this path. I cannot overstate how much I appreciate Kobin for his indispensable training and stimulating thoughts which have greatly shaped my analyses. Enormous thanks to him for his unparalleled guidance and his meticulous feedback on almost all my writings, for believing in my work, influencing my approach to data with his analytic rigor, and helping me to see the forest when I was lost for the trees.

I owe a great debt of gratitude to my examiners *Merran Toerien* and *Danielle Pillet-Shore* for their encouraging feedback, enthusiastic discussion, and useful suggestions for further research in the viva. Especial thanks to Merran for accepting to be in my Thesis Advisory Panel and being a source of inspiration throughout my PhD years. I deeply value the support and friendship of the Interactional Linguistic Lab and CASLC communities. I would especially like to mention *Paul Drew* for his generous time and insightful discussions on some of my data. I extend particular thanks to the Department of Language and Linguistic Science for sponsoring my PhD; to *Sam Hellmuth* for sharing with me her insight on intonation and to *Nino Grillo* for trusting in me and offering me a unique teaching experience. I would also like to thank *Barbara Fox* for the most enjoyable data sessions and the unforgettable poetry nights.

I owe so many thanks to my friends. To my lovely *Sara* and *Marjan* who kept sending me courage and hope. To my dear *Laleh* for her loving concern for this project and my future. To *Emanuela* for the most peaceful moments by the river, counting the stars and listening to birds. To *Marina* for her unfailing understanding and care, and to *Rose* for our lovely walks and amazing chats.

My most wholehearted thanks go to my family. They are the light in my dark nights, the alchemists whose smiles turn my saddest moments into uttermost happiness. Huge thanks to them for reminding me of my strengths when I felt the weakest and for helping me have faith that all ends well. Most especial thanks to my parents, the kindest angels of my life, without whose support this journey could not have even started let alone come this far.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References. A Part of this work has been published in *Research on Language and Social Interaction* [https://doi.org/10.1080/08351813.2020.1833590].

Chapter 1.   Introduction

1.1.   Prelude

Navigating our daily lives, we sometimes find ourselves in situations in which what we pursue is in collision with what our coparticipant pursues: We want a window open; they are closing it. We want a boardgame to go faster; they are taking their time to decide on their move. We want a mobile phone they are holding; they are still using it, and so forth. These situations can be seeds of conflict. They can sprout disaffiliative encounters in which we tell our co-interactant to do the action that we pursue, but they resist it. And these disaffiliative moments can in turn grow into conflicts if each party insists on pursuing their own course of action.

Consider the following episode of interaction between two Persian friends Roshana and Azita. They are waiting for the results of their university applications to be uploaded online. Prior to this extract, they had a rather long dispute on whether the results are released. Roshana says that the results are uploaded; Azita says that the web page does not display the actual results and only an announcement on the release of the results is uploaded. Azita eventually gives her mobile phone to Roshana to check the web page and picks up Roshana's book and starts reading it (Figure 1.1). The extract begins as Roshana reads a sentence from the online announcement: 'It {the result} has been uploaded' (line 1). Azita then challenges the announced news and performs embodied actions that project she is going to take her phone back. Then the two participants perform actions that are in collision with one another. For example, Roshana issues imperatives and requests more time to use Azita's phone ('wait'), but Azita issues imperatives and requests her phone back (give {it} give {it}). The participants 'throw' imperatives at one another and the courses of action that they are pursuing clash: one course of action (CoA) is designed to keep the phone and the other is to take it back.

Figure 1.1. Participants in Extract 1.1. Roshana (left) is using Azita's (right) mobile phone to check the results of their university application. Azita is reading Roshana's book.

Extract 1.1. RAY041a, 10.29

```
01    ROS    gharâr gereft-e           ((reads from the phone))
             stable take.3SG-PRF
             it's been uploaded.

02           (0.4)
```

```
03    AZI   +kh:ob\+ (.) khob ku::::. +#
            PRT     (.) PRT  whe::re
            but but whe::re.
            +......+throws hands in air+
      fig                              #1.2
```


#1.2

```
04    AZI   kodum\      kojâ:-[s.
            which       where-be.3SG
            where the\ whe:re is it.
```

```
->    05    ROS                    [enâ   sab  +[kon,=      +#
                                    Here patience do.IMP
                                    here wait,
```

```
06    AZI                                      +[((click))   +#
      azi                                      +puts book dwn+
      fig                                                    #1.3
```


#1.3

```
=>    07    ROS   =+ye  lahze  dahant-o      ±be-band,+        ±#
                    One second  mouth-OBJ.mar IMP-shut.2SG
                    shut your mouth a second,
            azi   +uncrosses legs--------------------+
            ros                               ±hold hnd vertically±
            fig                                               #1.4
```


#1.4
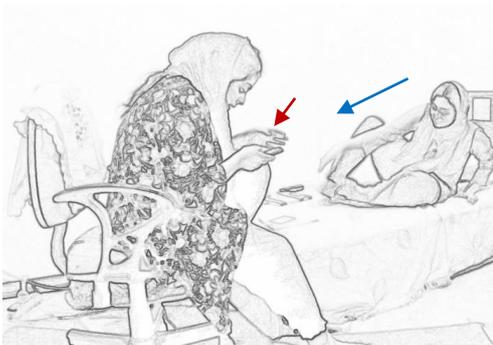
```
08    AZI   #+<<whispery>ah.>
                        INTJ
                        Damn.
             +crawls towards ROS-->
      fig   #1.5
```

```
09    ROS   zemanan  ân daste az dâv\    ((reads from the phone))
             In.the.meanwhile that group of applican\
             in the meanwhile those applican\

10          (1.1)

11    ROS   besmelâhe            rahmâ[ne Rahim
             In.the.name.of.Allah merciful compassionate
             in the name of God the Merciful the Compassionate[1]

12    AZI                            [be-de,      be-de,+
                                     IMP.give.2SG IMP.Give.2SG
                                     give {it}, give {it},
      azi                                              -->+
```

```
=>    13    ROS   ±.h #khafe    sho          khob?±=
                   .h suffocated IMP.become.2SG OK
                   .h Shut up OK?
                   ±hold arm between mobile and AZI---±
            fig      #1.6
```

```
=>    14    ROS   =[khafe     sho.
                   suffocated become.2SG
                   shut up.
```

---

[1] Among some Persian speakers, this phrase is sometimes used at the opening of activities or at the opening of the stages within an activity, especially when they are 'important' to the speakers. Uttering this phrase can potentially indicate that Roshana is getting closer to finding the results.

```
15   AZI   [<<f>bâbâ nemiâd;>
              PRT   NEG-come.3SG
              gosh/hey {the results} won't come up;

16   AZI   †m:::i[gam#          in    etelâiye†±\
           PRG-say-1SG          this announcemen\
           I'm telling you this is an announcemen\
```

```
->   17   ROS   †     [<<f,h>kh #eile khob      ††bâbâ bezâ::?>†
                       very    OK              PRT IMP-let.2SG
                       OK OK my gosh wait?
           azi   †raises right hnd twrd phone-----†.............†
           ros   ±pushes hand against AZI's hand--±
           fig               #1.7
```


#1.7

```
18   AZI   #†.h e:::::::::†
            †claws fingers†
           fig  #+1.8
```


#1.8

```
19         #⊤        (0.2)              ⊤ (0.5)⊤
           azi   ⊤throws book with left hnd⊤,,,,,,⊤
           fig  #1.9
```


#1.9

```
20   ROS  leh    kard-i    un  ketâb-o:::::::::::
          smashed do.PST-2SG that book-OBJ.mar
          you smashed that boo:::::::::k
```

This thesis is concerned with episodes of interaction similar to the above. All the episodes come from Persian data. They consist of sequences in which at least one interactant uses an imperative structure to get the co-interactant to perform an action; the co-interactant shows resistance against complying, one way or another, and the imperative-speaker pursues compliance. An imperative structure can be in service of performing various social actions (see §1.6 for a review), but in this thesis, the term imperative is a shorthand for an imperative verb that is designed to get the co-interactant to perform a physical action unless stated otherwise. This introductory example shows some of the recurring features and phenomena that we will see in the rest of the thesis.

Firstly, as briefly mentioned above, the extract depicts a sequence in which participants issue imperatives to get the co-interactant to perform an action. For example, after some disagreement on the status of the results (lines 1-4), Roshana issues the imperative 'wait' and requests more time to use Azita's phone (line 5). Then she issues another imperative 'shut your mouth a second' (line 7) designed to quieten Azita, who is producing vocal displays of exasperation on Roshana's positioning regarding the results (line 6).

Secondly, the extract shows that the imperative-recipient resists against the imperative-speaker's CoA (Resistance is conceptualized in Chapter 2). For example, immediately after Roshana requests more time with the phone (line 5, 'wait'), Azita performs embodied actions that project movement towards the phone, potentially to take it back: she puts the book down (line 6; Figure 1.3) and then uncrosses her legs (line 7, Figure 1.4). Both of these bodily actions can be seen as the preparation phase of moving towards Roshana. Moving towards Roshana, in turn, can potentially be the initiation of a CoA in which Azita takes her phone back. Indeed, after this preparation phase, she crawls towards Roshana (lines 8-11), and when she reaches her, she requests her phone back (line 12). Therefore, Azita's resistance against Roshana's imperative ('wait) is both embodied (reaching her) and verbal (counter-imperatives) (line 12). While in this extract verbal resistance against Roshana's CoA is done through counter-imperatives, as we will see in the following chapters, resistance can be done through other means such as rejection tokens and the like. What is recurrent is that at least one of the speakers issues an imperative and the co-interactant resists against the imperative-speaker's CoA.

Thirdly, what features in this extract is pursuing compliance ('Pursuit' in CA is reviewed in Chapter 2). Roshana pursues compliance in two turns (lines 13 and 17). One of her earlier imperatives was *ye lahze dahanto beband* 'shut your mouth a second' (line 7) and after Azita verbally requests her phone back, the imperative *khafe sho khob khafe sho* 'shut up ok shut up' pursues compliance (line 13). Her other imperative was *sab kon* 'wait' (line 5), and when Azita upgrades her embodied attempt to reach the phone (line 17), the imperative *bâbâ bezâ* 'my gosh wait' is delivered (line 17). In this extract the lexical design of the initial imperatives and their subsequent versions differ, but we will see that pursuit may be done by repeating a prior imperative.

Fourthly, this extract illustrates an instance of a main phenomenon that this research introduces: multimodal gradation (Chapters 3-4). A type of multimodal gradation is multimodal upgrading. While pursuing their CoAs, participants upgrade multiple design features of their turn simultaneously. Compare the imperative *sab kon* 'wait' (line 5) with its subsequent version <<f, h>*bâbâ bezâ::*> 'my gosh wait' (line 17). Relative to the initial version, the subsequent version is produced louder and higher, and it includes an interjection that displays exasperation (*bâbâ* 'my gosh'). The embodied behaviour is also upgraded. As shown in Figure 1.10, Azita (blue arrow) gets progressively closer to Roshana to take the phone back and Roshana (red arrow) embodies an increasingly 'bigger shield', protecting the phone from being grabbed. When Azita is still relatively further away, Roshana only holds her hand vertically next to the phone with the back of her hand oriented towards Azita as if she is protecting the phone

from being grabbed (line 7, Figure 1.4, Figure 1.10, panel A). When Azita reaches the phone and issues the imperatives 'give {it} give {it}' (line 12), Roshana upgrades her embodied shielding. She places her whole arm between Azita and the phone (line 13, Figure 1.6, Figure 1.10, panel B). Then Azita raises her hand and moves it towards the phone. This action is observably designed to 'invade' Roshana's shield, and hence is an upgraded movement in reaching the phone. Roshana also raises her hand and pushes Azita's hand back, an upgraded shielding (line 17, Figure 1.7, Figure 1.10, panel C).

| A | B | C |
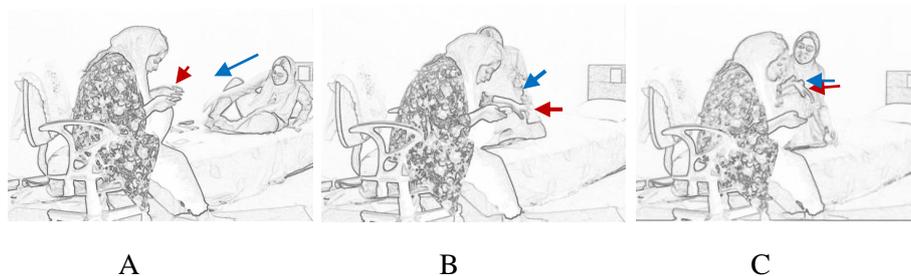
Figure 1.10. Embodied upgrading in the bodily conduct of the two participants. The participants' embodied clash is manifested in Azita's bodily actions designed to get the phone and Roshana's bodily actions designed to keep the phone.

Finally, the extract also includes another phenomenon that the thesis introduces: *mock aggression* (Chapter 6; Afshari Saleh, 2020). When the participants' disaffiliative exchange is heightened (lines 1-17), Azita claws her hand and produces an elongated open-mid central unrounded vowel with advanced tongue root [ɜ::::::]. The advanced tongue root gives a coarse quality to the vowel, one that resembles a sound that may be heard when one's throat is being squeezed. This can be seen/heard as a mock strangling action (line 18; Figure 1.8). Immediately, after this Roshana smiles, which can potentially show her orientation to Azita's gesture as physically non-threatening. The smile stops when Azita throws Roshana's book further away on the bed (line 19. Figure 1.11). Roshana complains about her book being smashed (line 20) and with this the participants transit from the imperative sequence to a complaint sequence.

Figure 1.11. A mock strangling gesture.

The participants are, therefore, in a conflictual episode with disaffiliative exchanges. They are trapped in a loop of disaffiliative conduct. Besides conflict and disaffiliation which are the running undercurrent of the sequences under investigation, a more specific common thread is 'pursuit', and not just pursuing *a* response but more specifically pursuing a preferred response.

Against this backdrop, the thesis aims to answer the following questions:

1. What verbal practices do participants use to pursue compliance in the face of resistance? (Chapter 2)

2. Do the details in the composition of the pursuing action have consequences for the trajectory of the interaction? More precisely, can the composition of the pursuit turn be projective of escalation or de-escalation? (Chapters 2, 3, 4)
3. What implications do the multimodal design features of a pursuit turn have for the structure of the sequence? (Chapters 3-4)
4. What implications does the action of pursuing have for the participants' agency? (Chapter 5)
5. Do different practices of pursuit vary in terms of the agentive stance that the pursuing party adopts? (Chapter 5)
6. How may participants exit an escalated conflict? (Chapter 6)

In seeking answers to these questions, this thesis uses a multimodal conversation analytic methodology. Therefore, in the following, some commonly used analytic terms from conversation analysis (CA) are briefly reviewed (§1.2). Next in §1.3, research on multimodality within and beyond CA is touched on. Then, the main theme of the thesis – conflictual interaction – is reviewed (§1.4). I save reviewing the literature on pursuit (the second over-arching theme of the thesis) for the introduction of Chapter 2 and other more specific concepts such as upgrading or downgrading will be reviewed at the beginning of each relevant chapter. The analyses in the forthcoming chapters rely on two concepts: preference and accountability, so after having a look at the literature on conflictual interaction, these two concepts are reviewed (§1.5-1.6). Then because imperatives are the featuring syntactic structure in the pursuit sequences that we examine, a review of CA studies on imperatives is presented (§1.7). Next, we will have a look at Persian grammar and then more specifically at Persian imperatives (§1.8). The data and methodology of the research are explained in §1.9 and §1.10, respectively. Finally, a preview of the organization of the thesis is presented in §1.11.

1.2.   Conversation Analysis

Conversation analysis (CA) "is an inductive, micro-analytic, and predominantly qualitative method for studying human social interactions" (Hoey & Kendrick, 2017, p. 151). The origin of this method is rooted in sociology and ethnomethodology, and then it gained popularity in a variety of disciplines such as linguistics, anthropology, psychology. CA is deeply rooted in the sociologist Erving Goffman and the ethnomethodologist Harold Garfinkel's approach to social interaction (Heritage & Clayman, 2010). Among Goffman's contribution to sociology is his shrewd observation that social interaction is orderly, and it is "a form of social organization in its own right" (Heritage & Clayman, 2010, p. 8). He argued that social interaction consists of natural *units of action* and a great deal of his work was to "uncover the normative order prevailing within and between these units" (Goffman, 1967, pp.1-2). Garfinkel (1967) highlights social members' commonsense and shared methods of understanding. He notes that common understanding between social members "consists of various methods whereby something that a person says or does is recognized to accord with a rule" (p. 25). So he, too, talks about rules and orderliness based on which members of a community make sense of one another's social actions. Influenced by Goffman and Garfinkel, Harvey Sacks, in collaboration with Emanuel Schegloff and Gail Jefferson in the 1960s, pioneered applying a novel set of analytic 'tools' to show the orderliness underlying the social actions that participants do to and with one another (Heritage, 1984; Heritage & Clayman, 2010). His analytic toolkit was soon known as *conversation analysis* (Sacks, 1992, I-II). This toolkit includes important interactional organizations, some of which are *turn-taking, turn design, sequential organization*, *adjacency pair*, and *preference* (for other organizations in CA, refer to Hoey & Kendrick, 2017; Sidnell & Stivers, 2013; and Sidnell, 2010). In the following, I briefly touch on turn-taking, turn design, sequential organization and adjacency pairs (§1.2.1-1.2.4). Preference is reviewed in more depth later in the chapter, with a focus on preference in conflicts and arguments (§1.5).

1.2.1. Turn-taking

This organization explores the orderliness involved in the allocation of a turn to participants (Sacks, Schegloff, & Jefferson, 1974). Turns are participants' interactional slots in which they can perform social actions and contribute to the activity in progress. Turn-taking deals with the *who* question: which

participant performs a social action next. A turn consists of at least one turn constructional unit (TCU). Turn constructional units are traditionally considered to be composed of "linguistic units (words, phrases, clauses, etc.) that form a recognizably complete utterance in a given context" (Hoey & Kendrick, 2017, p. 153). With the increasing popularity of a multimodal approach to interaction (see §1.3), embodied components are also considered in the building block of a turn (Mondada, 2007). TCUs are vehicles by which social actions are performed (Schegloff, 2007, p. 9). There are various grammatical, phonetic, and embodied clues that a TCU or a turn in progress is reaching its end (Ford, Fox, & Thompson, 1996; Auer, 1996; Rossano, 2013), at which point, the next speaker may take the turn. This point is called a transition-relevance place (TRP) in which the transition of turn from one speaker to the next may occur.

1.2.2. Turn Design

So a TCU is a vehicle for action, but how does the analyst know what social action is being performed with a turn at talk? Turn design is an aspect of interaction that can clue the co-interactants and the analyst on the type of action: "Turn design refers to how a speaker constructs a turn-at-talk – what is selected or what goes into ' building ' a turn to do the action it is designed to do, in such a way as to be understood as doing that action (Drew, 2013, p. 132). Turn design, therefore, deals with the *how* question. A turn can be designed with various multimodal features: lexical, syntactic, phonetic, embodied, and so forth, and it is the combination of these features plus the relative positioning of a turn in a sequence that contributes to action formation and ascription (Levinson, 2013). Consider the action of requesting an object like a mobile phone that a co-interactant is holding. One can perform this action through a solely embodied design: extending hand towards the mobile phone (Kendrick, 2020). Another possible design may involve embodied and lexical components, for example, pointing towards the phone and issuing the imperative 'Give it'. Then the imperative may be produced with noticeably high and loud prosodic features. The precise social action that is being done in these three scenarios is slightly different. In this thesis, CA's turn design toolkit is extensively used to analyze pursuit sequences.

1.2.3. Sequential Organization

Sequential organization "is any kind of organization which concerns the relative positioning of utterances or actions" (Schegloff, 2007, p. 2). So sequential organization deals with the question *where* in an interaction an action or an utterance is produced. For example, within a turn at talk, the relative positioning of the lexical components of that turn is a matter of sequential organization. The relative ordering of *actions* is referred to as *sequence organization* (Schegloff, 2007, p. 2; Kendrick et al, 2020). The main assumption is that interaction is not a random occurrence of a series of actions; rather, it is made up of sequences – "coherent, orderly, meaningful successions" of actions by which activities are done" (Schegloff, 2007, p. 2). We will pick up the concept of sequence organization in Chapter 2, where we dissect the structure of pursuit sequences. Here, suffice it to say that sequence organization is another aspect of interaction that can clue the kind of social action that is performed in a TCU:

> [T]hroughout the course of a conversation or other bout of talk-in-interaction, speakers display in their sequentially 'next' turns an understanding of what the 'prior' turn was about. That understanding may turn out to be what the prior speaker intended, or not; whichever it is, that itself is something which gets displayed in the next turn in the sequence. We describe this as a *next-turn proof procedure* (Hutchby and Wooffitt, 1998, p.15).

This is how sequentiality or 'nextness' matters. Next-turn proof procedure provides an analytic rigor. Rather than relying on their own understanding, conversation analysts base their analyses on the *participants*' emic understanding of one another's turns at talk. Participants' emic understanding of the prior talk is reflected by the action that they perform in the next turn. The analysis in this thesis follows this golden rule. Claims and arguments are backed up with interactants' observable orientations and evidence from the data.

1.2.4. Adjacency Pairs

Two rather well-established sequence types in CA are the organization of adjacency pairs (Sacks, 1967; Schegloff, 1968; 2007) and storytelling (Mandelbaum, 2013). The sequences in this thesis are structured in adjacency pairs, and, therefore, this section briefly summarizes some of the main concepts related to this sequence type, mainly based on Schegloff's (2007) book *sequence organization.*

Schegloff (2007) lists the main features of an adjacency pair structure: it is "composed of two turns", "by different speakers", "adjacently placed", "relatively ordered", and "pair-type related" (p. 13). While all of these are crucial in shaping an adjacency pair, I zero into the relative ordering of the actions in a sequence: whether an action is placed first or second in its pair. Some actions such as questions, offers, commands typically go first, and thus they are "first pair parts" (FPP) while some others such as answers, acceptance/rejection, and (non)compliance, go second; they are "second pair parts" (SPP) (p. 13). FPP actions *initiate* a sequence while SPP actions are *responsive* (p. 13). A basic minimal adjacency pair thus consists of two turns, an FPP and an SPP. In the sequences in this thesis, the imperative turns are in an initiating position unless stated otherwise. They perform a variety of relatable actions such as advice, instruction, request, commands, and offers, so they are FPPs (In Chapter 4 we will see an imperative turn in a non-initiating position, but other than that, the imperatives in this thesis are FPPs).

What operational rules do actions follow that convey they are not haphazard productions of two random social actions simply following one another? For actions in an adjacency pair, Schegloff (2007) explains, this rule is called the 'conditional relevance rule'. Conditional relevance is a relationship between an FPP and SPP when the FPP makes the production of an SPP by a co-interactant a conditionally relevant next action. After an FPP, the absence of an SPP comes off as noticeable, official, and consequential (p. 20). This is one piece of evidence for the operation of conditional relevance rule. Another piece of evidence is that in the absence of an SPP, the FPP-speaker may *expand* the sequence and pursue a response (see Kendrick et al., 2020, pp. 121-122 for various types of evidence on the operation of conditional relevance). Since this thesis is on response pursuits and sequence expansion, the whole of Chapter 2 is dedicated to show the implications of a response pursuit for the structure of an adjacency pair. A review of sequence expansion is provided at the beginning of Chapter 2.

As mentioned earlier, data in this research is analyzed through a multimodal conversation analytic approach. Having reviewed some main CA concepts, I now review multimodality within and beyond CA.

1.3. Multimodality

Multimodality has been addressed by researchers within and beyond CA. From early stages after the birth of CA, the field's pioneers started to use video data to find orderly patterns of social conduct in face-to-face interaction (see e.g., Goodwin, 1981; Heath, 1984, 1986; Sacks & Schegloff, 1975/2002). Influenced by gesture researchers such as Adam Kendon (e.g., 1972; 1975), these CA scholars drew the field's attention to embodiment and its consequentiality in interaction. The phoneticians of talk in interaction were also documenting ways in which the phonetic features are systematically used by interactants in formatting social actions or structuring turns or sequences (e.g., Local & French, 1983; Local & Kelly, 1986; Local, Kelly, & Wells 1986; Couper-Kuhlen, 1992). Linguists concerned with grammar and the verbal aspect of talk also presented evidence on the interplay between grammatical, embodied, and prosodic cues (e.g., Auer, 1996; Ford & Thompson, 1996). All this resulted in calls for an incorporated approach to interaction that takes into account the multiplicity of the resources that interactants employ (e.g., Local, 2003, Mondada, 2014, Walker, 2012 among others), an approach which has gained popularity as the multimodal approach to interaction (Deppermann, 2013; Goodwin, 2018; Heath & Luff, 2013; Mondada, 2014; Streeck, 2009b).

In CA, the term multimodality is used to "refer to the various resources mobilized by participants for organizing their action -- such as gesture, gaze, body postures, movements, prosody, lexis and grammar" (Mondada, 2014, p. 138). Each of these resources has been the focus of various enquiries.

Phonetic/prosodic resources, for example, are shown to be systematically used in managing various aspects of interaction. Among early work on the phonetics of talk in interaction is Local et al.'s (1986) in which they show systematic phonetic properties in relation to the turn structure in Tyneside English. They present different classes of phonetic features that occur in turn endings. The phonetic features in one of these classes, for example, are a slower tempo towards the end of the turn, a sudden increase and decrease in the loudness on the last stressed syllable, elongation of the last syllable, and a pitch step-up at the end of the turn (see the original work for details). On action formation, Ogden (2006) argues upgrading the prosodic design of a second assessment indicates a strong agreement with a first assessment while downgrading a second assessment is projective of disagreement (on prosody and action formation, see also Cantarutti, 2021; Gonzalez Temer & Ogden, 2021; Couper-Kuhlen, 2004; 2012; Ogden, 2010; Selting, 1996, Szczepek Reed, 2012, among others). On sequence structure, some clusters of phonetic features are shown to signal sequence ending while some others sequence continuation (e.g., Couper-Kuhlen, 2004; Curl, Local, & Walker, 2006; Local & Walker, 2005; Ogden, 2010; Persson, 2013, among others). Borrowing Ogden's (2017) terms, a low and quiet (LQ) cluster is shown to project sequence (or turn) ending while a high and loud (HL) cluster can indicate continuation. Prosody has been also discussed in terms of stance (in rough terms, stance can be understood as participants' positioning towards one another or something, see §5.3 on stance). Reber (2012) presents how affective stance can be reflected in 'sound objects', or for example Pillet-Shore (2012) shows that greetings with a cluster of 'large' phonetic features display a positive stance such as approval towards the coparticipant while a 'small' cluster signals "(no more than) a neutral stance" (p. 375).

Bodily resources have also attracted scholars' attention. Gaze, for instance, is shown to be consequential for sequence structure (Rossano, 2013), for action formation (Kidwell, 2005), for turn-taking (Holler & Kendrick, 2015), and for the organization of preference (Kendrick & Holler, 2017). Among bodily movements, nodding in the context of storytelling is shown to manage (dis)affiliation (Stivers, 2008), pointing can regulate turn-taking (Mondada, 2007), or manual actions such stretching hand towards an object can form social actions such as requesting (Kendrick, 2020) (see Nevil, 2015 for a review on research on embodiment).

As Mondada (2014) notes, studying various resources that are involved in formatting social actions can be approached in two main ways: by focusing on one particular resource (e.g., gaze, pointing, nodding, one aspect of prosody, or one lexical or grammatical feature, etc.) or by focusing on a web of diverse multimodal resources, or what she calls a complex multimodal Gestalt: "a large array of multimodal resources, mobilized and packaged in an emergent, incremental, dynamic way, in response to the contingencies of the setting and of the interaction" (p. 140). While the former allows for investigating a particular resource in terms of "its systematic distribution within sequential environments", it does not explore the interplay among the multitude of resources that are operative all at once (p. 139). Studies on multimodal Gestalts, Mondada continues, has been mainly focused on single case analyses since establishing systematic collections of complex multimodal Gestalts is more difficult (p. 139). This study is a contribution to the literature of multimodality since it explores the complex relationship between a diversity of multimodal resources based on collection of 140 cases.

Increasingly more scholars are considering multiple resources that are simultaneously operative in formatting social actions (e.g., Cantarutti, 2021; Gonzalez Temer & Ogden, 2021; Goodwin, 2018; Goodwin & Cekaite, 2014; Keevalik, 2014; 2018). For example, focused on the token 'mm' in the context of tasting unfamiliar food, Gonzalez Temer and Ogden (2021) report various prosodic designs of the token, the co-occurring bodily movements such as head movements and gaze direction, and the interactional functions that the token performs such as acknowledgement, assessment, continuer. They

conclude that firstly there is no one-to-one mapping between the phonetic and embodied design of the token and its interactional function, and secondly the boundaries of the multimodal resources involved in the production of the token are non-convergent. Focused on pursuit in interaction, this study will show convergent boundaries for multiple resources that are employed in a pursuit turn.

Beyond CA, various disciplines have shown interests in multimodality, for example, gesture researchers (Kendon, 2004; McNeill, 1992) neuroscientists (e.g., Hagoort & Berkum, 2007), and phoneticians (e.g., Loehr, 2012). Piling evidence from these fields suggests that not only are speech and co-speech gestures related, but they also form one integrated system (Goldin-Meadow & Brentari, 2017; Guellaï, Langus, & Nespor, 2104; Kendon, 2014; Perniss, 2018). Such evidence questions the traditional distinction between 'linguistic' and 'non-linguistic' cues. Instead, they suggest the notion of *multimodal language* which incorporates both vocal and visual cues into studying language (Perniss, 2018). Kendon (2014) puts forward a related concept. Based on the observation that kinesic and vocal expressions are essential in speakers' utterances, he proposes "languaging" or "doing language", the study of which goes beyond 'language' as "the formal abstract system of utterances" (p. 1).

Among the diverse findings on multimodality beyond CA, two classes of findings are more relevant to this study. The first one is related to the temporal coordination between speech and co-speech gestures (e.g., Kendon, 1980; Krivokapić, 2014; Loehr, 2012; McNeill, 1992) and the second one is on the communicative function of multimodal features (e.g., Ambrazaitis & House, 2017; Brown & Prieto, 2017; Gibbon, 2009; Gussenhoven, 2004, pp.71-96; McNeill, 1992). Temporal coordination is studied at various levels, for example whether gestural peaks align with prosodic peaks or whether gesture and speech are co-extensive. Loehr's (2012) work on English data shows that gestural apices align with prosodic peaks (pitch accents). Of different prosodic hierarchies (pitch accents, intermediate phrases, and intonational phrases, Nespor & Vogel, 1986), his findings indicate that gesture phrases align with intermediate phrases on the prosodic level. Regarding function, Brown & Prieto (2017) summarize findings from various fields in six categories: Multimodal resources manage information status, turn-taking, epistemic stance, (im)politeness, irony, and speaker identity. The present contribution adds to this line of research firstly by showing a fine temporal coordination between multiple modalities in naturalistic data. The findings illustrate that multimodal Gestalts can manage structuring sequences of social action.

Having reviewed some of the main concepts related to the selected methodology of this work, we now delve into reviewing the literature on the over-arching theme throughout the thesis: conflictual interaction.

1.4. Conflictual Interaction

Let us take as our point of departure Goffman's metaphor of individual as a "vehicular unit" and the rules by which social order is maintained as "traffic code" (Goffman, 1971/2010, pp. 5-6). Traffic code, Goffman notes, provides a "safe passage pattern" by which vehicular units avoid collision (p. 6). When these codes are not adhered to, collision is likely. Looking at it the other way around, interactants in a social encounter – the vehicular units on a road – follow a set of social norms – the traffic codes – and violations of the norms brings about potential collisions. Social norms are moral rules among members of a community that govern their social behavior (Garfinkel, 1967, p. 35).[2] Overall, there is a tendency to minimize collision (Goffman, 2010; Ross, 1901, p. I; Sacks, 1987); nevertheless, collision may occur when what two interactants pursue comes into clash, and as with vehicular incidents, social collisions vary in their scale of magnitude. They can be as subtle as a brief disaffiliative exchange, for example a passing disagreement between two interactants on a 'inconsequential' matter, a mitigated rejection to an invitation accompanied with excuses and expression of regrets, or an implicit declination of a proposal along with accounts. Or they may be more expressive such as a heated argument or dispute,

---

[2] *Moral* is used to refer to principles of acceptable social conduct.

or a fight. Regardless of their magnitude, they all share a common underlying ground: what one party pursues interferes with what the other party pursues. Conversation analysis, with its focus on the detailed sequential unfolding of social events, has great affordance to investigate such social collisions. In comparison, CA analytic toolkit works like the slow-motion effect that are sometimes used to inspect vehicular incidents. The following reviews some CA studies on conflicts or conflictual encounters.

Among early CA work on conflictual interaction is Coutler's (1979/1990) on the structure of arguments, M. Goodwin's series of enquiry on disagreements among black urban children (1980, 1982, 1983), and Maynard's (1985) on early stages of disputes among children (see also Atkinson & Drew, 1979 on interaction in court). Coutler's (1979/1990) study on declarative assertions show that an argument is minimally composed of an assertion followed by a counter-assertion. He explains that following this base adjacency pair, re-assertions are sequence expansive whereas backdowns are terminative. He further notes that explicit backdowns are preferred to backdown-implicative silences (see §1.5 on preference). Goodwin (1983) counts some turn design features of aggravated disagreement among children aged 12-13 in correction sequences. These include partial repeats of a prior assertion with 'aggravated' intonation contour (e.g., falling-rising intonation) or on-record disagreement tokens (e.g., 'No') with no mitigating prefaces. She continues to compare the trajectory of the correction sequences in her data on children with Schegloff et al.'s (1977) data on adults and concludes that while among adults, "movement towards termination of the sequence [occurs] rather quickly", in children's data the sequences are likely to get expanded: corrections embodying pre-disagreements turn into disagreements and then to disputes. We will see in this thesis that sequence expansion towards escalation is also observed among adults and that expansion or termination of the sequences is more the matter of the design of the action than the age of speakers. While Goodwin (1983) explicates how children's arguments unfolds, Maynard (1985) focuses on the initial stages of an argument or a dispute: the "arguable move" – the origins of an argument in the midst of children's ongoing activities which could be nonverbal (p. 3). Then argumentative moves are produced: "initial statements of opposition" that "only contingently turn into an element of argument or dispute episode" (p. 8). He shows that an opposition may be resolved at these early stages, for example, by "letting the opposition pass", so these initial stages do not necessarily end in a dispute (p. 7).

The field's scope of investigation was then broadened: conflictual interaction was explored in various contexts such as multi-party arguments (Goodwin & Goodwin, 1990), mediation (Bilmes, 1992; Garcia, 1991), and radio talk (Hutchby, 1992; 1996) and also in light of different organizations in CA like turn-taking (Hutchby, 1992) and preference (Kotthoff, 1993). In the context of mediation, for example, Garcia shows that the turn-taking system in mediation meetings in which parties in conflict address one another through a mediator is a contributing factor to the resolution of conflict. Again, on turn-taking, Hutchby's (1992) findings suggest that confrontational talk can be manifested in an interactant's bids for turn space in overlap with a co-interactant's talk without any indication of overlap minimization. These studies on conflict talk based on data from interaction among children and from institutional and broadcast contexts have been informative; however, they may not be great representatives of conflictual encounters that happen in daily situations among adults. As Dersley & Wootton (2000) note, in institutional contexts, for example, participants tend to pursue pre-specified objectives which has implication for how oppositions are actualized.

As the field continues to flourish with studies on conflictual talk in institutional settings or particular communities (e.g., Butler, Danby, & Emmison, 2011; Clayman, 2010; Gago, 2013; Heinemann, 2009; Glenn & Kuttner, 2013; Lee, 2019; Pino, 2018; Scott, 2002), more investigations on daily conflicts among adults came to CA horizon around and especially after the turn of the century (e.g., Afshari Saleh, 2020; Clift, 2020; Clift & Pino, 2020; Dersley & Wootton, 2001; 2000; Glenn, 2019; Haugh & Sinkeviciute, 2019; Hoey et al, 2020; Holt, 2012; Lloyd, 2017a; 2017b; Lloyd & Mlynář, 2021; Muntigl & Turnbull, 1998; Potter & Hepburn, 2020; Whitehead, 2015; Whitehead, Bowman, & Raymond, 2018; Yu, Wu, & Drew, 2019).

While all these studies illuminate important aspects of conflictual talk, of especial relevance to this thesis are the ones that inspect episodes of escalated conflicts (e.g., Clift & Pino, 2018; Hoey et al, 2020; Lloyd & Mlynář's, 2021; Whitehead et al. 2018). Some show various resources by which an initial clash gets escalated, and some others concern about possible resources that participants employ to deal with the conflict when it is escalated. Among resources that show an escalated state of affairs is expletive insertion (Hoey et al., 2020). Hoey et al. argue when dealing with some interactional difficulty, expletives are used to pursue a cooperative response after earlier attempts have failed to secure cooperation. Expletives, they show, contribute to an upgraded turn construction and are used after other upgraded turn features, such as 'serious' prosody and upgraded embodied actions, have been employed to no avail. Lloyd & Mlynář (2021) report similar features such as name-calling, swearing, and provocative bodily actions such as a 'wanker' gesture, as both embodying escalation and contributing to further escalation of the conflict talk. Whitehead et al. (2018) discuss that 'risk factors' such as race and class categories that non-situation based studies consider to be potential correlates that may result in violence are indeed turn resources that participants use to initiate or escalate conflicts in situ. In Chapters 2 and 3, we will see that the type of resources mentioned in these studies can be in fact a part of a multimodal Gestalt that contributes to escalating the ongoing disaffiliative state.

The second group of studies relevant to this thesis shows different ways in which participants may respond to or handle a conflictual turn. For example, Holt (2012) studies sequences where a participant is complaining about a third party, but the coparticipant does not fully affiliate with the complaint. Therefore, there is some discord between the two participants about a complainable. She shows that after the complaining party escalates the complaint, the coparticipant may resort to laughing to "[resist] the further development of the complaint" (p. 446). Clift and Pino (2020) focus on a particular practice which they call *conduct formulation*: formulations that "asserts something negative about the recipient's conduct in the prior turn" such as "I dunno why you're being so aggressive," "Don't put words in our mouth," or "why are you shouting at me." (pp. 1-2). As they show, this practice is used by a recipient of an accusation or admonishment at a point where the accusing or admonishing party has made escalated attempts to align the recipient with their project. They argue that the practice "turns the table" on the participant who has been accusing till then; that is, the accuser is accused of some negative conduct. In Chapter 6, we will see a different practice – mock aggression – by which participant handle an unfolding disaffiliative sequence.

But why does disaffiliation need handling? Where can we find evidence that interactants treat disaffiliation as a matter that calls for handling? It is these two questions that we turn to in the following.

1.5. (Dis)affiliation & (Dis)preference

Affiliation has close affinity with social solidarity, social cohesion, cooperation and collaboration among social units. In CA, the term is often used to describe actions or aspects of action composition that supports interactants' social relations, and in contrast, its negative form – disaffiliation – is used to describe actions or action design that weaken social bonds (Lindström & Sorjonen, 2013).

What lies in the heart of most forms of conflict is their potential for undermining social ties or in other words, their disaffiliative potentials. Being pro-social creatures by nature (Tomasello, 2014), human beings typically show an overall tendency to promote affiliation and accord in their interaction with one another and demote disaffiliation and discord (Goffman, 2010; Ross, 1910, p. I; Sacks, 1987). A disaffiliative move in the course of an interaction goes against this overall tendency and as such raises up a flag for a breach of a normative expectation (Garfinkel, 1967). Participants orient to this 'flag' as one that normatively calls for some managing. A great locus to trace the footprints of such tendency and orientations is the organization of preference in CA (Clayman, 2002; Heritage, 1984, pp. 265-269; Kendrick & Torreira, 2015; Lerner, 1996; Nishizaka & Hayano, 2015; Pillet-Shore, 2017; Pomerantz & Heritage, 2013, pp. 210-228; Robinson, 2020; Schegloff, 2007, pp. 58-96). In the following, after briefly overviewing the concept of preference in CA, I explain two observations on preference that

some scholars have (over)generalized to the context of conflict. Then I clarify the approach that I take in this thesis regarding preference in conflictual talk.

The early sparks of the concept of preference in CA flashed in Sacks's lectures (e.g., 1992, pp. 414-15; see also Schegloff, 1992, pp. xxxi-xxxiii). These remarks were then shaped into some early crucial papers, e.g., Sacks & Schegloff (1979), Pomerantz (1984a), and Sacks (1987), which built the groundwork for much of the following CA investigations on preference. What features in all discussions about preference is the asymmetry between some alternative options and different weighting of these alternatives in terms of social norms. For example, in their work on person reference, Sacks and Schegloff (1979) compare different alternatives for referring to people: single versus compound forms and also recognitional versus non-recognitional. They conclude that single recognitionals are preferred over compound non-recognitionals. That is, interactants tend to use single reference terms by which their co-interactants can recognize the reference. Later, Sacks (1987) compares two alternatives 'yes' and 'no' answers doing agreement and disagreements, respectively. Among other observations, he notes that 'yes' tokens tend to be actually produced when a participant does agreement whereas 'no' tokens are likely to be deleted or delayed in producing a disagreeing response. These observable features make public the *collective expectations* of members of a society on some underlying rules of conduct, for example, if possible, minimize disagreement. In this context, then agreeing is the preferred alternative and disagreeing the dispreferred option (Pomerantz, 1984a). So preference is not about an individual member's psychological desires. It is about cultural principles (Pomerantz & Heritage, 2013) and "how people systematically design their actions to either support or undermine social solidarity" (Pillet-Shore, 2017, p. 2). Overall, there has been this consensus that preferred actions are typically pro-social and supports members' social bonds whereas the reverse is true for actions that weaken social relations (e.g., Heritage, 1984, pp. 265-269).

Preference came to the scene of CA with the observation of two types of tendency (which was later 'over'-generalized by some scholars to conflict contexts). The first observation is on the relationship between actions and their design and the other one on the frequency of preferred actions. The first issue is summarized in Heritage's (1984) conceptualization of preferred and dispreferred actions in an adjacency pair structure: "Actions which are characteristically performed straightforwardly and without delay are termed 'preferred' actions, while those which are delayed, qualified and accounted for are termed 'dispreferred'." (p. 267). He notes that between the two broad alternative responses to requests (granting and rejecting), granting is often performed straightforwardly, that is with no hesitation, accounts, apologies and the like. Rejecting, on the other hand, is characteristically delayed or accompanied with extra work such as accounting and apologizing. So granting is the preferred and rejecting the dispreferred alternative. The 'non-straightforward' design of the rejection responses reveals participants' orientations to some underlying norm: "If possible, minimize stated rejections of requests" (Pomerantz & Heritage, 2013, p. 210; Clayman, 2002; Pillet-Shore, 2017).

The second observation is about the higher frequency of preferred actions compared to dispreferred alternatives. This is first mentioned in Sacks and Schegloff's (1979) and Sack's own (1987) writings when they mentioned the higher frequency of the preferred alternative. For example, in the context of agreement ('yes') versus disagreement ('no'), Sacks (1979) states "[T]he blandest look would say that if you examine only answer turns, then 'yes's' are a lot more frequent than 'no's' are." (p. 11). Or in the person reference paper, the writers use descriptors such as the "overwhelming" or "heavy" use of preferred alternative (pp. 17-18).

Soon some scholars examined the above observations in the contexts of conflicts (Bilmes, 1988, p. 175; Garcia, 1991, p. 821; Gruber, 1998, p. 477; Kotthoff, 1993, p. 203), and concluded that in conflict talk, for example disputes, the preference structure is reversed. That is, for example, disagreement is preferred over agreement. The basis for such conclusions is not always explicitly elaborated, but Kotthoff (1993) bases her argument on the preference for disagreement in conflict talk on the frequent

act of disagreement without 'reluctance markers'. But this conclusion does not sit well with the consensus that preferred actions are pro-social (see for example Clayman, 2002; Heritage, 1984, pp. 210-228; Pillet-Shore, 2017; Pomerantz & Heritage, 2013). It goes against a long-held agreement on humans' moral expectations and their tendency on maintaining social order: "The members of an orderly community do not go out of their way to aggress upon one another. Moreover, whenever their pursuits interfere they make the adjustment necessary to escape collision, and make it according to some conventional rule" (Ross, 1910, p. 1). If disagreement is the preferred alternative, we should not see any orientations in the behavior of participants that something is happening against expectations. But, for example, the very fact that some participants attend mediation meetings to resolve their conflict shows that they are orienting to their disagreement as an action contrary to cultural norms, and that cultural norm is the tendency to reduce conflicts (see also Dersley & Wootton, 2000, questioning the claim on the reversal of preference structure in institutional talk).

Two sets of untangling can perhaps resolve this clash: untangling action from its design and moral expectations from statistical expectations. One of developments on the first issue is manifested in, for example, Kendrick & Torreira's (2015) work in which they separate action from its design in terms of their preference status. They show that a preferred action (e.g., accepting an invitation) may be performed in a preferred design (e.g., straightforwardly) or dispreferred design (e.g., with hesitations and so forth). Therefore, rather than concluding that in conflict talk, disagreement is preferred, it can be said that interactants perform a dispreferred action (disagreement) in a preferred design (straightforwardly).

The second untangling concerns the types of expectations. Is it moral expectations that matters in the organization of preference or statistical expectations? Put differently, is the underlying issue normativity (moral expectations) or normality (statistical expectations)?[3] It seems that it is the cultural expectations and principles that are the central issue. This is especially clear in Heritage and Pomerantz's (2013) notes where they explain preference in terms of culturally shared principles and the "orderly ways of speaking and acting that are produced in accord with those principles" (p. 210). Sacks and Schegloff's (1979) writing alludes to both types of expectations. Note how they translate the social rule of conduct that is manifested in the preference of single reference use: "On occasions when reference is to be done, it *should* preferredly be done with a single reference form" (p. 16, emphasis added). Here, the modal 'should' tacitly shows the moral expectations underlying the choice of the preferred alternatives. But, as mentioned earlier, they also mention the higher frequency of the preferred alternatives. Indeed, in a non-conflict context, moral expectations are likely to be translatable to statistical expectations. But when it comes to conflict talk, the untangling is especially needed. Does the fact that disagreement is normal in conflict talk make it normative? Despite Sacks and Schegloff's early allusions to frequency counts and the follow-up interpretations that some scholars have made based on that (see Griffiths, Merrison & Bloomer, 2010, p. 8, on preference being a matter of probability), it seems that it is normativity and collective cultural principles that plays a central role. There are other types of orientations than 'straightforward' versus 'non-straightforward' design of the turns and frequency counts that can reveal participants' orientations to disagreements as a culturally dispreferred option (Pillet-Shore, 2017 for a review; see §1.6 below).

The question remains: if the design of an action does not necessarily indicate whether or not the action is preferred, where should we start? Where to find participants' orientations to an action as preferred or dispreferred? The field perhaps needs some progress in this regard especially when it comes to initiating actions (but see e.g., Pillet-Shore, 2010; Kendrick & Drew, 2014; Robinson & Bolden, 2010), but about responsive actions, Schegloff's (2007) later characterization is noteworthy:

---

[3] I'm grateful to Paul Drew for bringing to my attention the distinction between normativity and normality, and for our insightful discussion around the issue that normality of a pattern does not necessarily make it normative.

> Sequences are the vehicle for getting some activity accomplished, and that response to the first pair part which embodies or favors furthering or the accomplishment of the activity is the favored – or, as we shall term it, the preferred – second pair part (p. 59).

And this is the conceptualization that I will rely on in this thesis in my analysis of actions as preferred or dispreferred. That is, in the sequences under examination (i.e., sequences that include initiating imperatives after which granting or rejecting are two broad relevant next actions), granting is the preferred next action since it furthers the course of action that the imperative has initiated, and rejection is the dispreferred alternative as it hinders the progress of the initiated course of action (Clayman, 2002; Pillet-Shore, 2017; Stivers & Robinson, 2006, on the preference for progressivity in interaction). Put differently, even though in this collection rejecting is more frequent than granting and although they tend to be performed in a 'straightforward' design, they are still analyzed as the dispreferred alternative. Participants' emic orientation towards rejection as the dispreferred alternative is revealed when they hold the rejecting party accountable for hindering their course of action. We will turn to the issue of accountability in the following.

1.6.   Accountability & Negative Sanctioning

In this thesis, the concept accountability is used as a type of orientation that participants display towards each other's actions. Generally speaking, in a dyadic interaction, there are two main types of orientations that participants can reveal towards a particular action: the orientations that the performer of the action makes public towards their own action, and the ones that the coparticipant displays. The literature on preference in CA has mainly focused on the former: the performer's treatment of their own conduct manifested in the action's constructional features, for example, whether the action is performed straightforwardly or with hesitations, delays, and so forth. But in the next slot, the co-interactant can also make public their treatment of the just performed action. As is established in CA, it is through the co-interactant's orientations that intersubjectivity is maintained or cultural norms are re-enforced. In the next slot after an action, the co-interactant can reveal that something contrary to normative patterns of conduct has just happened or in other words a dispreferred action has just been performed.

There are various ways that co-interactants can reveal their orientation towards an action as awry to expectations, one of which is holding the performer of the action accountable. Before reviewing some studies on accountability in social interaction, a brief note on the double use of the term accountability in CA is worthy of mention (for an extensive discussion on this, see Robinson, 2016, pp. 1-46). The term account-ability, as Robinson (2016) puts it, is sometimes used to refer to the intelligibility of a social action: "interlocutors' abilities to form and ascribe (i.e., recognize and understand) possible actions" (p. 11). Thus, in this sense, the term is closely related to intersubjectivity, action formation and ascription (Levinson, 2013). In another sense, accountability is used to refer to the responsibility of participants for their actions (Robinson, 2016, p. 12). In this usage, it is tied to moral expectations and normative patterns of conduct. That is, participants are accountable (responsible) for their deviation from normative expectations. It is this second sense that is intended throughout this thesis.

The idea that interlocutors in a social interaction hold one another accountable for deviating from rules of conduct can be traced back to Garfinkel's (1967) breach experiments. Through those experiments, Garfinkel shows that an interactant's breach of the mutual 'trust' – violations of some agreed upon cultural norm – warrants *negative sanctioning*. Social sanctions, according to Goffman (1971/2010) are "techniques for ensuring conformance" with social norms (p. 95). Goffman categorizes sanctions in two broad groups: negative and positive sanctions. The former is "penalties for infraction" and the latter is "rewards for exemplary compliance" (p. 95). Negative sanctioning is one recurrent concept in this manuscript; it is used as evidence behind some of my main arguments. Given that we do not deal with positive sanctions in this thesis, throughout the manuscript, the term sanction(ing) is a shorthand for *negative sanction(ing).*

Sanctions can be formal or informal. Formal sanctions are "administered by specialized agents designated for the purpose" such as imprisonment (Goffman, 1971/2010, p. 347). Informal sanctions are performed locally by the parties in an interaction themselves. They can be as subtle as a disapproving look (e.g., Kidwell, 2005) to more explicit ones such as soliciting accounts (Bolden & Robinson, 2011). Subtle or not, negative sanctions are the flags raised by cultural members when social transgressions or deviations from normative patterns of conduct have occurred. Through negative sanctions participants highlight red lines and re-enforce normative expectations.

CA studies have reported various ways for sanctioning deviations from norms: extreme case formulations (Edwards, 2000); sanctioning looks (Kidwell, 2005), repairs (Schegloff, 2005), polar interrogatives that are unanswerable (Heinemann, 2008); soliciting accounts with why-interrogatives (Bolden & Robinson, 2011); grammatical conditionals and negative assessments (Hepburn & Potter, 2011); imperative directives issued after the relevance of directed actions (Kent & Kendrick, 2016); factual declaratives doing criticism (Rossi, 2018); 'shaming' interrogatives (Potter & Hupburn, 2019), conduct formulations (Clift & Pino, 2020), mock aggression (Afshari Saleh, 2020, Chapter 6); and others. Among these, Kent & Kendrick's (2016) work holds some especial similarity to this investigation in that in their study it is resistance against a directive that is treated as sanctionable. Here, too, recipients are held accountable for their resistance against completing imperative-speakers' CoAs. Another similarity is that they too focus on the imperative structure as means of getting the recipient to perform an action, and it is this syntactic structure that is reviewed next.

1.7. Imperatives

Imperative is one of the major sentence types in many languages (König & Siemund, 2007, pp. 276-324; Sadock & Zwicky, 2007) including Persian (Farshidvard, 2009). Although imperatives can be used to do a variety of social actions (see for example, Aikhevanvald, 2010, pp. 198-199; Kendrick, 2020, p. 114), they are widely known as a grammatical format which gets the addressee to do something (Aikhenvald, 2010; Sorjonen, Raevaara & Couper-Kuhlen, 2017 on imperatives across languages; see also Floyd, Rossi & Enfield, 2020 for various methods of getting a co-interactant to do an action, including imperatives, across a range of languages). This can be in part under the influence of Searle's (1975) speech act theory in which he proposes that there are five basic types of action: representatives, directives, commissives, expressives, and declarations. Among these, directives are classified as a group of actions such as ordering, commanding, requesting, begging, instructing, that attempt to get the interlocutor to do something (Searle, 1975). This class of action was then associated with imperative utterance type: "The imperative indicates the speaker's desire to influence future events. It is of service in making requests, giving orders, making suggestions, and the like" (Sadock & Zwicky, 1985, p. 160).

CA, however, has shown that the precise function of an imperative depends on its sequential position and other interactional details. Imperatives that initiate a course of action are shown to be functionally different from responsive imperatives or those that are issued after the directed action is performed. For example, in Heinemann & Steensig's (2017) study, responsive imperatives that are produced after requests for permission function to grant permission (see also Zinken & Deppermann, 2017). Based on data on sparring sessions, Okada (2018) shows that a coach's imperatives that follow the performance of the relevant action by the trainee "articulates the rule guiding when to choose the targeted boxing move" (p. 81). These sequential and interactional observations have motivated CA scholars to cast a more detailed look at this sentence type and go beyond the general characterization of 'getting someone to do something'.

Imperatives are studied from two broad perspectives: on a more macro-level, how observations of face, rights, power, and social distance or relationship among individuals may motivate the use of this sentence type (on face: Brown & Levinson, 1987; on entitlement and deontics: Antaki & Kent, 2012; Craven & Potter, 2010; Frick & Palola, 2022; Lindström 2005; Stevanovic, 2017). The second perspective involves more micro-level details of the interaction as it unfolds (Deppermann, 2018; Kim

& Kim, 2020; Okada, 2018; Rossi, 2012; 2017; Taleghani-Nikazm, et al., 2020; Zinken & Ogiermann, 2013; Kendrick, 2020, pp. 12-123). These more local factors include, but not limited to, the urgency of the action and its relation to the overall ongoing activity (see Sorjonen, Raevaara & Couper-Kuhlen, 2017, pp. 1-23, for four pragmatic dimensions of the situations in which an imperative may be used).

The first group include, for example, Craven and Potter's (2010) study on directives in parent-child interaction. Building on Curl and Drew's (2008) work on various degrees on entitlement and contingency in different requests formats, they compare imperative-directives with other type of directives and conclude that imperative-directives display "high entitlement to direct the other speaker and little or no orientation to the contingencies on which the compliance with the directive may rest" (p. 426). They also show that when the child resists complying, the "directive is upgraded by using constructional features such as volume, lexical choice and intonation. Then extending Craven and Potter's (2010) work, Antaki and Kent (2012) argue that in interaction between care-home staff and clients with intellectual disabilities, imperatives are the main linguistic format used by the care-home staff to get the client to do an action and this can be in part motivated by the kind of relationship between the two parties, the imperative-speaker being responsible for the recipient.

The more micro-level factors include orientations to the urgency or immediacy of the directed action (Mondada, 2017; Raevaara, 2017), to accountability (Deppermann, 2018; Kent & Kendrick, 2016; Taleghiani-Nikazm, 2020), to compliance and availability of the recipient (Bolden, 2017), and to the compatibility of the action with the overall ongoing activity (Rossi, 2017; Wootton, 1997; Zinken & Deppermann, 2017). On timing and urgency, for example, Mondada (2017) shows that in French, an imperative structure can be used when a timely performance of an immediate action is consequential to the success of ongoing activity. Imperatives may be issued after an interactional failure of some sort, in which the imperative may be oriented to the recipient's accountability. For example, issuing an imperative after "the relevance of the directed action not only direct the recipient to act but also find fault in his or her current actions or inaction" (Kent & Kendrick, 2016, p. 273). Similarly, in German driving lessons, one function of an imperative structure is suggested to be correcting fault in the recipient's ongoing action (Deppermann, 2018). Also, in card game interactions, Taleghani-Nikazm et al. (2020) argue, an imperative format functions to resume a halted game when the recipient is already oriented to the game but fails to perform the next relevant action. On compatibility, Rossi (2012), building on Wootton's (1997) work, argues that in Italian "actions that are integral to an already established joint project between requester and recipient" are likely to be requested through imperatives (p. 426). In a later paper (Rossi, 2017), he adds a secondary use of Italian imperatives to do requests that are "*compatible* with the current situation of the requestee, either because the requested action "piggybacks" on what the requestee is currently doing, or because the requestee is momentarily doing nothing" (Rossi, 2017, p. 104, italic original; see also Zinken & Deppermann, 2017, pp. 27-63 on a similar usage in Polish and German).

This thesis is different from the above-mentioned studies on imperatives in that although imperatives feature in the sequences that we will analyze, their interactional functions or the interactional details that warrant their use is not our main concern. Rather, this syntactic structure is the criterion observed in the collection building procedure. I will return to this point in the Data section (§1.9) where I explain the scope of the collection, but before that an overview on Persian and Persian imperatives is due.

1.8.   Persian Grammar and Persian Imperatives

This section outlines some of the grammatical features of Persian that helps the reader make sense of glossing in the transcripts and the analysis of extracts in the forthcoming chapters. It is not intended to be a comprehensive review of Persian grammar. Persian has some specific or rare features, such as a high frequency of light verb constructions (Megerdoomian, 2020, p. 472) and a wealth of affective tokens, interjections, or idiomatic phrases, some of which have not even found their way to dictionaries.

In the following, first some of these overall features of Persian grammar are lightly touched on (§1.8.1). Then, we briefly see the morphological structure of imperatives in Persian (§1.8.2).

1.8.1. Some General Properties of Persian Grammar

Before engaging with imperative structure in Persian, let us review some notes on a few general grammatical properties of Persian that better clarify the transcripts in the analytic chapters. These properties are related to sentence word ordering, light verb construction, objectless transitive verbs, sentence subjects, and particles.

Word ordering

The canonical word order of Persian sentences is reported to be SOV by some scholars (Karimi, 2018; Mahootian, 1997). Some others, however, have mentioned SVO to be the 'natural' word ordering in an informal context (Jasbi, 2020, pp. 135-136). These different views may stem from the fact that Persian is among languages with 'scrambling' syntactic property in which "phrasal categories may appear in different positions" in a sentence without affecting the meaning or grammaticality of the sentence (Karimi, 2018, p. 165). Therefore, not only SOV or SVO, but also a whole range of various ordering of words, involving other syntactic categories than subject, verb, and object, yield grammatical sentences in Persian. Take the following examples from this naturalistic corpus:

Extract 1.2. RAY010b, 13.19

```
01    MAS   khast-om      be khoda   ma-am.
            tired-be.1SG  to God     1SG-too
            adj.v         adv        s-adv
            tired am {I swear} to God I too.
```

Consider the word ordering reflected in the third gloss line. Besides adverbs that are inserted in the middle and end of the sentence (*be khoda* and *am*), the scrambling property is especially reflected in placing the subject pronoun *ma* 'I' almost at the end of the sentence.[4] This is so while prescriptive Persian grammarians, as mentioned above, claim that Persian word ordering is SOV. Those who consider 'informal' settings have considered it to be SVO. In both models, the subject is pre-supposed to appear at the beginning of a sentence, which is clearly not the case in our example here. Consider another example:

Extract 1.3. RAY014c, 12.03

```
01    MIN   tekun      na-de    mâmân      saret-o?
            movement   NEG-give  Mom        head-2SG-OBJ.mar
            head.of.v  v         vocative   obj
            Don't move Mom your head?
```

Whereas some Persian grammar books such as Anvari and Ahmadi Givi's (2011) presume Persian vocatives are placed at the beginning of a sentence, in this example from spoken Persian, the vocative *mâmân* 'Mom' is placed between the light verb construction *tekun nade* 'Don't move' and its object *sareto* 'your head' (more on light verbs, below). Here is another example:

Extract 1.4. RAY048a, 05.28

```
01    ROZ   sefra-ro           cherâ         injuri      andakhtin.
            Food.cloth-OBJ.mar  why           like.this   throw.2PL
            obj                 question.word adv         v
            Food cloth⁵ why have you like this set up.
```

---

[4] *Ma* in this speaker's dialect is a reduced form of *man* 'I'.

[5] Food cloth is a piece of cloth or plastic on which food is served. It is spread on the ground and people sit around it to have food.

Again, the question word *cherâ* 'why', which is traditionally assigned the first slot in interrogative sentences is placed after the object in this example.

Traditionally, deviations from the 'default' SOV word ordering are attributed to 'focus' or 'topic interpretations' (e.g., Karimi, 2018, p. 165); however, this seems to be too simplistic especially for spoken Persian: for one thing, the 'default' SOV word ordering seems to be based on written Persian or prescriptive grammar books; the dominant word ordering for spoken Persian requires rigorous frequency counts for all sentence types based corpora of naturally occurring interaction. For another thing, without considering the prosodic properties of a sentence, associating word ordering with broad analytic categories such as 'focus' does not seem to give a complete picture of what is involved. I have refrained to base the analysis of the collection on (over)generalizations such as SOV or SVO models. Where relevant, I have considered the ordering of the components of utterances in light of other design features such as the prosodic, multimodal, or other interactional details of the talk.

Light verb construction

Light verbs are the verb component in compound verb constructions that "neither retain their full semantic predicational content, nor are they semantically completely empty. Rather, they appear to be semantically light in some manner that is difficult to identify" (Butt, 2003, p. 1). In Extract 1.3 above, *tekun nade* 'Don't move' (literally 'movement don't give') is a light verb construction consisting of a noun *tekun* 'movement' and a negative form of the light verb *nade* 'don't give'. Light verb constructions are highly frequent in Persian (see Megerdoomian, 2018, p. 472) and they are a great locus of analytic work in terms of both theoretical syntax and interaction, but the main point in this thesis is not to read and interpret the light verb component separately from its 'head' (in the above example *tekun nade* 'movement don't give', the noun *tekun* 'movement' is the head and not the object of the verb).

Karimi (2018, pp. 169-170) gives a list of some of verbs that frequently function as 'light' in Persian compound verbs, for example, verbs out of infinitives *kardan* 'to-do', *dâshtan* 'to-have', *shodan* 'to-become', *dâdan* 'to-give', and so forth. The following is just a few examples from the pool of light verb constructions that you will face in the forthcoming chapters.

- ***garm kon-in***
  warm IMP.do-2PL
  Warm it up (second person PL)

- ***bus   kon***
  kiss IMP.do-Ø
  Kiss (second person SG)

- ***kâr   dâr-e***
  action have-
  He/she/it takes (third person singular)

- ***gonah dâr-e***
  sin   have-3SG
  Pity him/her/it (Idiomatic expression to show sympathy)

- ***hasudi-t   shod***
  jealousy-2SG become-Ø →
  you were being jealous

Given the high frequency of this construction in the extracts in this thesis, I will not comment on these verbs as light verb compound constructions so that we remain focused on the analytic point.

Objectless transitive verbs

Persian has both transitive and intransitive verbs. The main point on transitivity to make here is that the object of a transitive verb can be omitted in spoken Persian if it can be inferred from the context. In the following extract, the object of the verb *bokhor* 'Eat' is clear through the speaker's embodied conduct.

Extract 1.5. RAY045a, 03.03

```
01   FAR       +bo-khor.#
               IMP-eat.2SG
                eat.
            -->+hold out a fig towards AMI-->>
         fig              #1.12
```


#1.12

```
02            (0.5)

03   AMI   +((click))+[anji dust na-dâra-m.
                       fig  like NEG-have-1SG
            ((click)) I don't like figs.
```

The object of the imperative verb *bokhor* 'eat' can be inferred from the context: the speaker is holding out a fig towards the recipient (line 1). This is evidenced in the recipient's response where he displays his understanding of the object of the verb/offer to be the fig (line 3).

Subject

Persian is a pro-drop language (Ghaffari, 2020, p. 554; Mahootian, 1997, p. 5). This means that a sentence is grammatical even without an overt grammatical subject. Having a rich person/number inflection system, the subject of the verb can be disambiguated through the verbs and the context. Consider the following extract. In the first turn (line 1), the subject pronoun is produced while in the second turn (line 2), it is dropped.

Extract 1.6. RAY010b, 13.19

```
01     SUF   m:an   khaste-am=hehehehe[hehe
             1SG    tired-be.1SG
             I'm tired

02     MAS                           [kheili porru-i-â.
                                      very   cheeky-be.2SG-PRT
                                      {you} are too cheeky.
```

In line 1, *man* 'I' is a first-person singular pronoun; the verb is the first person to-be verb *hastam* which is reduced to *am*, similar to English first-person to-be verb 'am'. This same turn could have been produced without the subject as *khaste-am* 'tired-am'. In Persian the sentence 'tired am' is grammatical and in fact many sentences are produced without a subject. Consider the second turn, for example. In line 2, the sentence is produced without the subject pronoun *to* 'You'. This subject is reflected in the verb: *hasti*, which is reduced to *i* 'be.2SG, translated to 'are'. Note that because English is not a pro-drop language, in the translation line, the subject pronoun 'You' is inserted. This difference between Persian and English can sometimes result in some confusion especially when discussing agency. In

agency discussions, the main issue is *who* performs an action. The way *who* is reflected in Persian and English is different. I will return to this issue in one of the extracts in Chapter 2.

Interjections, particles, and idiomatic expressions

Spoken Persian is rich in terms of the variety of interjections, particles, and formulaic expressions used by speakers. These often show speakers' stance (Du Bois, 2007; Englebretson, 2007, Jaffe, 2009, Wu, 2004, on stance) especially affective stance, or perform affective actions such as expressing love, sympathizing, criticizing, and the like (see Farshidvard, 2009, pp. 314-326). Interjections are those vocalizations that are conventionally associated with particular social actions; for example the interjection *bah* is used responsive to pleasant stimuli such as good food or nice weather, and displays pleasure (Afshari Saleh & Ogden, in prep). Sometimes these interjections have a variety of functions. For example, *eh* can be used responsive to unexpected stimuli, displaying surprise, or responsive to unpleasant stimuli, displaying annoyance, disappointment, and the like. Here is an example. Faraneh has asked her husband, Ebi, to perform an action. Ebi has rejected Faraneh's directive and provided an account for the rejection. Line 1 is the second piece of account that Ebi is offering. In overlap, Faraneh vocalizes the interjection *eh* followed by an 'offensive' assessment of her husband *nekbat* lousy.

Extract 1.7. RAY020a, 16.55

```
    01   EBI    [bâyad be-r-am      jalase
                must  SBJV-go-1SG meeting
                I must go to a meeting.


->  02   FAR    [eh    nekbat
                INTJ lousy
                eh lousy
```

Here the interjection *eh* displays the speaker's negative stance, for example annoyance or disappointment, about her husband's rejection. It is not always easy to find an English equivalent interjection, in which case, in the forthcoming chapters, I have just reproduced the Persian interjection itself on the translation line and provided some explanation on the overall function of the interjection.

By particle, I mean lexical items that in some contexts have content meaning, but in some others, they function to display some (affective) stance. Take the following example. The literal meaning of the word *âghâ* is 'sir' or 'man'. But here it is in the service of displaying some affective stance. Minuvash is addressing her *sister* – a female participant – with the particle *âghâ* and then announces what their pet cat is doing 'it's sleeping next to the heater'. This particle functions as a vocative here, but a vocative that is emotionally loaded.

Extract 1.8. R014a, 05.33

```
    01   MIN    âghâ kenâre shufâzh     khâbid-e.=      [gonâh dar-e.
                sir  next.to-Ez heater asleep-be.3SG   sin   have-3SG
                âghâ it's sleeping next to the heater.= Poor thing.
```

The word *âghâ* here is not used in its common dictionary meaning as 'sir' or 'man': for one thing, the speaker is addressing a female co-interactant; for another, even if the addressee was male, the interaction is between two family members and the 'formal' vocative 'sir' is not normatively relevant. This lexical item is in service of some other interactional function. Persian particles like this are hugely understudied. Pinning down what interactional functions they perform in the midst of an interaction requires systematic scrutinization of them in various contexts. Given that they are not the focus of this manuscript, I have refrained from translating them. Wherever possible, I have mentioned the overall affective stance that is displayed through them, but more specific and analytically grounded accounts should be the aim of future investigation. In any case, the analytic point made in the extracts in which these particles appear does not depend on the particles' function or meaning.

Persian is also special in terms of the vast variety of formulaic expressions, many of which are affectively loaded. Take, for example, the formulaic expression *gonâh dare* literally, 'it {the cat} has sins' in the extract just above (Extract 1.8). The literal meaning of the lexical components of this TCU is not intended here; rather, the TCU is an expression used to display the speaker's particular affective stance, in this case, about the cat that is sleeping next to the heater. For these expressions, I have provided a free translation (as opposed to literal translation). Attempt has been made to find English idiomatic expressions that English speakers may use in similar interactional contexts.

1.8.2. Imperatives in Persian

Having reviewed some general properties of Persian grammar, we turn to the structure of imperatives. Imperative verbs are the basic building block of the imperative sentence type. Imperative sentence type is one of the four main sentence types in Persian: imperatives, declaratives, interrogatives, and exclamatory (Farshidvard, 2009, p. 66; Anvari & Ahmadi Givi, 2011, p. 321). Here, we examine imperative verbs in terms of their morphological structure: imperative prefix and person/number suffix. We also consider imperatives in terms of their grammatical subject, and finally we have a brief look at T/V distinction in Persian.

Morphological structure: Imperative prefix

For many Persian verbs, the imperative mood is conveyed morphologically through the prefix *be* [be] added to the *present stem* of the verb (Anvari & Ahmadi Givi, 2011, p. 66). The prefix *be* is phonetically realized differently as *bâ* [bɑ], *bo* [bo], and *bi* [bi] depending on its phonological context. Here are some examples of imperative verbs for second person singular. The imperative morpheme is shown as IMP in the gloss line. Singularity is shown by SG, second person is shown by the numeric 2.

Table 1.1. Prefix 'be' in second person singular imperative verbs.

Person	Second person singular
Structure	*be*–present stem–∅ (no person/number suffix)
Example	• **be-de** IMP-give *give.2SG* • **bo-ro** IMP-go *go.2SG* • **bi-â** IMP-come *come.2SG*

For some verbs, the prefix *be* can be dropped. The imperative form of these verbs is simply the *present stem* of the verb for the second person singular.

- ∅.**sho**
  ∅.become
  become.2SG

In compound verb constructions, the prefix *be* is frequently dropped, but if it is used, it is added to verb component:

- **garm kon**
  warm do
  warm up.2SG

- **garm bo-kon**
  warm IMP-do
  warm up.2SG

Syllable *be* in Persian has other prefix functions as well: it can be used to make other moods for verbs such as the subjunctive mood. It is sometimes added to nouns to make particular adjectives and so forth. Therefore, not every affix *be* in the data is an imperative maker.

Morphological structure: person & number inflection

Persian verbs have rich person and number inflection system. Some scholars have postulated imperatives for all persons: first, second, and third persons, both singular and plural (Anvari & Ahmadi Givi, 2011, p. 66), but the collection under investigation only includes second person, mostly singular but also a few examples of plural. So here I focus on verb inflections for second person singular and plural imperatives.

For the second person singular, as shown above, there is no verb inflection in the imperative mood. Put differently, the prefix *be* added to the *present stem* on its own is indicative of an imperative verb for second person singular. For the second person plural, however, the suffix *in* is added. In spoken Persian, this suffix (a) can affect the phonetic realization of the preceding stem and (b) it itself is phonetically realized variously in different phonological contexts.

Table 1.2. Person/number affix for second person plural imperative verbs.

Person	Second person singular	Second person plural
Structure	*be*–present stem–∅ (no person/number suffix)	*be*–present stem–person/number suffix *in*

Examples	• **be-de**
	  IMP-give
	  give.2SG
		• **be-d-in**
		  IMP-give-2PL
		  give.2PL

	• **bo-ro**
	  IMP-go
	  go.2SG
		• **be-r-in**
		  IMP-go-2PL
		  go.2PL

	• **bi-â**
	  IMP-come
	  come.2SG
		• **bi-â-yn**
		  IMP-come-2PL
		  come.2PL

The first column was discussed above. Here only the second column on second person plural is explained. In the first two examples, the suffix *in* has made the preceding stem lose its vowel on the phonetic level. For example, in *bedin,* the present stem *de* 'give' has lost the vowel [e]. In the last example *biâyn*, [in] is phonetically realized as [ɑɪn] with Romanized transliteration *yn*.

Imperative turns with and without subjects

In §1.8.1, I mentioned that Persian is a pro-drop language: the subject of the sentence can be dropped without harming the meaning or grammaticality of the sentence. In this collection a great majority of imperative turns are constructed without a subject; however, Persian imperatives can be produced *with* a subject. Consider the following example:

Extract 1.9. RAY048b, 08.34

```
01    EBI    amir amir to  bi-â         injâ  be-shin,
              Name Name 2SG IMP-come.2SG here  IMP-sit.2SG
              Amir Amir you come sit here,

              ((8 lines omitted))

09    EBI    bi-â          injâ  be-shin
              IMP-come.2SG here  IMP-sit.2SG
              come sit here
```

In line 1, After the vocative *Amir*, the subject of the imperative verb is produced: *to* 'you' even though the verb on its own conveys that the subject of the verbs *biâ* and *beshin* is second person singular. In

the second rendition of the same imperative verbs, however, the subject is dispensed with. Future investigations should address the interactional grounds based on which the subject of an imperative verb is retained.

T-V distinction

Persian is among languages with a pronominal T-V distinction (after *tu* 'you.SG' and *vos* 'you.PL' in Latin). In some languages with a T-V distinction, a singular co-interactant is sometimes addressed with a plural pronoun. Persian is one of these languages (Nanbakhsh, 2012). This usage of the pronouns is traditionally attributed to factors such as power, solidarity, politeness, or overall, the formality of the context across a variety of languages with a T-V distinction (Brown & Gilman, 1968; Brown & Levinson, 1987; Clyne et al. 2006, p. 284). However, this is suggested to be a simplistic explanation (Afshari Saleh, 2022). In the forthcoming chapters, there are a few instances of second-person plural imperative verb addressed at a singular co-interactant. In none of those instances do various possible interpretations on the T-V distinction have a bearing on the analytic point and, therefore, they are not commented on. Nevertheless, this explanation is provided here to rectify any potential confusion for the readers when interpreting the extracts.

1.9. Data

Having covered some features of Persian language, we now address the data based on which this study is conducted: the overall corpus, participants, the collection, and the transcription conventions.

1.9.1. Corpus

The overall database for this study consists of 55 hours of video recordings of naturalistic face-to-face interaction among Persian speakers in Iran, obtained in 53 recording sessions. The recordings were made during my field trips to Iran in 2018 and 2019. The cameras and audio recording equipment are as follows: a wide-angle lens Zoom video camera and a Canon EOS 5D Mark II camera for video recordings and Tascam DR-10L Audio Recorder with Lavalier Microphone for audio recordings. Separate microphones attached to each participant provided high quality audio, making acoustic phonetic analysis possible. Most recordings were made using one camera only, but for some, where different angles were needed to capture the physical setting better, two cameras were used. Recordings were made at the participants' own residence while they were involved in their daily activities such as having their afternoon tea, having food, preparing meals, playing card or board games, watching TV, and so forth. For some recording sessions, I set up the camera and microphones and asked the participants to turn them off whenever they felt like stopping. For some others, the participants were given the equipment and necessary instructions on starting and ending the recordings. This second option provided some participants more convenience and allowed more flexibility in the time that recordings were made. Informed written consents were obtained prior to each recording.

1.9.2. Participants

Participants in each recording are family members or intimate friends. There are 70 different participants, ranging from 6 to 84 years old. Except for three children (age 6-10), and six late teenagers (age 17-19), the rest of the participants are all adults. Participants speak standard Persian. The number of participants in each recording vary. In the majority of recordings there are two to four participants. In three recordings, the number is up to 12. Reporting the exact number of participants per recording is not straightforward since, involved in their daily tasks and plans, some participants join in the middle of the session and then some other leave, so the number varies even during one single session.

1.9.3. Collection

The findings of this thesis are based on two collections. One on an embodied phenomenon which was later called *mock aggression.* The second is a collection of pursuit.

Prior to data collection, I did not have any a priori research focus. The initial examination of the corpus brough to my attention a particular embodiment recurrent in the corpus, which I later named *mock aggression* – ostensibly aggressive embodiments which are not designed, or oriented to, as physically threatening (Afshari Saleh, 2020; this thesis, Chapter 6). The corpus was then scrutinized with attention focused on the participants' body movements to collect instances of mock aggression, Initially, 40 instances of such embodiments were collected and analyzed.

One of my observations on the mock aggression phenomenon was that in half of the collected instances, it occurred in sequences in which a social action is repeated despite the co-interactant's resistance against it. Therefore, a new collection that included repetition in the face of resistance was built. The occurrence of mock aggression was not an inclusion criterion in this phase.

This second collection was narrowed three times: based on the overall action, then based on the grammatical structure, and finally based on the repetition of the grammatical structure. Firstly, given the vast variety of sequences in which an action is repeated, I narrowed the scope of the collection based on the overall action: pursuing compliance with a directive. This excluded sequences in which disagreements, insults, complaints, and the like were repeated. The result was still a big collection, which included different syntactic and embodied means to direct a coparticipant. The scope of the collection was narrowed to include only one syntactic structure as means of getting another to perform an action, and *imperatives* were selected to be the criterion of inclusion.[6] Interested in multimodality and especially the intersection between phonetics and embodiment in naturally-occurring interaction, I limited the collection to only include sequences that provide grounds for a rigorous phonetic analysis: only those sequences were selected in which an imperative is repeated to pursue compliance. The repetition provides a natural control, making it possible to compare the phonetic structure of two imperative turns more readily. The result was 140 sequences in which (a) an imperative sentence type is used to get the co-interactant to do something; (b) the co-interactant resists against completing the sequence initiated by the imperative; and (c) the imperative is repeated, and this process is recursive. Figure 1.13 is the overall schematics of the sequences in the pursuit collection.
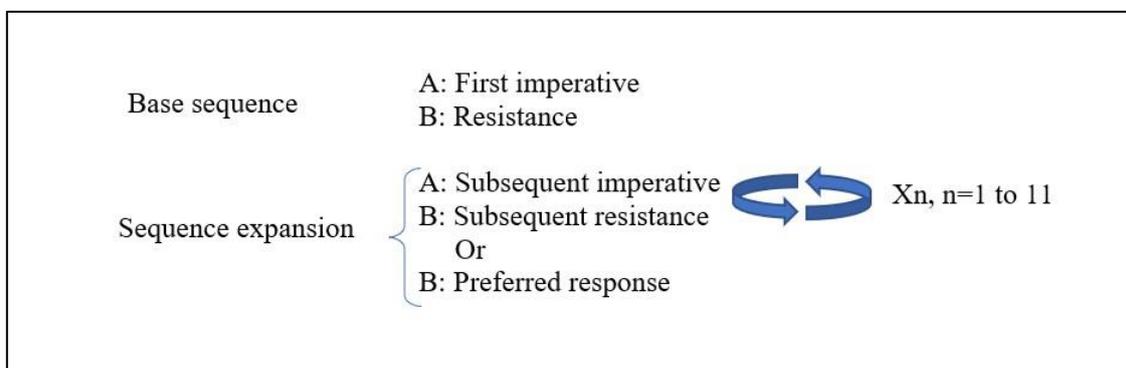
Figure 1.13. A simplified schematic of pursuit sequences

The number of subsequent imperatives in the collection varies from 1 to 11. The overall schematic is inevitably a simplified representation of the collection. Expectedly, sequences in the collection still differ from one another in details, for example, in some sequences, other means (e.g., embodied actions) are used to initiate the sequence and then facing resistance, a series of imperatives are issued. We will see this later where one participant stretches her hand, holding some fruit in it, towards the recipient, doing offering. The coparticipant rejects the embodied offer, and then multiple imperatives *bokhor* 'eat' are issued by the offering party. The minimum commonality in all the collected sequences is that an

---

[6] Even though prohibitives are considered, by some scholars, to be a form of imperative (negative imperative), to have a grammatically tight collection, prohibitives were also excluded.

imperative is repeated as means of pursuing compliance. In Chapter 2, we re-address the scope of the collection with examples. There, different kinds of repetition (full vs. modified) are also exemplified.

1.9.4. Transcription & Anonymization

After building the collection, the selected video clips were transcribed and anonymized. The audio excerpts of all imperative turns were then extracted and imported to Praat to analyze the phonetic properties of imperative turns (see the list of phonetic properties below in §1.10.2, phonetic analysis). Each video clip was also imported to ELAN to annotate relevant bodily conduct. Segmenting bodily movements into phases was done based on Kendon (1980) and Kita, van Gijn, and van der Hulst's (1998) conventions. Prosodic and bodily properties were then added to transcripts. For transcribing the embodiments, the conventions in Mondada (2019a) and for phonetic details, GAT2 (Selting et al., 2011) were used (Appendix A and B, respectively). Leipzig glossing rules (2015) were the guideline for glossing the transcripts (Appendix C). Finally, the extracts were analyzed through a combination of methods: conversation analysis, phonetic analysis, and multimodal analysis.

1.10.   Methods

1.10.1. Conversation Analysis

The analysis of the data in this thesis is done according to analytic methods in conversation analysis. Some main concepts of CA are already reviewed, so here I only mention that, following CA guidelines, in analyzing the interactional affordance of any turn, attention was focused on its sequential placement, turn design features, the accompanying actions or practices that may be produced with a turn, and co-participants' orientations. These are unpicked throughout the analysis of each extract in the forthcoming chapters (see §1.2 for more details on these concepts).

1.10.2. Phonetic Analysis

In analyzing the phonetic aspect of talk, I have followed the methodological guidelines observed in the phonetics-of-talk-in-interaction school (see Ogden, 2021, pp. 657-681; Local & Walker, 2005, for a full list of guidelines). The phonetic analysis of the data was done based on a combination of acoustic and impressionistic approach to data. In the forthcoming chapters, the phonetic details of the data can be found in various places: the flow of the text, Praat images, and in tables that summarize the design features of the imperative turns (called *design features summary* tables). Here is a list of phonetic features considered in analyzing an imperative turn:

- Intensity
- Average fundamental frequency or F_0 of the imperative turns
- Speakers' pitch range
- Speakers' overall F_0 level
- Pitch span
- Rate of articulation
- Articulation setting

The following briefly explains the measures that we will face in the analyses of the extracts and the procedure by which they were obtained.

Intensity

Intensity (measured in the scaled unit decibel or dB) is a physical measure related to the pressure which moves the air particles through the vocal tract (Stevens, 1998, p. 225). The greater the pressure, the greater the *amplitude* or the size of the movement of air particles to and from their rest position. "When the amplitude of the sound increases, listeners judge the sound to be *louder*" (see e.g., Stevens, 1998, p. 225). Therefore, loudness is a subjective perceptual correlate of intensity (see e.g., Laver, 1994; p. 501). The perception of loudness also depends on two other physical concepts: the fundamental frequency and spectral quality of sounds (Lehiste, 1970, p. 114).

In the visual presentations in this thesis, intensity line is shown with thin maroon colored dashed-dotted curves, and the top and bottom intensity measures of the imperative TCUs are shown on the right side (Figure 1.14). In this example, which we will analyze in Chapter 3, the imperative *beshin* 'sit down' is issued twice. The top and bottom intensity levels, taken both imperatives together, are 43 and 89 dB respectively. The average intensity of the first imperative is 58 dB and for the second, it is 80 dB. In the rest of the thesis, you will only see top and bottom intensity measures (not the average values). This is because in the visual presentations, frequency and intensity measures are overlaid in one single Praat image. To avoid inserting too many lines and values, which potentially creates confusion, only top and bottom intensity measure are shown in the images. In the design features summary tables, however, average values are provided.
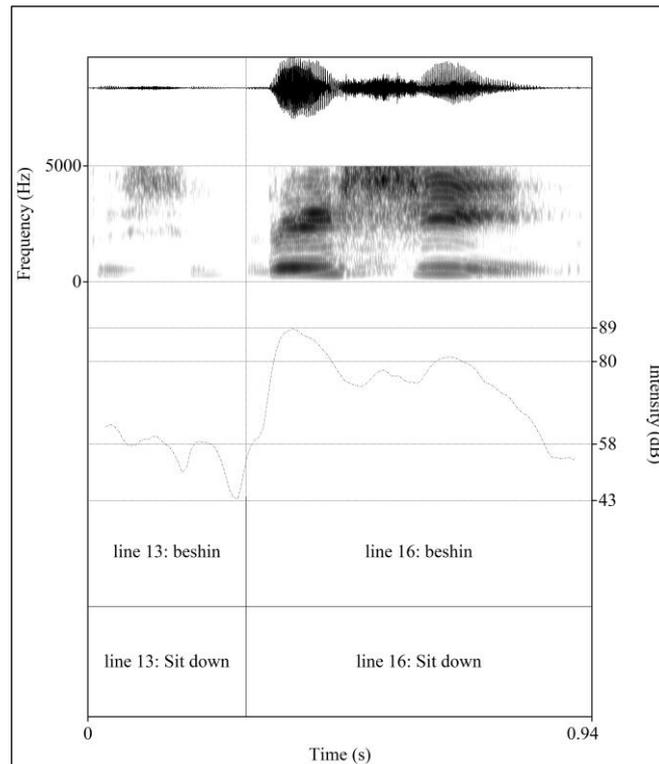
Figure 1.14. Intensity curve. The thin maroon dashed-dotted curves show intensity in decibel.

Average F_0 of imperatives

Fundamental frequency or F_0 (commonly measured in Hertz or Hz) is another acoustic measure, which refers to the "lowest frequency component in a complex sound wave" (Crystal, 2008, p. 203). In phonetic terms, the measure indicates "the frequency (or rate) of vibration of the vocal folds during the voicing of segments" (Laver, 1994, p. 450). Pitch is the perceptual correlate of F_0 (Stevens, 1998, p. 227). The higher the rate of vibrations of the vocal tracts during the production of a sound, the higher in pitch the sound is perceived. However, the relationship between the rate of vibration and the perception of pitch is not linear. Human ear is more sensitive to lower frequencies; that is, at higher frequencies a relatively bigger change in F_0 is needed for the listener to perceptually register a change in pitch. Therefore, pitch is presented through a logarithmic scale (rather than linear) in the graphs in this chapter (see Nolan, 2003).

To obtain the average F_0 of the imperative TCU, the audio of each imperative turn was separately imported to Praat. The pitch trace of the imperative TCU was obtained, and pitch correction was performed in order to keep the pitch trace as close to the analyst's perception as possible. The median of the F_0 values was used as the measure of average F_0. This is because the median is less sensitive to

outliers compared to average values (Gries, 2013, p. 119). In Praat visual presentations, thick black dotted curves show the pitch trace of the imperatives (Figure 1.15).
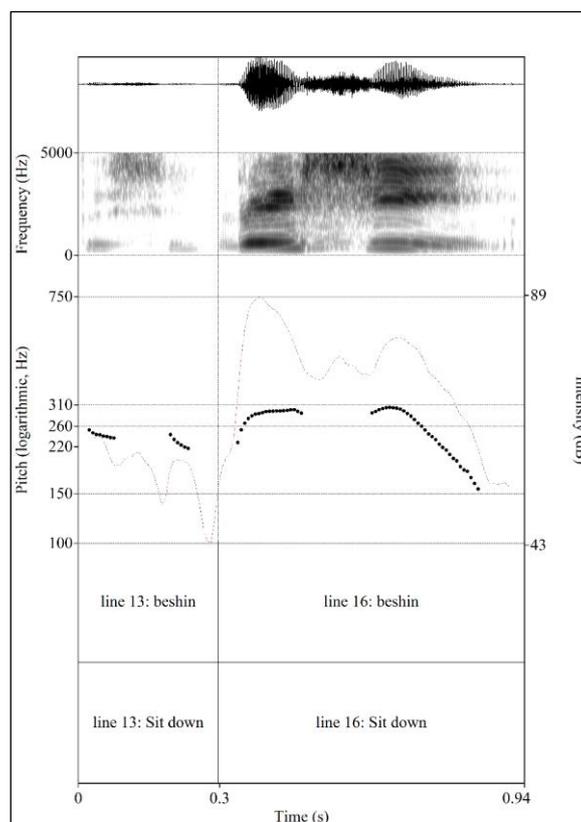
Figure 1.15. Pitch trace curve. The thick black dotted curves show pitch traces in Hertz.

In this example, the average F_0 for the first imperative is 234 Hz and for the second it is 287 Hz. The values on the left side are the guidelines to helps the reader spot the imperatives average F_0 relative to (a) the speaker's *overall F_0 level* and their *pitch range*. These two concepts are explained below.

Speaker's overall F_0 level

As well as the average F_0 of each imperative TCU, participants' overall F_0 was calculated based on one-minute sampling of their talk (see Walker, 2017). To do so, a one-minute audio file of each participant's clear talk (i.e., not in overlap) was constructed. To construct this file, bits of each participant's talk from different recordings at different times and from different interactional slots were extracted. The one-minute sampling for each participant contained stretches of talk which impressionistically sounded from very high to normal and to very low. The collected bits of talk (for each participant separately) were then concatenated in Praat. Thus, a sample audio file which was at least one minute long was created for each participant. All the recordings were made using the same audio recording equipment with the same setting; therefore, it was possible to concatenate stretches of talk from different recordings. Then pitch trace files of one-minute sample audios were created and corrected to ensure maximum consistency between the Praat pitch trace and the analyst's overall perception of the speaker's pitch movement. Finally, the median of the F_0 values was registered as the speaker's overall F_0 level. In all Praat images, the speaker's overall F_0 level is shown on the left side. In the *beshin* 'Sit down' example just above, the speaker's overall F_0 level is 220 Hz which is one of the values provided on the left.

Speaker's pitch range

The phrase *pitch range* can be used to refer to three different, but similar, concepts: organic pitch range, linguistic pitch range, and paralinguistic pitch range (Laver, 1994, p. 155). Organic pitch range or maximum phonational frequency range refers to the lowest and highest pitch limit of one's voice,

imposed by their laryngeal anatomy (Laver, 1994, p. 155). The range within which a speaker would ordinarily converse with others, which is usually smaller than their organic pitch range, is called linguistic pitch range or speaking fundamental frequency range, and finally, "if this range is adjusted for purposes of paralinguistic communication, it can be called the speaker's paralinguistic pitch range" (Laver, 1994, p. 155). In most studies (in linguistics literature), *pitch range* simply implies the linguistic pitch range; that is, the bottom and top pitch limits of a speaker's habitual speaking. An average speaker's conversational pitch range is reported to be around two octaves (24 semitones). But as explained above, the one-minute voice samples contain moments in which the participants are speaking noticeably higher or lower than their ordinary conversational pitch limits. Therefore, the maximum and minimum F_0 values reported for most participants in this chapter may be wider than 24 Semitones and delimit their *paralinguistic* pitch range. The choice to consider the speakers' paralinguistic pitch range (rather than the linguistic pitch range) was motivated by the observation that in this study, while repeating imperative TCUs to do persisting, participants sometimes speak a lot higher than their habitual conversational pitch limits. In vernacular terms, they can be heard to be shouting.

Speaker's paralinguistic pitch range was obtained using the corrected one-minute sampling Praat pitch trace files. In the *beshin* 'sit down' example, the top and bottom paralinguistic F_0 values are 750 and 100 Hz respectively, shown on the left side of the image (Figure 1.16)
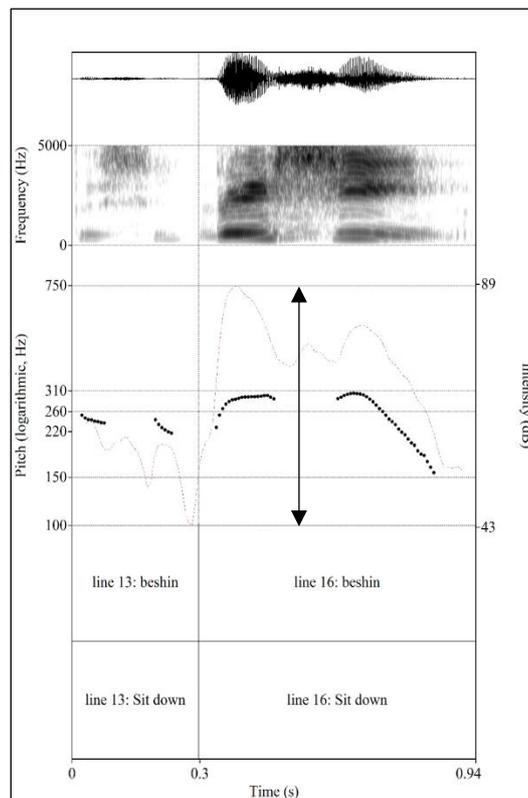
Figure 1.16. Pitch range. The paralinguistic pitch range of the speakers are shown by the bottom and top vertical lines in the pitch panel.

Note that in the literature *pitch range* is sometimes used to refer to the bottom and top F_0 values within a specific intonational phrase (rather than the speaker's overall pitch range), but I will reserve the phrase *pitch span* for that purpose.

Pitch span

"Within a speaker's linguistic [pitch] range, the momentary **pitch values** of the voice during a given utterance will vary between a local maximum and minimum. This can be called the speaker's **pitch-**

span" (Laver, 1994, p. 155, highlights original). 't Hart, Collier, and Cohen (1990, p.75) call this the "excursion size" of an utterance. Cruttenden (1997, p. 123) describes the same concept as the "width of pitch range over the whole intonation group" and refers to it as "key". This width can be measured by "computing the interval between some sort of average of high-pitch syllables and a similar average for low-pitch ones" (p. 124). Among the terms used in the literature, I use the phrase *pitch span* or simply span, reported in Semitones (ST). Pitch span for each imperative TCU is measured based on the corrected pitch trace of the imperative TCU (Figure 1.17).
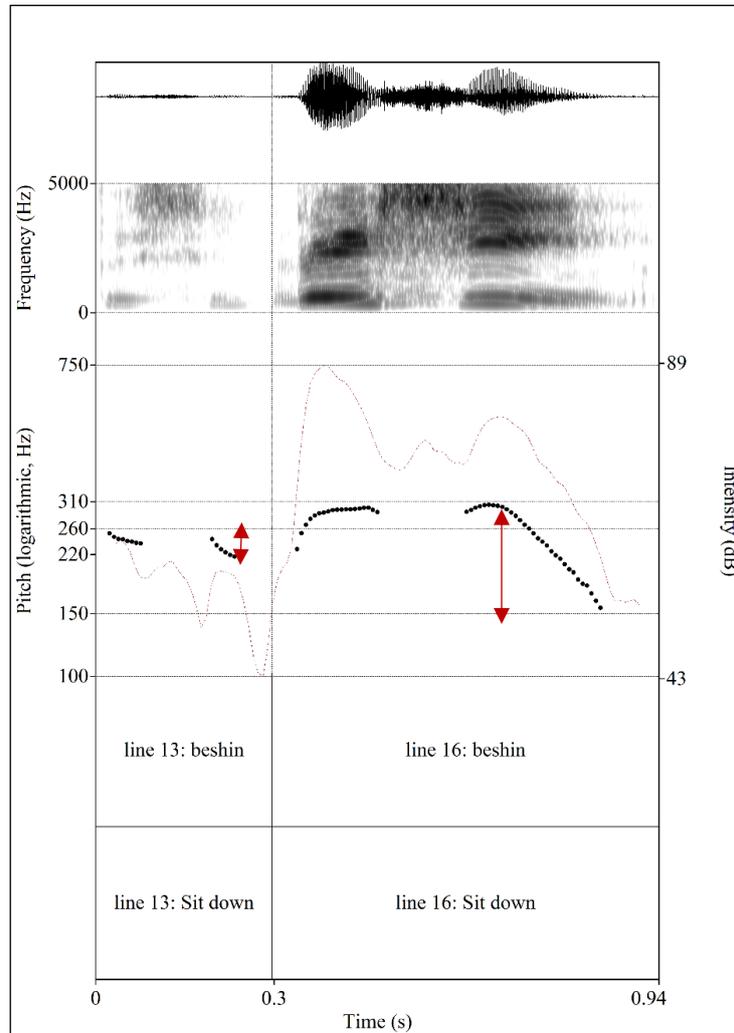
Figure 1.17. Pitch span. The pitch span of an imperative TCU is considered to be the distance between the minimum and maximum F_0 value in that imperative TCU.

The calculated pitch span for the first *beshin* is 3 ST and 11 ST for the second. The red arrows will not be shown in the rest of the thesis. The calculated pitch span values for each imperative will be provided in design features summary tables.

Rate of articulation

Articulation rate "refers to the tempo of performance of all audible, 'vocalized' speech within an individual utterance, whether that speech consists of the manifestation of linguistic units or of paralinguistic signals of hesitation" (Laver, 1994, p. 539). In simpler terms, it refers to how fast or slow

an utterance is vocalized, and it has a relationship with the duration of the utterance.[7] In this thesis, we compare the tempo of the production of two or more renditions of the same imperative. For two identical imperative TCUs, it is safe to say that if the production of one TCU takes shorter, it has a faster rate of articulation, and vice versa. In Praat visual presentations, this measure is shown on the bottom horizontal *Time* axis. Praat shows time in seconds, but following CA conventions, in the design feature summary tables, duration measures are reported in milliseconds. In the *beshin* 'Sit down' example, the articulation rate of the first imperative is faster (shorter in duration) than the second (300 ms vs. 640 ms) (Figure 1.17).

Finally, a guiding principle in analyzing the phonetic features of talk is noteworthy. Even though acoustic measures on intensity, duration, and various features related to F_0 are extracted and reported separately, human perception is not governed by analyses of a stretch of speech along various dimensions separately. Rather, it is a holistic impression of various segmental and suprasegmental features working together and affecting one another that makes us perceive an utterance the way we do. In analyzing the phonetic features of each extract, the design feature summary table under each extract lists all the phonetic properties outlined above for each imperative turn. The information is given so that the readers can develop their own independent analyses of data if they wish. But not all these measures may be relevant for a particular analytic claim that is made for a particular extract. Therefore, in the text of analyses, only relevant properties to that particular extract are commented on. For example, when analyzing an imperative turn as *upgraded* relative a prior imperative, not *all* the outlined measures above may be increased. Nevertheless, impressionistically, a sense of upgrading may be perceived. This is because, as mentioned earlier, a cluster or Gestalt of phonetic features feed our perception and this Gestalt can have different compositional details. In any case, the analytic claims are backed up with evidence from the data. In short, the phonetic analysis in this thesis is guided by both acoustic and impressionistic analyses together. The impressionistic analysis is verified by participants' own orientations.

1.10.3. Analysis of Bodily Movements

The analysis of bodily actions was guided by considering five aspects:

- the type of bodily actions
- the temporal organization of bodily actions
- the coordination between bodily actions and participants' talk
- the functional relationship between bodily actions and participants' talk
- the physical properties of bodily actions

Type of bodily actions

Overall types of bodily actions can be specified according to the observable end they are designed to achieve. Bodily actions referred to as *gesture* have some communicative purpose. Kendon (2004) defines gesture as "visible actions when it is used as an utterance or as a part of an utterance" (p. 7), and an utterance, according to him, is "any unit of activity that is treated by those co-present as a communicative 'move', 'turn' or 'contribution'" (p. 7). He excludes three other types of bodily actions from the gesture category in that they are not communicative (pp. 7-8): (a) bodily actions that give off feelings such as smiling (Ekman & Friesen, 1982), (b) self-adaptors such as touching and grooming oneself (Ekman & Friesen 1969, pp. 67-69; Streeck, 2020); and practical actions, sometimes called instrumental actions, such as eating, smoking (on practical actions, see Streeck, 2009b, p. 70).

To analyze bodily movements in this thesis, no bodily action – gesture or otherwise – was treated noncommunicative or irrelevant a priori. Initial analysis treated every observable bodily movement

---

[7] This relationship is not always straightforward especially if the utterance consists of pauses and gaps. Sometimes researchers count syllable per second as a measure for rate of articulation, but this complexity is not relevant in this thesis (see Laver, 1994, for a more comprehensive discussion).

potentially as relevant. Take, for example, the very first extract in this chapter where one of the participants (Azita) threw away the book belonging to the coparticipant (Roshana). Extract 1.1 is reproduced in part as Extract 1.9 below. (Embodied annotations and figure numbering are adjusted.)

Extract 1.9. RAY041a, 10.29 [excerpt of Extract 1.1]

```
->    17   ROS        [<<f,h>kheile khob  *bâbâ ⊤bezâ::?>*
                        very   OK     PRT IMP-let.2SG
                       OK OK my gosh wait?
           azi                               *..............*
           azi                                 ⊤moves hand twrds foot-->

      18   AZI   *.h ⊤e:::::::::    *⊤
                 *gaze twrds foot   *
                  -->⊤pulls out book⊤

      19         #⊤    (0.2)      ⊤(0.5) ⊤
           azi    ⊤throws book away⊤,,,,,,⊤
           fig   #1.18
```


#1.18

```
      20   ROS   leh    kard-i    un   ketâb-o:::::::::::
                 smashed do.PST-2SG that book-OBJ.mar
                 You smashed that boo:::::::::k
```

The throwing of the book was analyzed as a sanctioning move by Azita in response to Roshana's continued resistance in returning the phone. Now in the video clip, this throwing action looks to have initiated as a matter of managing the practicality of the situation – as a practical/instrumental action. As Azita approaches Roshana to take her phone back, she puts her foot on Roshana's book. Azita displays sudden noticing of this when she changes her face orientation rather abruptly: she turns her gaze back towards her foot and breaks eye contact with Roshana right in the middle of their 'heated' dispute (lines 17-18). Concurrently, she stretches her hand towards her foot (line 17) and pulls out the book from beneath her foot (line 18). But then, instead of putting the book away, she throws it away, making a rather loud noise (line 19; Figure 1.18). Therefore, Azita has moved beyond performing just an instrumental action. By throwing the book away, rather than putting it away, she changes the instrumental action of untangling the book to an interactional action of admonishing or negatively sanctioning Roshana. Azita's 'exaggerated' bodily action (throwing vs. putting) is oriented to by Roshana herself where she complains to Azita about 'smashing that book' (line 20). To reiterate then no bodily action – communicative or otherwise – was treated a priori as irrelevant. The guiding principle was tracing participants' own orientations towards what is done.

A note on transcribing bodily actions is due here. If we compare the bodily annotations in Extract 1.1 with the ones in its reproduced version just above, Extract 1.9, we will notice a clear difference. Extract

1.1 captured the overall movements of Azita's right and left hand; here the focus is on her left hand and gaze. So in each version, the annotated embodiments are selectively chosen (and with different degrees of details) to fit the argument, and this is a truism about most embodied transcriptions. Presenting each and every single detail in co-interactants' bodily movements would result in rather unintelligible transcripts. So bodily transcriptions throughout this thesis are inevitably selective.

Temporal organization of bodily actions

A bodily action is initiated when a body part begins its excursion away from the body and is ended when the body part goes back to the 'rest' position (Kendon, 1980; Kita et al., 1998). The movement can consist of a sequence of temporally distinct phases: preparation, stroke, and recovery or retraction (Kita et al., 1998, p. 3). A stroke is the obligatory phase of an embodied action: "a moment of accented movement", the peak, or the most meaningful part of a bodily action (Kendon, 1980, p. 212). A stroke may precede by a preparation phase in which the body part moves towards the stroke position and/or followed by a phase called recovery or retraction during which the body part moves towards the rest position. *Hold* is also another phase in which the bodily articulators involved in the stroke are held still. A hold phase may occur before or after a stroke (Kita et al., 1998). In the example above, moving hand towards the book is the preparation phase (shown by dots '…') of pulling the book out and immediately throwing it away. Then comes the retraction (shown by commas ',,,') during which Azita moves her hand back towards her body.

Where relevant, these phases are reflected in the transcripts following Mondada's (2019a) conventions. They are especially useful when the temporal coordination of bodily actions and prosodic events are discussed.

Temporal coordination between bodily actions and speech

The temporal coordination between speech and bodily actions is illustrated in Chapters 3 and 4. Here suffice it to say that such a relationship can be studied from three perspectives

- the coordination of the start and end marks of bodily actions with the performer's co-occurring speech (their co-extensiveness)
- the coordination of the bodily action with prosodic events such as prominent syllables, pitch accents and the like (see Schegloff, 1984; Loehr, 2012).
- the relationship of bodily actions with the coparticipant's multimodal productions (e.g., Mondada, 2016)

These aspects require extended exemplifications, which are provided in the following chapters.

Functional relationship between bodily actions and speech

Bodily actions can converge with the verbal or vocal productions that accompany them in terms of their interactional function. In Extract 1.9, for example, Azita produces an elongated interjection *e:::::::*, analyzed as displaying a negative stance towards her coparticipant's social behavior. Throwing the book looks to be in the service of a similar interactional function. Therefore, the two modalities converge in terms of their function. But, as shown later, this is not always the case. In fact, the functional divergence of embodied and verbal resources is the basis for one the main analytic claims in this thesis which we will cover in length in Chapter 4.

Physical properties of bodily actions

In analyzing bodily actions, attention was paid to physical properties of participants' movements: pace or speed of movement, size or length of excursions, direction of movements, force of actions (evidenced by observable cues), and in terms of haptic action, the magnitude of contact. In the example above (Extract 1.9), the force of the action is judged by the loud noise that the throwing action produces. The judgment on whether the noise is relevant or not is made based on the book owner's response to that noise (i.e., she raises her gaze off the phone, looks at the book, and issues a complaint about her book

being smashed). Observably also the movement of the book is fast (notice the blur in the figure which is indicative of the continues movement of the book), which is again another indication of the force behind the throwing action. Examples of other physical properties will be provided throughout the thesis.

1.10.4. Multimodal Analysis

This study adopts a multimodal approach to analyzing data. Although the details of the resources along various modalities were calculated or recorded separately, ultimately the totality of these observations together were considered to deliver the analytic claims. Such a holistic approach to data was guided by the principles of conducting multimodal research in CA. In §1.5, we reviewed the gradual rise and increasing popularity of multimodal investigations in CA. Here, I only list the principles, informing the analysis in this research, which are best summarized by Mondada (2014, 2016, & 2019b):

- Language is only one integrated resource within multiple other resources without any a priori priority (2014, p. 138).
- In formatting an action, a 'web of resources' from various modalities, shaping a 'complex multimodal Gestalt' can be operative (2014, p. 139).
- "[P]otentially every detail can be turned into a resource for social interaction, depending on the local ecology and the activity" (2016, p. 341).
- Some ecologies and types of activities might favor verbal resources along with gestures and body movements, whereas other ecologies and activities might favor distinctive and specific embodied resources over talk (2016, p. 341)
- Bodily actions "adapt to the particularities of [the] environment and cannot be understood without considering the embeddedness of the action within that environment" (2014, p. 142).
- Some resources may be conventionalized such as ritualized embodied actions (2016; p. 341).

These mean that a multimodal perspective to data questions the heavy weighting that is traditionally assigned to the verbal mode of communication in semiotics. This thesis will show that interactants use complex multimodal Gestalt in pursuing their conflicting CoAs and that the details of the Gestalt depend on the local ecology of the situation.

1.11. Organization of the Thesis

As previously stated, this study sets out to dissect moments of interaction when what an interactant pursues clashes with what the co-interactant pursues. Broadly speaking we will examine turn constructional features of a pursuit turn, the implications that these features have for the sequence structure, conflict management, and the pursuing party's agentive stance.

Chapter 2 has two overall aims. The first one to set a more detailed scene than the one presented in this chapter for the rest of the analytic chapters. It closely examines the organization of pursuit sequences and outlines some verbal practices by which a participant may pursue compliance. One of practices is repeating a prior imperative. It is this method of pursuit that this study mainly concentrates on. Two main types of repetition – full and modified – are presented. Then the second aim of the chapter is to introduce two types of pursuit: *persisting* and *gradual desisting*. It will be shown that these two types have different projective potentials in terms of the structure of the sequence. Persisting is sequence *expansive* while gradual desisting is likely to be sequence *terminative*.

In Chapter 3 and 4, a novel interactional phenomenon is introduced: *multimodal gradation*. Multimodal gradation, as defined in this thesis, involves a temporally coordinated modification of various design features from verbal, prosodic, and embodied modalities. In Chapter 3 *multimodal upgrading* in a pursuit turn is examined and in Chapter 4, *multimodal downgrading* will be analyzed. It will be shown that multimodal upgrading occurs when various constructional resources across a multitude of modalities are simultaneously 'intensified'. Multimodal downgrading, on the other hand, is achieved

when multimodal turn design features are simultaneously 'weakened'. These two chapters show that upgraded multimodal Gestalts can contribute to escalation of discord while downgraded Gestalts project withdrawal from the sequence and de-escalaation. From this perspective, multimodal upgrading is more associated with persistent pursuits while multimodal downgrading with gradual desisting.

In Chapter 5, we seek to find out the potentials of up- and downgraded Gestalts for the pursuing party's agentive stance. It will be argued that firstly by pursuing compliance, the pursuing participant displays a stronger agentive stance than abandoning the sequence. Furthermore, pursuits performed through upgraded Gestalts embody stronger agency for the pursuing participant than downgraded pursuits. From a more detailed perspective, pursuits with upgraded Gestalts can be divided into two broad groups: those pursuit turns that offer the recipient another opportunity for a responsive action and those that project that the pursuing party, herself, may 'force' the pursued interactional outcome to occur. It is claimed that the latter afford the pursuing interactant stronger agency than the former.

In Chapter 6 we will take the road to inspect another novel phenomenon: *mock aggression* (Afshari Saleh, 2020). Mock aggression refers to those embodied actions that appear to be aggressive (e.g., slapping, kicking, etc.) but they are not designed to physically harm the recipient and the recipient does not treat them as physically threatening. They are associated with 'play' sequences. In this chapter, we will widen our analytic lens and add other sequence types (than imperatives) to the picture. It will be argued that using *mock aggression* between intimate interactants provides an exit from a disaffiliative sequence into more affiliative ones. When it is used in an escalated pursuit sequence, they can contribute to de-escalation.

Finally, Chapter 7 provides a summary of the findings and discusses their implications in terms of sequence structure, escalation and de-escalation, preference organization, affiliation, and multimodality.

Chapter 2. Two Pursuit Types: Persisting and Gradual Desisting

2.1. Introduction

One integral ingredient of human sociality is our omnipresent capacity of getting others to perform actions. This capacity is materialized from very early age and in different forms as when a child raises both hands in the air for their mama to pick them up. It manifests itself in an array of different action types: we may make a request or issue a command, we may advise or instruct an action, or we may even beg or threaten another to do an action. What these action types have in common is that they all initiate a course of action, the completion of which relies on the co-interactant(s) to perform the action that is requested, commanded, advised, instructed and so forth. But what happens when the other party refuses to complete the initiated course of action? What moves does the initiator of the course of action make next? What practices can they employ to solicit a cooperative response after the refusal? What implications do these next moves and practices have for the trajectory of interaction? In this chapter we explore these questions by zooming in on episodes of interaction in which an imperative is issued to get another to perform an action, but the other refuses, one way or another, to comply.

Take the following example in which Faraneh stretches her hand towards her husband, Ebi, and issues the imperative *destam bus kon* 'kiss my hand' (line 2). The imperative and the arm stretch initiate a course of action which progresses once Ebi performs kissing. Ebi shows resistance against completing Faraneh's CoA by challenging Faraneh and soliciting an account (Bolden & Robinson, 2011). Faraneh orients to the incomplete status of her course of action and the ongoing relevance of kissing by holding her hand stretched even after Ebi's resistance (Raymond & Lerner, 2014). At the beginning of the extract, Ebi is outside the camera shot.

Extract 2.1. RAY040b, 09.01

```
    01              (0.3)+(1.2)
            far       -->+holds arm stretched-->>

->  02      FAR     #dastam   bus   kon.
                    hand-1SG kiss do.IMP
                    kiss my hand.
            fig     #2.1
```


#2.1

```
    03              (0.2)±(0.7)
            ebi         ±walks in front of the camera shot-->
```

48

```
->    04    FAR    bus kon.
                   kiss do.IMP
                   Kiss.

      05    EBI    cherâ.
                   why
                   Why.

      06           (1.7)+
            ebi       -->+stands in front of FAR-->>

->    07    FAR    bus kon
                   kiss do.IMP
                   kiss.

      08           (0.2)

      09    EBI    cherâ âkhe.
                   why   PRT
                   but why.
```

The imperative-speaker produces a first pair-part and names an action – kiss – for the imperative-recipient to perform. She facilitates a kissing action through her bodily conduct: she reduces distance between her hand and Ebi by stretching her arm towards him and exposes the back of her hand; hence she shows where on her hand the kissing action should be performed. As Ebi moves to the front of the camera shot, Faraneh re-issues the imperative (line 3-4).

A preferred second pair-part is compliance and doing the kissing since it progresses Faraneh's CoA. But Ebi defers a response; he initiates an insert sequence, asking Faraneh for an account (line 5). The account solicitation puts Ebi's resistance against completing Faraneh's CoA on display. His resistance becomes explicit when he reaches Faraneh, stands in front of her, but does not do the directed action (line 6). Therefore, he hinders the progressivity of the CoA projected by the imperative.

Facing the recipient's resistance, an imperative-speaker has a wide range of possible next actions, but two most general ones are to abandon the sequence or to pursue a response that furthers the hindered CoA – a preferred response. Here, Faraneh does the latter by repeating the imperative (line 7). Again, Ebi defers a response with the same practice: soliciting an account (line 9). The sequence develops as the participants fall in a loop of directing and account soliciting, which is not presented here. This extract is representative of the cases in most of this thesis, examples that illustrate one of the moves that an imperative-speaker may take after their initiated CoA is hindered by the recipient, namely, pursuing a preferred response by repeating the prior imperative.

This chapter uses imperative sequences similar to the one presented above to introduce two types of pursuit which I call *persisting* and *gradual desisting*. These two pursuit types differ in terms of the projected trajectory of the sequence. In the persisting type, the pursuit turns project sequence expansion while with gradual desisting, a sequence closure and an upcoming disengagement is likely. To establish solid grounds for this claim, after a brief review of the literature on pursuits and sequence structure in interaction (§2.2-2.3), I present possible sequential developments when a participant tells another to perform an action by means of an imperative. I start with the first position (the imperative turn itself, §2.4); then different ways by which an imperative may be responded are explored in §2.5. In §2.6, we see various interactional moves that the imperative-speaker may make after a recipient's resistance against compliance. These orientations include pursuing a preferred response. Six verbal practices by which participants can pursue a preferred response are outlined: addressing, incrementing, accounting, replacing, reformatting, and repeating. Then, we zoom into the last of these practices: pursuit by repetition (§2.7). It is this practice that this thesis centers around. The main analytic claim on the

distinction between persisting and gradual desisting comes towards the end of the chapter (§2.8). Finally, §2.9 concludes the chapter.

2.2. Pursuit

Broadly speaking, the term pursuit has been used in two main ways in CA. One is a more general usage based on one common meaning of the term; that is, seeking to achieve some objective. And another is a more specific usage, which, although hinges on the same semantics, it encompasses a more specific interactional move in which an interactant seeks a response: a response pursuit. Examples of the former usage of the term in CA include pursuing a topic (Button & Casey, 1985), intimacy (Jefferson, Sacks, & Schegloff, 1987), an uptake (Kim, 2007), alignment and affiliation (Kaneyasu, 2020), and so forth. And among works on response pursuits is Pomerantz (1984b), Stivers and Rossano (2010), and Bolden, Mandelbaum, and Wilkinson (2012). It is this second kind of pursuit that this thesis focuses on and the following overview orients towards. The essence of the concept is that after an initiating action, which makes a response conditionally relevant, when a response is absent or inadequate, the speaker of the initiating turn may seek, prompt, or encourage a(n) (adequate) response.

To the best of my knowledge, the first focused study on response pursuit was done by Jefferson (1981) in her investigation on "post-response pursuit of response". In this paper, Jefferson accounts for the production of tag questions such as *right?* to solicit responses after the coparticipant has produced a responsive action and relinquished the floor. Here is an example:

Extract 2.2. [BC:II:R:144] (adapted from Jefferson, 1981, p. 60)

```
        01      CAR     It was a good thing, it was a very great thing. It w- in
                        fact the greatest thing in the world has ever known, at
                        its time.
->      02      B.C     Mm hm?
->      03      CLR     Right?
        04      B.C     Mm hm
```

She notices that the coparticipant's response after which such response pursuits are produced is minimal, very short, often tokens such as continuers, and argues that these instances of pursuits communicate that "although the recipient may have taken it that prior speaker was not yet finished, he was indeed finished, and it is now recipient's turn to do some extended talk" (p. 60). Therefore, they are "a "Turn-Exit Device" when turn-transfer has not been adequately accomplished" (p. 60).

The term was then popularized by Pomerantz (1984b) through her work on assertion sequences. She observes that "[i]f a recipient does not give a coherent response, the speaker routinely sees the recipient's behavior as manifesting some problem and deals with it" (p. 152). One way for dealing with the problem, she argues, is to pursue a response through three main operations: "clarifying an understanding problem", "checking presumed common knowledge", "changing one's position" (pp. 152-163). Through these operations a speaker solicits agreement with their assertions. Here is an example of pursuing an agreeing response through clarifying an understanding problem:

Extract 2.3. [SBL:1:1:10] (Pomerantz, 1984b, pp. 154-155)

```
        01      B       And that went wrong.
        02              (1.0)
        03      A       Well, uh --
->      04      B       That surgery, I mean.
```

Pomerantz (1984b) explains that A's hesitation in providing a response (lines 2-3) can be implicative of some problems in understanding the reference of 'that' in B's turn in line 1. By clarifying the reference of 'that', B pursues an agreeing response.

Studying sequences more similar to the ones this thesis is based on, Davidson (1984) illustrates that subsequent offers, invitations, proposals, and requests can pursue 'acceptance' after the recipient has issued rejection or rejection-implicative responses, for example:

Extract 2.4. [NB:38, p. 92] (Davidson, 1984, p. 105)

```
      01    A      C'mon down he:re,=it's oka:y,
      02           (0.2)
->    03    A      I got lotta stuff,=I got be:er en stuff 'n
```

The gap in line 2 is rejection implicative and the subsequent version in line 3 is a potential account for why B may accept the invitation and thus encourages an acceptance.

After these early works, more CA investigations continue to examine pursuits. Rossano (2006) showed that gaze can be used to pursue a response (Rossano, 2006). Later, Stivers and Rossano (2010) counted four design features by which a pursuing party puts 'pressure' on the recipient to produces a response: interrogative morphosyntax, interrogative prosody, speaker gaze, and recipient-tilted epistemic asymmetry. Their claim is that first pair parts which they describe as 'noncanonical' (e.g., assessments, noticings, announcements) are more 'coercive" in mobilizing a response when they are designed with some or all of these features. They then extended their claim on the degrees of 'pressure' to actions that they call 'canonical' (e.g., offers, requests, invitations). They used this claim in favor of their argument on a scalar model of conditional relevance, which is questioned by some of the scholars in the field (see Couper-Kuhlen, 2010; Schegloff, 2010). This chapter and in general this thesis focus on design features of the pursuit turns, so in this regard this work and Stivers and Rossano's share an element of focus; however, as we will see later, I analyze the design features of pursuit turns in light of the trajectory of interaction rather than degrees of 'pressure' which may not be straightforwardly evidenced with the kind of argumentation that I am taking in this thesis.

Bolden et al. (2012) review some devices for pursuing a response: response prompts, preference reversals, turn extensions, and increments. Then they introduce a new device, namely self-initiated self-repair to indexical references. They note that while the first three devices "[*expose*] the lack of a response as a problem", increments or transition space repairs on an indexical component *obscure* the turn transition problem, and tacitly treat the speaker's own turn as problematic (p. 140). Here is an example taken from their work.

Extract 2.5. [6] Helpline (Bolden et al, 2012, p. 141)

```
      01    Clt    Oka:y. A:nd how did you hear about us:.
      02           (.)
->    03    Clt    Thee (.) helpline.
      04    Clr    Uh:::m (.) hhh I'm just tryin' to thi-
      05           <oh well I think it was throu:gh uh:m (.) uh: Christine
      06           Craggs-Hinton:'s book.
```

The repair in line 3 both clarifies the reference of *us* in line 1 and renews the relevance of a response from the recipient. The clarification masks the problem of the recipient's lack of response and frames the repair to be oriented to some inadequacy in the pursuing party's own question in line 1. Later in this chapter, I count some verbal practices by which a speaker pursues a preferred response after the recipient's resistance-implicative (non)response in imperative sequences. Some of them are in common with a part of Bolden et al.'s review, for example, incrementing. But then I go on to show how addressing, accounting, replacing, reformatting, and repeating function to pursue a preferred response. Focused on pursuits after an interactant's insufficient or inadequate response, Romaniuk's (2013) study on broadcast interviews illustrates four design features conveying that the interviewer is pursuing an adequate response rather than, for example, initiating a new question. These design features include explicit references to the initial question, verbatim repeats, indexical links, and implicit formulations.

This chapter also shows verbatim repeats as one verbal practice for pursuing, and indeed it is this practice that the thesis is based on. However, there is a difference between how verbatim repeats are conceptualized in this thesis and in Romaniuk's work. As I show later, I take verbatim repeats (full repeats) to copy all and only those verbal components that appear in a prior version. Romaniuk's examples of verbatim repeats includes those that are described as 'modified repeats' in this thesis. Here is an example in which she takes a 'verbatim repeat' is produced:

Extract 2.6. [#16 BBC Omnibus] (Clayman & Heritage, 2002, cited in Romaniuk, 2013, p. 151)

Initial Question: `Do you quite like him?`
First Pursuit: `<But do you like> him?`
Second Pursuit: `But do y'like him.`

The initial question includes a lexical component which the subsequent versions lack (i.e., 'quite') and also the conjunction 'But' is added to the subsequent versions. In this thesis, as we see later, this type of reproduction is taken as a modified repeat rather than a verbatim or full one. To argue that a repeat is doing pursuing rather than doing a new question, Romaniuk (2013) relies on contrastive markers such as 'but' that are produced in the pursuit turn. These markers, she argues, function to contrast an interviewee's response. It is not clear from her analysis how a *full* repeat, without these additional contrastive markers, may convey that the repeat is a pursuit rather than an initial attempt. Later, I argue that in a full repeat after a dispreferred response, it is its sequential position and other multimodal design features that convey that the turn is a pursuit. We will see that a full repeat after a dispreferred response treats the prior version as unproblematic, one that does not need any adjusting and tacitly passes accountability for the halted sequence to the recipient. See also Rauniomaa (2008) for a study on repetitions as a 'recovery' practice by which the speaker displays their stance, for example, towards the recipient's uptake as problematic.

The literature is growing with more investigations on various verbal practices for response pursuit in various languages and interactional contexts. Betz, Taleghani-nikazm, Drake, & Golato (2013) study third-position repeats in German sequences where a party requests information. They discuss that the repeat expands the sequence and pursues a more elaborate response from the coparticipant. Stevanovic (2013) focuses on proposal sequences in Finnish workplace interactions and shows that asserting one's thought after an initial proposal is a method to pursue an adequate response. In her later work on violin instructions, she studies the Finnish particle *-pA* in imperative and hortative turns to pursue compliance (Stevanovic, 2017). Again, in Finnish, Etelämäki and Couper-Kuhlen (2017) show that when an imperatively formatted directive faces the recipient's resistance, the speaker pursues a compliant response through a second-person declarative structure. They discuss this syntactic change in terms of its implication for agency.

Using bodily resources to pursue a response has also attracted some scholars' attention. As well as Rossano (2006) and Stivers and Rossano (2010), mentioned above, Peräkylä and Ruusuvuori (2012) also study facial pursuits in the context of tellings. When a due display of stance is delayed by the recipient of a telling or it is not sufficient, the researchers show, the telling-speaker may use facial expressions to reinforce, explicate, or sometimes modify their own displayed emotional stance and that can elicit an affiliating response. Children as young as 2 to 3 years old have also been shown to use their bodies to pursue a response to their assessment when there has been no uptake from the caregiver (Keel, 2015).

More recently, Hoey et al. (2020) have shown people use various turn design resources – verbal, prosodic, and embodied – to pursue a sufficient response. The verbal resource they focus on is expletive insertions in sequence-initiating actions, by which a speaker upgrades the turn-constructional method of the pursuit. Expletive insertion embodies an escalated attempt in pursuing a response. This chapter and the forthcoming chapters focus on pursuits, this ubiquitous sequence structure which is shown to exist in a variety of languages (Kendrick et al., 2020). I will argue that the various resources that

different scholars have shown to pursue a response may be in fact parts of a multimodal Gestalt. It will be shown that upgrading is only one possible turn design for a pursuing turn as is escalation only one possible trajectory. Another alternative, we shall see, is downgrading and withdrawal from the sequence.

2.3. Overall Structure of Pursuit Sequences

The main claim of this chapter is that the design features of a pursuit turn has implications for the structure of the sequence projected by the pursuit turn. In this section, we will have an overall look at the structure of a pursuit sequence without considering turn constructional features. As reviewed in Chapter 1 (§1.2), Sequentiality in interaction, that is the orderly arrangement of actions, turns, and turn constructional components, is one the main underpinnings of CA, and it is studied under the rubric *sequential organization* (Schegloff, 2007). One type of sequential organization is sequence organization: "the organization of courses of action enacted through turns-at-talk" (Schegloff, 2007, p. 2; Kendrick et al, 2020). Again, as mentioned in Chapter 1 (§1.2), in an adjacency pair structure, the first and second pair parts together make a base sequence. This section shows that a base sequence can be expanded in different ways. If the expansion occurs before the base sequence, it is a pre-expansion; if it comes between the first and the potential second pair-part, it is an insert expansion, and post-expansions are produced after the base SPP slot (see Schegloff, 2007, for a comprehensive discussion). The main habitat of pursuit turns, of the type that we will analyze in this thesis, is in post-expansions. The following extract includes a pre-expansion (line 1-2), but our main focus is on lines 6-11 where two rounds of post-expansions occur (this sequence does not have an insert expansion). Amir is a university student and has an exam. His wife, Farnaz, is directing him to get up and study.

Extract 2.7. RAY011a, 10.00

pre	01	FAR	amir jân? name dear *dear Amir?*
	02	AMI	hâ. INTJ *Mhm.*
base	03	FAR	pâsho dars-e-t-o be-khun. get.up.2SG lesson-of-you-OBJ.PAR IMP-study *get up {and} study your lessons.*
	04	AMI	na no *no.*
	05		(0.6)
Post-1	06	FAR	ye bâr dige mi-g-am. one time another PRG-say-1SG *I'm saying {it} another time.*
	07	FAR	pâsho dars-e-t-o be-khun. get.up.2SG lesson-of-you-OBJ.PAR IMP-study *get up {and} study your lessons.*
	08	AMI	ne-mkhâ-m. NEG-want-1SG *I don't want to.*

```
09    FAR   baar-e   sevom elaam   mi-kon-am  ke
            time-of  third announce PRO-do-1SG that
            I'm announcing {it} for the third time that

Post-2  10            paash-i   dars-e-t-o be-kho[on-i
                      get.up-2SG lesson-of-2SG-OBJ.PAR SBJV-study-2SG
                      you get up {and} study your lessons.⁸

        11    AMI                            [<<p.all>ne-mikh-am>
                                                NEG-want-1SG
                                                I don't want to.
```

With Amir's recipiency secured in the pre-expansion sequence (lines 1-2), the base imperative turn – the base FPP – is issued (line 3). The FPP projects the relevance of an SPP, and in the next slot, Amir produces a noncompliant response. With Amir's response, the base sequence is complete. At this point Farnaz can treat the sequence as closed or she can expand it. She does the latter and pursues a compliant response (line 6-7). In this extract, the status of Farnaz's pursuit as a post-expansion FPP, rather than a new base FPP, is especially clear in the adverbial phrase *ye bâr dige* 'another time', which shows the orientation of the imperative-speaker herself that she is re-doing an action that has been already performed. Rather than being a new base FPP with merely a similar topic or agenda to the previous base sequence, Farnaz's imperative turn (lines 6-7) is structurally built on the based sequence and shows that even though the base sequence is complete, the overall sequence is not yet closed for her. Amir then produces another noncompliant response: a post-expansion SPP (line 8). And the process is recursive. At this point, too, Farnaz could have oriented to the sequence as closed, but she further expands the sequence (lines 9-10). Again, the status of her turn as another post-expansion (a post-post-expansion) can be seen in her counting practice ('I'm announcing it for the *third* time', line 9). This makes it public that the imperative-speaker herself is tracing her CoA back to the base sequence such that her action in lines 9-10 is the third in a series rather than a new first. And Amir completes the second post-expansion sequence with another noncompliant SPP.

This extract was presented to show a brief structural overview of the sequences that appear in this thesis. While not all the sequences have pre- or insert expansions, all are expanded after the base SPP slot.⁹ In the following sections, we take a more detailed look at three crucial positions in an imperative sequence: the first position (the imperative turn itself), the second position where a responsive action is expected to occur, and the third or subsequent position which is the slot after the responding position where a pursuit may or may not happen. We will see various 'moves' that participants can make in the second and third/subsequent positions. This will set the scene for the main claim of this chapter: pursuit turns are of two main types. Some project further sequence expansions (persisting) and some project a contingent sequence closure (gradual desisting).

2.4. First position: Imperatives

As reviewed in Chapter 1, an imperative structure can be in the service of different social actions and interactional outcomes. This thesis revolves around those interactional episodes when an imperative is used to get a coparticipant to perform an action. To reiterate from Chapter 1, in this thesis, for the ease of reference, the term imperative is a shorthand to refer to an imperative when it is in the service of this broad aim unless stated otherwise. With this operationalization, an imperative-recipient is under a normative expectation to, in general terms, grant or reject the performance of the action that imperative

---

⁸ The third pursuit is not formatted as an imperative.

⁹ For a common type of insert expansion in the sequences in this collection, refer to the first extract in this chapter (Extract 2.1), where after the imperative 'Kiss my hand', the imperative-recipient initiated an insert sequence: 'Why.' Such insert expansions are in the service of challenging the action projected by the imperative.

has made relevant. Whatever the recipient's next action, it is most likely interpreted for its import for one of these two alternatives. Consider the following case where Masumeh, who is setting the dinner table, addresses her adult daughter, Sufia, who is lying on the couch, with two imperatives: *pâsho* 'get up' and *komak kon* 'help' (line 1). A preferred response, from Sufia, would have been to physically get up and contribute to the activity of table-setting. However, in the response slot, after a noticeable silence during which Sufia continues to lie on the couch (line 2), she produces a declarative, informing the imperative-speaker that she is tired (line 3).

Extract 2.8. RAY010b, 13.19 [An overlapping commentary on a TV program is not transcribed.]

```
01    MAS    pâsho      komak kon    dokhtare tanbal
             Get.up.IMP help  do.IMP girl     lazy
             Get up {and} help lazy girl
      suf    >>lies on the couch>>

02           (0.7)

03    SUF    m:::an   khaste-am=hehehehe[hehe
             1SG      tired-be.1SG
             I'm tired

04    MAS                              [kheili porru-i-â
                                       very   cheeky-be.2SG-PRT
                                       You're too cheeky
```

The imperative verbs project a CoA in which the imperative-recipient (Sufia) performs the named action. The imperative-recipient herself orients to the imperative turn as a resource used by the imperative-speaker (Masumeh) to get her to perform the named actions. This is evident in Sufia's next action where she accounts for not performing the imperative actions (line 3). Since performing the actions take physical energy, the informing 'I'm tired' (hence, lacking physical energy) functions as an inability account (Heritage, 1984). The account functions as an implicit rejection. This informing is interpreted by the Masumeh herself for its import for either granting or rejecting, evidenced by her next action where she disaffiliates with Sufia's account (line 4). Generally, a speaker's negative assessment of her physical condition (e.g., tired, ill, etc.) can occasion an affiliative response, for example, a display of sympathy; however, Sufia receives a disaffiliative response (line 4). This is because by virtue of its sequential placement after the imperative, Sufia's negative assessment of her physical status in line 3 is interpreted by Masumeh as rejection-implicative and receives disaffiliation through a sanctioning response rather than affiliation by, say, sympathizing.

2.5. Second Position: Responses to Imperatives

In the second position, the recipient can respond to an imperative FPP in a variety of ways of which two alternatives are granting (complying) and rejecting (noncomplying). Granting is the preferred and rejection is the dispreferred next action. As reviewed in Chapter 1 (§1.5), these two alternatives differ in their potential affiliative valences and the interactional trajectories that they afford. The preferred alternative promotes the interactional outcome of the imperative and furthers the CoA that the imperative has initiated whereas the dispreferred alternative suppresses the imperative's interactional outcome and impedes the progression of the launched CoA (Clayman, 2002; Heritage, 1984; Schegloff, 2007, p. 59; Pillet-Shore, 2017; Pomerantz & Heritage, 2013). The former has cooperation- and the latter has resistance implications. The following two extracts show these alternatives. The participants in Extract 2.9 (not from the pursuit collection), are having lunch. Mona and Haleh (mother-late teenage daughter) are sitting next to each other such that Mona's space is rather cramped; she tells Haleh to move further away (line 1).

Extract 2.9. RAY048a, 09.21

```
01    MON    hâleh ye kam  unvartar        *boro      mâmân*
             name  one bit that.side.more  IMP-go.2SG Mom
             Haleh go to that side a bit sweetie
```

```
        hal                              *..............*

->   02        *(1.5)            *
        hal   *moves further away*
```

Haleh starts to move her body away from Mona at the first instance where the rest of Mona's TCU is projectable even before the imperative verb is issued (note the preparation phase of Haleh's movement, line 1). She produces a preferred response, the response type that advances Mona's CoA. In the following extract the opposite happens: the imperative-speaker's CoA stops progressing when the recipient refuses to perform the named actions (*boro* 'go' and *biâr* 'bring'). Dara and Safura (couples) are at Safura's mother's.

Extract 2.10. RAY013, 10.18

```
    01   DAR   boro        châyi bi-âr
               IMP-go-2SG tea    IMP-bring.2SG
               Go bring some tea

->  02   SAF   ne-mikh-âm
               NEG-want-1SG
               I don't want to
```

Here the response to the imperatives is dispreferred. The CoA that Dara's imperatives has initiated involves Safura's going to the kitchen and bringing tea for him and potentially him drinking it. This CoA is impeded by Safura's noncomplying response. Even though the recipient completes the sequence by producing a dispreferred second pair part, the sought interactional outcome is suppressed.

Compliance and noncompliance are not categorical but should rather be understood as a continuum. If we take these two response types as two extremes on the continuum of possible responses to an imperative, somewhere in the middle grounds lie other responses which have both preferred and dispreferred features. Take the following example in which the recipient performs a responsive action that facilitates for the imperative-speaker to do the directed action herself. Haleh (late teenage) and her adult brother, Amir (adult), are sitting on opposite sides of the lunch food cloth (The two participants in front in Figure 2.2. Haleh, left; Amir, Right). A bottle of water is next to Amir (in Figure 2.2, the bottle is covered by Amir's right arm. Its whereabouts is indicated by the red arrow). Haleh addresses Amir with the imperative 'pour water' (line 1). This is not the first time that Haleh is telling Amir to pour her water. The imperative is prosodically loud and high, orienting to Amir's failure to do the action earlier. In response, Amir puts the bottle of water in front of Haleh (lines 3-4).

Extract 2.11. RAY049a-b, 24.27

```
    01    HAL    <<f, h>#amir âb    be-riz>
                       Name water IMP-pour
                       Amir pour water
          fig          #2.2
```



```
                                                    #2.2

    02           (0.4)

->  03    AMI    +biâ
```

```
                       come
                       here
                       +pick up bottle from its tip-->

04                     (0.5)+(0.4)#
          ami            -->+moves bottle toward HAL-->>
          fig                       #2.3
```


#2.3

```
05     HAL     be-riz     amir
               IMP-pour Name
               pour Amir
```

A complying response – the preferred alternative to the imperative FPP – would have involved the recipient's picking the bottle up and pouring water into Haleh's glass. Amir picks the bottle, but the way he does so projects that he is just moving the bottle towards Haleh. Rather than picking up the bottle from its body, Amir picks it up from the tip, blocking the entrance of the bottle through which water should move out (lines 3-4; Figure 2.3). This suggests that he is going to only pass the bottle. Only halfway through Amir's moving the bottle, Haleh repeats the imperative (line 5) and hence displays her understanding of Amir's embodied action as one that is not going to complete the CoA. The repeated imperative is a pursuit for a preferred response and creates another interactional slot for Amir to complete the CoA. The prosodic design of the pursuit is downgraded; it is not loud or high anymore which displays Haleh's orientation to Amir's action as one that at least heads towards making water accessible. The object of the verb (i.e., water) which was produced in the prior version is dispensed with in the subsequent. By repeating the imperative without the syntactic object, Haleh topicalizes the action that should be done (i.e., pouring vs. passing) and narrows the focus of attention to the expected action.

Amir's response creates an opportunity for Haleh to further her own initiated CoA. Although he himself does not make the CoA reach a completion point, by making the bottle accessible to Haleh, he facilitates the progression of Haleh's CoA. These sorts of responses are not entirely complying, nor are they totally noncomplying. So on the continuum of responses to the imperatives, they sit somewhere in between with both preferred and dispreferred features. The preferred features are manifested in the moves that potentially allow the progression of the CoA and the dispreferred status is grounded in the fact that the response itself does not make the CoA reach a completion point. In this example, the participant's emic orientation to the mixed nature of these responses is evident in pursuing a response but with prosodically downgraded features. These types of responses are *less than preferred*. They have implications of both cooperation and resistance simultaneously and, as I will show in the forthcoming chapters, the imperative-speaker may orient to one of these implications more than the other.

2.6.   Third Position: Orientations to Resistance

Having seen a range of different responses to imperatives, we now turn to the third position in an imperative sequence where the imperative-speaker may show their orientation to the recipient's response. The continuum of responses to an imperative includes preferred, less than preferred, and dispreferred responses. In this thesis, cooperation or a cooperative response is conceptualized to occur when the recipient produces a preferred response – one that *furthers* the CoA projected by the imperative turn. Resistance, on the other hand, occurs when a recipient produces a dispreferred or a less than preferred response. These responses *hinder* or *slow* the progressivity of the initiated CoA. This

thesis mainly focuses on those imperative sequences with resistance-implicative responses. Therefore, in this section, I illustrate three main ways by which an imperative-speaker deals, in the third position, with the recipient's resistance: abandoning the sequence, closing the sequence, or pursuing a preferred response.

2.6.1. Abandoning the sequence

One way to deal with a recipient's resistance is to abandon the disaffiliative sequence: the imperative-speaker does not pursue a preferred response or does not even close the sequence; they do not expand the sequence with the imperative-recipient.[10] They tacitly let the CoA be incomplete (at least temporarily) and withdraw from the sequence straightforwardly. I use the terms 'withdraw' and 'disengage' interchangeably in this thesis (see Chapter 4 for a review of disengagement and withdrawal in CA). I save the term 'abandon' for a particular type of withdrawal: straightforward withdrawal from the sequence without any sequence expansion. In the following extract, for example, Minuvash (late teenage), her mother, Melisa, and a third participant are just about to start a second round of a card game. Concurrent with Minuvash's *khob* (line 1), marking a transition toward the next phase of the activity (Afshari Saleh, Taleghani-Nikazm, Monfaredi & Rezaee, 2022), Melisa taps on her mobile phone upon which a speech starts playing (lines 1-2). Responsive to her mother's playing out a speech in their play time, Minuvash produces a 'click', a vocal display of negative stance, and then moves the phone out of her reach (lines 3-6) as she accounts for her embodied conduct (line 3, *mikhâym bâzi konim* 'we want to play'). The turns in focus are when Melisa issues an imperative to regain access to her own phone (line 7, *bedesh mandesh* 'give it to me that'), Minuvash produces an overt rejection (line 8, *nakheir* 'no'), and Melisa abandons the sequence (line 9).

Extract 2.12. RAY050a, 04.45

```
01   MIN    +khob+
             PRT
             OK
     mel    +taps on mobile+

02          (0.7)  ((a speech starts playing on the mobile device))

03   MIN    ((click))±*mikhâym          bâ ±zi kon-im
                      PRG-want-1PL       play  do-1PL
            ((click)) we want to play
                      ±....................±puts mobile away-->
     mel              *gaze at mobile-->
04          (1.2)

05   MEL    .he::

06          (1.3)±
     min       -->±

07   MEL    +be-de-sh#          [man-desh]+*
             IMP-give.2SG-3SG 1SG-3SG
             give it to me that

08   MIN    +        #          [na-kheir]+*
                                 No-no
                                 No
     mel    +stretches hand------------+
     mel                               -->*
     fig            #2.4
```

---

[10] She may expand the sequence addressing another participant though.

#2.4

```
09        +  (1.0) #+*±(10.0)
->   mel  +,,,,,,,,+
          *gaze down*
     min             ±deals cards-->>
     fig           #2.5
```


#2.5

The imperative speaker's straightforward withdrawal from the imperative sequence is evident in her embodied conduct. Concurrent to the imperative (line 7), she stretches her hand, palm open, towards Minuvash, who has hold of her (Melisa's) mobile device, with her gaze directed at it (Figure 2.4). As well as the imperative, this embodied conduct – hand stretch and gaze direction – makes relevant the passing of the mobile (see Kendrick, 2020, p. 107 on the orientation and shape of manual actions specifying a relevant next action). Both the sequential and morphological design of the rejection (line 8) makes it a dispreferred response delivered in a preferred format (Kendrick & Torriera, 2015): it comes in overlap and, morphologically, it is a compound word *nakheir,* composed of two free morphemes *na* and *kheir*, both of which can be used as stand-alone rejecting resources. This double negation colors the rejection a 'strong' disaffiliative shade. Facing the resistance, Melisa, immediately, retracts her hand, rests her arm drop on her knee, and shifts her gaze down (Figure 2.5). She does not show any negative stance towards Minuvash's resistance or ask for her mobile anymore; that is, she does not negatively sanction the dispreferred response or pursue a preferred response; she does not even close the sequence. She simply quits the sequence without expanding it whether to pursue compliance or close the sequence: she abandons her CoA. Soon after this, Minuvash deals the cards, and the participants resume the game.

2.6.2. Closing the sequence

Another way by which an imperative-speaker may deal with the recipient's resistance is closing the sequence: the imperative speaker expands the sequence minimally and produces, for example, sequence closing thirds such as acceptance tokens or assessment turns that shows her stance towards the recipient's response (Schegloff, 2007, p. 118). The following two extracts show these two different ways of sequence closure. In Extract 2.13, Kati (adult daughter) addresses Faraneh (mother) with an imperative, requesting information on the whereabouts of a pair of knitting needles for Faraneh to knit something (lines 1-2). After Faraneh refuses to complete Kati's CoA through an account (line 4), Kati produces a sequence closing third (SCT) *bâshe* 'OK' (line 6).

```
     01   KAT     khob be-g-u         *mil    bâftani-â   kojâ-s=
                  PRT  IMP-say.2SG     needle  knitting-PL where-be.3SG
                  PRT tell me where the knitting needles are
                                          *gaze on FAR-->

     02   KAT     =man be-r-am      bi-ar-am
                   1SG SUBJ-go-1SG SBJV-bring-1SG
                  {so that} I go get {them}

     03           (0.3)

     04   FAR     •alân hese-sh    na-dâr-am•
                  now   mood-3SG   NEG-have-1SG
                  I'm not in mood now
                  •gaze at KAT------------•

     05           (0.3)*
          kat       -->*

->   06   KAT     *bâsh-e
                   Be-3SG
                   OK
                  *gaze down-->>
```

Through the SCT *bâshe* 'OK', Kati explicitly closes the imperative sequence and treat Faraneh's account for not offering information on the whereabouts of the needles an adequate response (see Kendrick, et al, 2020; Schegloff, 2007, pp. 120-123). Here, *bâshe* is a minimal sequence expansion that closes the sequence and projects the imperative-speaker's withdrawal. Her withdrawal from the sequence is also embodied in her gaze aversion away from the imperative-recipient (Rossano, 2012). With her *bâshe* 'OK' and bodily action, Kati takes an affiliative stance towards Faraneh's rejection in that she endorses Faraneh's choice about not providing the requested information.

Sequence closing thirds are not always affiliative. In the following extract, the imperative-speaker disaffiliates with the recipient's rejection through a negative assessment that closes the sequence (Schegloff, 2007, pp. 123-127). Faraneh, who has been mixing the content of a bowl, suspends her mixing activity and initiates an imperative sequence addressing her husband, Ebi, as soon as he enters the room (lines 1-2, 'Take my car key'). She just starts her second TCU (line 3; presumably an instruction for Ebi on what to do with the key afterwards) when Ebi produces a rejection in overlap, thwarting Faraneh's initiated CoA through an inability account (line 4, 'Sorry I'm late'). Then, concurrent with Ebi's second piece of inability account (line 6, 'I must go to a meeting'), Faraneh produces our line of interest: a disaffiliative SCT (line 7, *eh nekbat* 'Eh lousy').

```
     01           +±(0.3)
          ebi     >>+walks to the living room in work clothes-->
          far       ±stops mixing-->

     02   FAR     +chan tâ      he:u\          suiche mâshine mano vardâr,+=
                  some  number hesitation.PAR key    car     1SG  IMP.pick
                  some um\ take my car key,
          ebi     +stops walking----------------------------------------+

     03   FAR     +[e\
                   hesitation.PAR
                   um

     04   EBI     +[sharmande man dir-a-m      shode
```

```
                 ashamed   1SG late-of-1SG become.3SG
                  sorry I'm late
              +resumes walking towards the door-->

     05          (0.5)

     06   EBI    [bâyad be-r-am      jalase
                  must   SBJV-go-1SG meeting
                  I must go to a meeting.

->   07   FAR    [eh    nekbat
                  INTJ lousy
                  eh lousy

     08             +*±(1.1)
          ebi   -->+walks out of the living room-->>
          far   -->*gaze down-->>
          far   -->±resumes mixing-->>
```

After the recipient's dispreferred response (lines 4-6), the imperative-speaker closes the sequence by verbal and embodied means. Firstly, she produces a composite minimal expansion SCT (Schegloff, 2007, pp. 127-142). The SCT *eh nekbat* 'eh lousy' is made up of two components: the Persian interjection *eh*, sometimes used responsively to unpleasant stimuli and the curse word *nekbat* 'lousy', addressed at Ebi (line 7). It is this third-positioned negative assessment on Ebi that Faraneh employs to both close the sequence and disaffiliate with Ebi's uncooperative response. Even though Faraneh takes a negative stance towards Ebi's rejection, she explicitly treats the sequence as closed. This is evidenced by her embodied conduct which displays her disengagement from the sequence: she averts her gaze away from Ebi and resumes the activity of mixing (line 8). By re-engaging in the activity that she had previously suspended, Faraneh shows her orientation towards the interjecting imperative sequence as closed (see Raymond & Lerner, 2014, on the suspension of activities by interjecting actions).

2.6.3. Pursuing a Preferred Response

The two previous sections showed that following the recipient's resistance, the imperative-speaker may simply abandon the sequence or close it by means of affiliative or disaffiliative sequence closing thirds. Both ways indicate the speaker's disengagement from the sequence. A third way in which the sequence may unfold in the third position is a non-minimal expansion of the sequence by the imperative-speaker to pursue a preferred response (Schegloff, 2007, pp. 148-168 on non-minimal expansions). By pursuing a preferred response, the imperative-speaker tacitly shows their orientation towards the sequence as not yet closed. In the sequences which we investigate in this thesis, since the recipient has already displayed resistance one way or another, the response pursuit limits the recipient's response options to only one alternative – a preferred response. Therefore, a response pursuit post-resistance is in fact a pursuit of a *preferred* response, or in other words, the response pursuit is biased towards compliance.

There are a variety of verbal practices by which a response pursuit can be achieved*: incrementing, addressing, accounting, replacing, reformatting,* and *repeating*. These practices vary in terms of the operation they perform on the base first pair part. Incrementing operates at the level of TCU: the pursuit turn does not include a new TCU; rather, it *extends* the same imperative TCU. Some operate beyond TCU level: replacing, reformatting, and repeating. In this collection, they *initiate* a new TCU.[11] Addressing and accounting can belong to either of these two groups depending on their prosodic and

---

[11] Replacing, reformatting, and repeating are practices that are sometimes used in the organization of repair in CA (Schegloff, 2007). In that organization, they do not necessarily result in a new TCU. So the observation that they initiate a new TCU is limited to this collection.

grammatical design. They can be heard as extensions of the same imperative-TCU or a new TCU, which will be explained in the following sections.

The following touches on these practices. The outline is intentionally brief; it is presented to situate *repeating* – the practice of our concern – among some other verbal practices of pursuit identified in the overall corpus. Then we will explore the main message of the chapter: there are two types of pursuit sequences: persisting and gradual desisting.

Incrementing

After a recipient's resistance against complying, that is in the third position, the imperative-speaker may pursue compliance by incrementing. By definition, increments – grammatical extensions to a prior TCU by the same speaker (Couper-Kuhlen & Ono, 2007; Lerner, 2004; Schegloff, 1996, pp. 90-92; Schegloff, 2016, pp. 239-263) – recomplete an already completed TCU and, as such, renew the relevance of whatever type(s) of response which the host TCU had already made relevant. Therefore, an increment to an imperative turn renews the relevance of, broadly speaking, compliance or rejection. However, as mentioned earlier, increments which are added after the recipient's rejection, are biased towards the preferred alternative. An increment produced after a coparticipant's response is termed variously in the literature: "post-response pursuit of response" by Jefferson (1981) in the context of pursuit or "third turn increments" by Rickford (2019). Time and manner adverbials (e.g., *alan* 'now' or *tond* 'quick') and courtesy or pleading terms (e.g., *lotfan* 'please') are among the observed increments to the imperatives in this corpus. The following is an example of a post-response pursuit of response, in which Nasima tells her sister, Minuvash to bring their pet cat (line 1). After Minuvash's tacit rejection (an account for not performing the action, line 2), Nasima extends her already completed TCU with the courtesy word *lotfan* 'please' (line 3). By doing so, she renews the relevance of a response, this time, biased towards compliance.

Extract 2.15. R014a, 05.33

```
       01   NAS      minuvash bi-âre-sh.
                     Name     IMP-bring-3SG
                     Minuvash bring it.


       02   MIN      âghâ kenâre shufâzh    khâbid-e.=      [gonâh dar-e.
                     sir  next.to-Ez heater asleep-be.3SG   sin   have-3SG
                     But it's sleeping next to the heater.= Poor thing.


->     03   NAS                                             [<<p.l.cr>lotfan.>
                                                             please
                                                             please


       04            (1.0)
```

The imperative sequence, initiated at line 1, is already completed by the recipient's rejection implicative account – a dispreferred response (line 2). By adding the increment *lotfan* to an already responded-to TCU, the imperative-speaker, in effect, rules out the dispreferred alternative from the acceptable responses, renews the relevance of another response, and hence, pursues the preferred alternative only (line 3).

Addressing

Address terms perform various interactional functions (Butler, et al., 2021; Clayman 2010, 2013; Rendle-Short, 2010). But after an imperative-recipient's resistance against compliance, they can be at the service of prompting them to performing the directed action. Depending on their prosodic design, terms of address can be heard to be the continuation of the previous TCU or a new TCU. Without getting into details, an address term with a pitch configuration that roughly follows where the imperative-turn was ended can probably be heard to be incremented to the imperative-turn whereas if, for example, there is a pitch reset, it can be heard as a new turn, doing pursuing. In the former case, the address term

operates at the TCU level and in the latter case, it operates beyond the TCU level. In the following example, the address term is issued with a pitch reset and is heard as a new TCU.

Melisa (mother), and her two late-teenage daughters (Minuvash and Nasima) are sitting around a table playing cards. Nasima has stretched her leg such that her foot touches Minuvash's hip. Prior to this extract, Minuvash had asked, told, commanded her sister to move her leg away multiple times. Nasima refused to comply sometimes even by means of overt rejection tokens. The beginning of the extract shows Minuvash's yet another attempt to make her sister stop touching her hip. She uses a prohibitive (line 1, 'Nasima don't hit me with your foot'), along with a slapping action on Nasima's foot. Two seconds pass, but Nasima does not move her foot or show any other orientations to compliance. Minuvash then turns her gaze towards her mother, requesting an intervention from her (line 3, 'Look'). The first three lines are presented to show that the extract comes from a 'compliant' sequence and are not the focus here. The focus is on Melisa's imperative addressed at Nasima, designed to resolve the complaint (line 4, 'Nasima sweetie move to that side a bit more my daughter'). Almost two and a half seconds pass during which Nasima continues to look down at her cards, providing no response, embodied or otherwise (lines 5-7), which can, especially given the prior background, be implicative of resistance. Melisa, then, addresses Nasima using her name accompanied with an endearment word (line 8, 'dear Nasima'), following which Nasima produces a very affiliative compliance token (line 9).

Extract 2.16. RAY014b, 06.00

```
01   MIN   nasimâ pâ-t-o              +na-zan          + be man
           Name   foot-2SG-OBJ.MAR NEG-IMP.hit.2SG  to 1SG
           Nasima don't hit me with your foot.
     min                              +slap NAS's foot+
     nas   >>gaze at her cards>>

02         (2.0)

03   MIN    negâ *kon
            Look IMP.do.2SG
            look.
               *gaze at MEL-->

04   MEL   nasimâ mâmân* ye kam unvar     tar  bo-ro     dokhtar-am?
           Name   Mom  one bit that.side more IMP-go.2SG daughter-my
           Nasima sweetie move to that side a bit more my daughter
     min          -->*

05         (0.2)

06   MIN   pâsh     daghighan pish-e  man-e
           foot-3SG exactly   next-EZ 1SG-be.3SG      2.4 S
           her foot is just next to me

07         (0.6)

->  08   MEL   nasimâ jun
               Name   dear
               dear Nasima

09   NAS   chashm
           eye
           on my eyes ((polite OK))

10         (0.4)
```
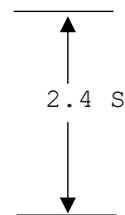
```
11    MEL    mersi azizam
             thanks sweetie
             thank you sweetie
```

Two response options that the imperative makes relevant includes compliance – Nasima moving away – and rejection in some form (line 4), but the recipient does not respond (lines 5-7). Her nonresponse, similar to a rejection, hinders the progressivity of Melisa's CoA, and is implicative of resistance. At this point, Melisa can abandon the sequence or expand the sequence to further her CoA. She opts for the second option by pursuing a response through an address phrase (line 8). The prosodic design of the address phrase conveys it to be a new TCU: Melisa's imperative turn (line 4) is produced with a final rise pitch movement (which is one recurrently observed pitch contour for imperatives in Persian). So it ends high in Melisa's pitch range at 360 Hz, 6.3 ST above her average F_0 level. The term of address, on the other hand, starts at her average F_0 level, at 235 Hz, 1 ST below her average F_0 level.

Addressing in this context renews the relevance of a response. Nasima's orientation to the term of address renewing the relevance of a response is clear in her compliance token (line 9, *chashm*). In Persian, a response to a summonsing term of address is typically *baleh* 'yes' or interjections such as *ha(n)*. By producing a compliance token *chashm* (line 9), rather than a response token, Nasima displays her understanding of the address phrase as a response pursuit device. This response pursuit is biased towards compliance. This is because pursuing a response after the recipient has already shown resistance to completing the CoA tacitly treats resistance as not adequate and, therefore, pursues only a preferred response.

When address terms are used after a resistance-implicative *non*response, as in this case, they can pursue a preferred response without going on-record (see Bolden et al., 2012 on off or on-record operations of pursuit). Since one conventional function of address terms is securing the addressee's recipiency, an imperative-speaker may use addressing to frame the recipient's nonresponse as a consequence of no recipiency (e.g., trouble in hearing etc.) rather than resistance. Put differently, by giving the recipient grounds to treat the term of address as a summonsing practice, the imperative-speaker makes an off-record attempt to pursue a preferred response. This can potentially save the interaction from getting overtly conflictual.

Accounting

Pursuing a preferred response can be done by means of accounting for the imperative (see Curl & Drew, 2008 on accounting for requests; Kendrick, 2020). Accounts for the imperatives often elaborate how it is that the imperative is relevant or benefits either of the participants in that context (see also Clayman & Heritage, 2014). After an initially resisted imperative, such an elaboration functions to pursue a preferred response. Again, accounts can be grammatically and prosodically designed as new TCUs or extensions to the previous imperative-TCU. The following example includes an account designed as a new TCU. Minuvash and Melisa (left and right, respectively, in Figure 2.6) are sitting on opposite sides of the sofa and playing a game. Prior to the extract, Minuvash reports a trouble on how she had attached the research microphone to her T-shirt's collar. As Melisa leans forward to assist, Minuvash claims she just solved the trouble herself, so Melisa leans back (Kendrick & Drew, 2016). However, she keeps her gaze on Minuvash's microphone even though Minuvash has resumed the game. The extract starts just after Melisa starts to lean forward again with both her arms stretched towards Minuvash's collar. This embodied conduct projects Melisa's intervention with the supposedly resolved trouble and displays her orientation to the trouble as still ongoing: Minuvash's necklace chain is still tangled in the microphone's hook. First, she issues the imperative *biâ* 'Come' (line 1), and after Minuvash resists, she accounts for the imperative (line 2). In line 2, the first *in* 'this' refers to the necklace chain, and the second one refers to the microphone hook.

Extract 2.17. RAY051b, 12.36

```
01    MEL    biâ+ɬ::,       +#
             IMP-come.2SG
             come,
```

```
mel        >>leans twrd MIN-->>
mel        >>strtches hand twrd MIN-->>
min          +lean bck+#
min          ±holds MEL's hands-->
fig                    #2.6
```


#2.6

```
->   02  MEL    hatâ ine-sh-o          bâyad az  tu in± darâr-im.
               even  this.of-3SG-OBJ.PRT should from in this remove.1PL
               even {yet} we should remove this from this.
        min
        min                                      -->±
        fig                                      #2.7
```


#2.7

The imperative *biâ* 'Come' is designed to reduce distance between the two participants (line 1). Melisa's arm-stretch towards Minuvash's neck conveys that the reduction of distance will be in the service of re-fixing the microphone trouble. Immediately after the imperative is issued, Minuvash performs two bodily actions that puts her resistance on display: she leans back, an action which has the opposite effect to that of the directed action, and she holds Melisa's hands which have already reached very close to the microphone (embodied annotation in line 1, Figure 2.6).[12] Melisa then accounts why a reduction of distance is still relevant (line 2, 'yet we should remove this {chain} of it {the necklace} from this {microphone hook}). Concurrently, Minuvash drops her own hands and allows Melisa's intervention (embodied annotation in line 2, Figure 2.7).

Accounting after a resisted imperative can be a practice to pursue compliance. In this extract, the account is designed as a new TCU: there is a pitch resent at the beginning of the account TCU and the Persian line includes no subordinating conjunction (such as *chon* 'because') to make it a dependent clause. It is syntactically a simple declarative. An account as a practice of pursuit treats a recipient's resistance to be potentially a consequence of their lack of epistemic access to the grounds that made the imperative relevant in the first place. By providing the grounds based on which the imperative is issued,

---

[12] The bodily action of leaning back is performed for a brief moment and cannot be easily captured in a still image.

the imperative-speaker makes the directed action accountable (intelligible, sensible) in that context, and this can potentially encourage compliance and/or make more resistance sanctionable.

Replacing

In the context of self-repair, Schegloff (2013) defines replacing as "speaker's substituting for a wholly or partially articulated element of a TCU-in-progress another, different element, while retaining the sense that [this] is the same utterance" (p. 43). Adapted for pursuits, replacing refers to the imperative-speaker's substituting for the imperative verb another imperative which has the same outcome of the prior, for example, the imperative *begin* 'say.PL' may be replace by *sohbat konin* 'talk.PL' or *sâket bash* 'be quiet' by *khafe sho* 'shut up'.

The following example illustrates this. Prior to this, Somayeh launched a telling about someone wearing an old coat. Marjan (Somayeh's adult daughter) halted the story to tacitly reprimand Somayeh for assessing the coat as old, an assessment which should not have been made in front of the research camera. After that, Marjan, the imperative-speaker, used different practices to get Somayeh, the recipient, to resume the storytelling, but Somayeh is still engaged in the prior sequence: she is putting forward candidate understanding on why she was just reprimanded. Replacing is used when, after multiple times issuing the imperative *begin* 'say' (line 1 and 4), Marjan uses the imperative *sohbat konin* 'talk' (line 7).

Extract 2.18. RAY019a, 07.07

```
->    01    MAR    be-g-in
                   IMP-say-2PL
                   say
             som   >>gaze at MAR

      02    SOM    na-g-am      yani    kohne boode
                   NEG-say-1SG meaning old   be.3SG
                   Do you mean I shouldn't say it was old?

      03           (0.3)

->    04    MAR    tch h.::<<giggles>begin        dige
                   INT                 IMP-say-2PL PRT
                   Tch say dige.

      05           (1.0)

      06    SOM    hâ [mi-g-i\
                   INTJ PRG-say-2SG
                   Ah, you mean\

->    07    MAR       [sohbat kon-in.
                      talk   do-2PL
                      Talk.

      08           (0.7)
```

Replacing is employed when the imperative speaker changes the imperative *begin* 'say.PL' (lines 1, 4) to *Sohbat konin* 'talk.PL' (line 7).[13] The imperative *begin* 'say.PL' is issued twice and after both times, the recipient shows resistance against compliance. The first time is in line 1. Somayeh shows resistance through continuing her own CoA and puts forward candidate understanding for being reprimanded (line 2). The second time is in line 4. Marjan lexically upgrades *begin* (line 1) to *begin dige* (line 4). The

---

[13] The plural suffix seems to be oriented to Somayeh's seniority.

particle *dige* in this context implies that the directed action should have been performed earlier.[14] Changing *begin* to *begin dige* is a practice of repetition which is not our focus here. This second time is also met with Somayeh's resistance. Somayeh's resistance is displayed in a noticeable silence during which the participants have mutual gaze (line 5), and her initiation of another candidate understanding (line 6). After these two 'failed' attempts to get Somayeh to relaunch the storytelling, Marjan changes the imperative *begin dige* 'say *dige*' (line 4) to *sohbat konin* 'talk' in a post-transition onset overlap (Drew, 2009) (line 7).

The first imperative (*begin* 'tell') is transitive and semantically more specific: one tells a story, a joke, a lie, and so forth. In this context, even though the object of the verb is not specified since it can be inferred from the context, the imperative still conveys that the recipient should tell *the story*. The subsequent imperative (*sohbat konin* 'talk') is intransitive in Persian and is semantically broader: it basically involves the process of producing words without specifying the particular words or topic that should be spoken. So in attempting to get Somayeh to relaunch the story, Marjan moves from a specific imperative to a more general one. Changing 'tell' to 'talk' can be oriented to the noticeable gap in line 5. This silence is attributable to Somayeh, so she can be seen to be resisting Marjan's CoA through not talking at all. It is in this context that the more general imperative 'speak' is issued to prompt her to talk (even if she does not tell the story).

Replacing an imperative turn can deal with potential problems in the design of the resisted imperative turn as the context unfolds. By adopting a new lexical and syntactic design for the pursuit turn, the imperative-speaker orients to her own imperative turn as one that may be 'not adequate enough', 'not strong enough', or otherwise problematic as the context is shaped by the recipient's response (e.g., the imperative may be too specific or too broad or vague, etc.). This aligns with Bolden et al.'s (2012) analysis on response pursuit. They note that self-initiated self-repair as a resource for pursuit tacitly treats aspects of the speaker's own turn design as potentially responsible for the recipient's inadequate response. However, in this collection, while some examples may fit this characterization, this is not always the case. For example, when the verb 'be quiet' is replaced with 'shut up', by choosing a semantically stronger and a socially 'offensive' verb, the imperative-speaker is treating the recipient accountable for failing to comply, and in fact re-designing the lexical composition of the turn is in the service of passing accountability for the halted progression of the CoA to the recipient as well as pursuing compliance. Thus, the precise interactional work that replacing achieves depends on the details of the lexical and syntactic design in the pursuit turn. A similar argument can be made about reformatting, presented below.

Reformatting

Again, in the context of self-repair, Schegloff (2013) conceptualizes reformatting as a grammatical change in the syntactic structure of the TCU, for example changing a declarative to an interrogative (p. 62). In pursuit sequences, reformatting happens when a different syntactic design than the one in the prior turn is used to prompt the recipient to perform the directed action. In this practice, too, a new TCU is initiated to pursue a response, so it operates beyond the TCU level. Let us examine this in an extract that was presented earlier (Extract 2.7, above). Farnaz issues imperatives to get her husband to get up and study his lessons. After issuing two imperatives (lines 3, 7), she uses a declarative to pursue a preferred response (lines 9-10).

Extract 2.7. RAY011a, 10.00, reproduced

```
->    03    FAR    pâsho       dars-e-t-o              be-khun.
                   get.up.2SG  lesson-of-you-OBJ.PAR   IMP-study
                   get up {and} study your lessons.

      04    AMI    na
```

---

[14] More investigation is needed for a more technical and specific function of this particle.

```
                      no
                      no.

        05            (0.6)

        06    FAR   ye  bâr dige    mi-g-am.
                    one time another PRG-say-1SG
                    I'm saying {it} another time.

->      07    FAR   pâsho     dars-e-t-o          be-khun.
                    get.up.2SG lesson-of-you-OBJ.PAR IMP-study
                    get up {and} study your lessons.

        08    AMI   ne-mkhâ-m.
                    NEG-want-1SG
                    I don't want to.


->      09    FAR   baar-e   sevom elâm    mi-kon-am  ke
                    time-of  third announce PRO-do-1SG that
                    I'm announcing {it} for the third time that

        10          pâsh-i     dars-e-t-o be-kho[on-i
                    get.up-2SG lesson-of-2SG-OBJ.PAR SBJV-study-2SG
                    you get up {and} study your lessons.

        11    AMI                            [<<p.all>ne-mikh-âm>
                                             NEG-want-1SG
                                             I don't want to.
```

After two rounds of resistance from Amir (lines 4, 8), Farnaz reformats the imperative 'get up {and} study your lesson' to a declarative with a similar interactional outcome as the imperative turns (lines 9-10). She upgrades the design of the declarative through a counting practice and a more 'formal' lexicon. The verb *elâm mikonam* 'I'm announcing' is rather formal for an interaction between two intimate participants. In this way, Farnaz is doing 'being distant' to Amir. The counting practice is culturally associated with higher entitlement. Through the combination of these lexical choices, Farnaz can be seen to be doing 'being a distant authority'. Using a declarative to upgrade an imperative directive goes against some claims made in the literature that associate imperatives with higher entitlement (e.g., Craven & Potter, 2010). The syntax of the TCU is only one of the design features, and the interactional affordance of the TCU, as this example shows, depends on a complicated web of various design features, including lexical ones.

Repeating

The last practice that this chapter presents is pursuit by *repeating* a prior imperative. As mentioned earlier, it is this practice which the rest of the thesis is based on. To reiterate from Chapter 1, the collection includes two main types of repeats: full and modified (see Kendrick, 2015; Robinson, 2013, Robinson & Kevoe-Feldman, 2010; and Stivers, 2005, on types of repeats). By full repeats I mean a verbatim reproduction of the whole of a prior imperative turn. The prosodic and embodied design features of the prior turn may change, but the verbal features are replicated without any addition, deletion, or reordering. Modified repeats are operationalized as those pursuit turns in which the imperative verb remains the same as a prior, but some other verbal features are reworked. Modified repeats are produced by inserting at least one extra verbal component in the pursuit turn, deleting at least one existing verbal component from a prior imperative turn, and/or reordering of the components of a prior imperative turn. Common lexical components which are added or deleted include particles, terms of address, adverbs, verb objects or pronouns. Extract 2.18 just above included an example of a modified repeat by insertion: the speaker added the particle *dige* to the prior imperative; the verb *begin* 'say' was changed to *begin dige* 'say dige'. The very first example of this chapter (Extract 2.1) included an example of modified repeat by deletion: the object of the verb was deleted: the imperative *dastam*

bus kon 'kiss my hand' was changed to *bus kon* 'kiss'. In the following, samples of full repeat and modified repeat by reordering are illustrated.

Even though in some other contexts a full repeat may covertly attribute a recipient's nonresponse to problems of hearing, in the context where there have been observable displays of resistance as instances in this collection, repetition is an on-record practice that, to borrow Bolden et al.'s (2012) terminology, "exposes" the recipient's failure in producing an adequate response. This is because by recycling the verbal design of a prior turn in full, the speaker tacitly treats the prior version as 'fitted' or 'adequate enough', one that can be even recycled with no modification. There are prosodic and embodied design features by which the speaker displays her stance towards this failure (we will focus on design features in Chapters 3-4). Full repeats are great loci for comparing prosodic-embodied design features in multiple productions of the same turn because of the natural control on the morpho-syntactic structure that they provide. These design features are the focus on the forthcoming chapters (Chapters 3-4). In this example, we remain focused on the verbal composition of the pursuit turn with only a passing mention of other multimodal features. Azita, who is visiting her friend Rosa, is talking to her (Azita's) brother, Kaveh, on the phone. In the middle of Azita's sequence with Kaveh (lines 1-2), Rosa produces the first imperative in focus (line 3, 'say hi to Kaveh'). When Azita resists, Rosa pursues a preferred response by repeating the very same imperative turn twice (lines 5, 7).

Extract 2.19. RAY041a, 12.51

```
      01   AZI   bi-shur                ((talking to Kaveh))
                 without-intelligence
                 idiot

      02   Azi   [khob be to che        ((talking to Kaveh))
                 PRT  to 2SG what
                 it's none of your business

->    03   ROS   [<<smiles>>be kâveh salâm be-resun>#
                 to Name  hello IMP-send.2SG
                 say hi to Kaveh.
            fig                                        #2.8
```
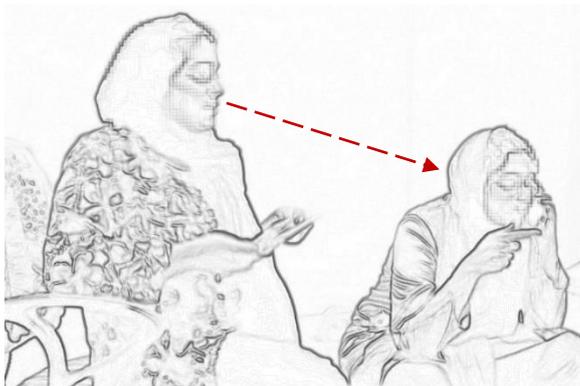


```
                                       #2.8

      04   AZI   be [t\                 ((talking to Kaveh))
                   to 2SG
                 it's none of\

->    05   ROS      [+be kâveh *salâm be-#+resun+ hehehe
                    to  Name   hello   IMP-send.2SG
                  say hi to Kaveh.
            ros    +gets closer to AZI----+,,,,,+
            azi              *fixated gaze at ROS-->
            fig                     #2.9
```
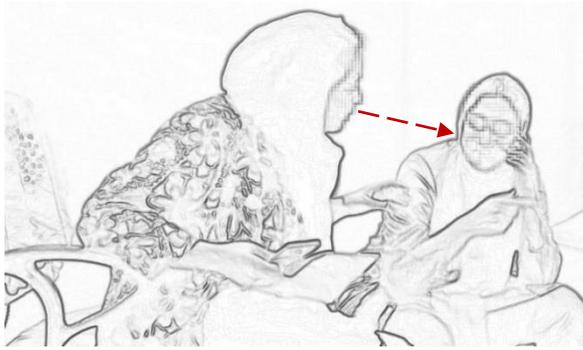
#2.9

```
       06          (0.1)

->     07   ROS    +<<lgh>be  kâveh +±#salâm  ±be-resun>*+±hehehe#
                          to  Name   hello    IMP-send.2SG
                        say hi to Kaveh.
            ros    +get clser to AZI+,,,,,,,,,,,,,,,,,,,,,+
            azi                                 -->*
            azi                   ±mock slap±,,,,,,,,,,,±
            fig                   #2.10                    #2.11
```
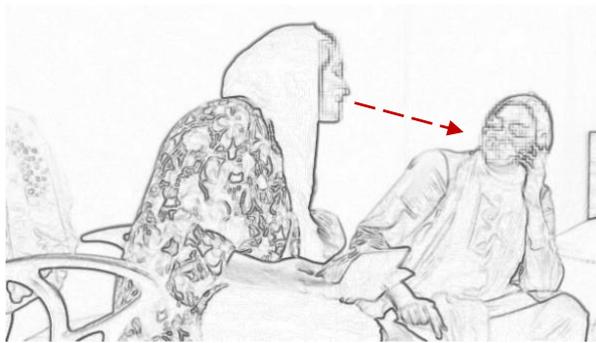

#2.10

#2.11

The choice and order of the lexemes of the subsequent imperatives (lines 5, 7) are exactly the same as the first (line 3). The accompanying smile to the original imperative turns into laughter in the subsequent versions and the imperative-speaker moves progressively closer to the recipient and talks progressively louder with each successive imperative; however, the actual lexical design and order do not change, so she is repeating a prior version in full.

Against the first imperative, the recipient shows resistance by continuing her own sequence with Kaveh (line 4, Figure 2.8), and the resistance against the second and third imperatives is manifested in Azita's embodied conduct: a long-fixated gaze (Figures 2.9-2.10); increasing distance between herself and Rosa (Figure 2.10), and a mock slapping gesture targeted at Rosa's mouth (Figure 2.11). With her gaze and

mock slap, Azita negatively sanctions Rosa's repeated interruption in her interaction with Kaveh (see Kidwell, 2005, on sanctioning looks; Afshari Saleh, 2020 on mock aggression as sanctioning). Her bodily actions also make it public that Azita's noncompliance (i.e., not passing Rosa's greetings to Kaveh) is not due to problems of hearing or understanding. Rather, it is an overt resistance against completing Rosa's initiated CoA. Repeating the very same imperative after Azita's display of resistance shows Rosa's orientation to the morpho-syntactic design of her imperative turn as unproblematic: it is treated as fitted to the context and the interactional outcome that it is pursuing such that it can be recycled with no modification. Whether the recipient is treated as accountable for her failure to comply with an already 'adequately' designed prior version depends in part on the subsequent imperative's multimodal design features and accompanying practices. We will explore this claim when two types of pursuit are introduced below (§2.8) and also in the following chapters (Chapters 2-3).

In a modified repeat as means of pursuing a preferred response, the imperative-issuer's orientation towards their prior imperative turn depends on the details of modification: what components are added, what are deleted, how the imperative turn is reordered and how all these manage the details of the unfolding context; therefore, the exact interactional affordance of a modified repeat is discussed on a case-by-case basis whenever they appear in the thesis. We have already seen some examples of modified repeats by addition (Extract 2.18) and deletion (Extract 2.1), so here only an example of a modified repeat by reordering is presented. The participants are having lunch and talking about tetanus vaccination. Amir is sitting next to a pot (indicated by a red arrow in Figure 2.12). Mona addresses Amir and issues the first imperative as she points to the pot (line 2, 'Amir give {it to} me'), by which she requests the pot. Amir does not fully comply (line 4-11), and Mona produces a lexically reordered version of the first imperative (line 11, 'Give {it to} me Amir).

Figure 2.12. Participants in Extract 2.20. Left to right: Haleh, Mona, Ebi, and Amir. A pot is to the right of Amir (indicated by the red arrow). It cannot be seen in the still image.

Extract 2.20. RAY049b, 04.10

```
     01   AMI    #*shomâ barâye chi  vâk[san * be-zan-in;         #
                  2PL   for    what vaccine  SBJV-take-2PL
                  why would you take a vaccine;


->   02   MON    #*                      [amir* +be-de      man?#
                                         Name   IMP-give.2SG 1SG
                                         Amir give {it to} me?
          ami    *gaze at EBI--------------*gaze at mon-->
          mon                                +point twrd pot-->
          fig    2.13                                        #2.14
```

71
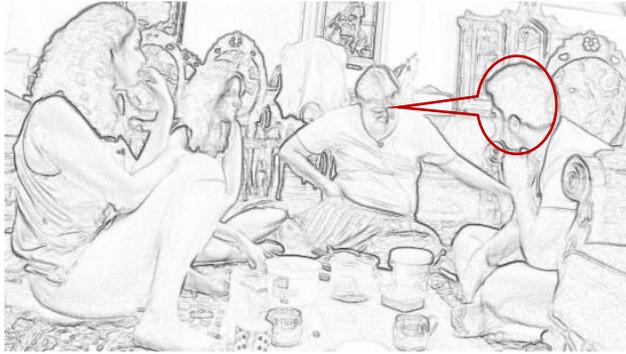
#2.13
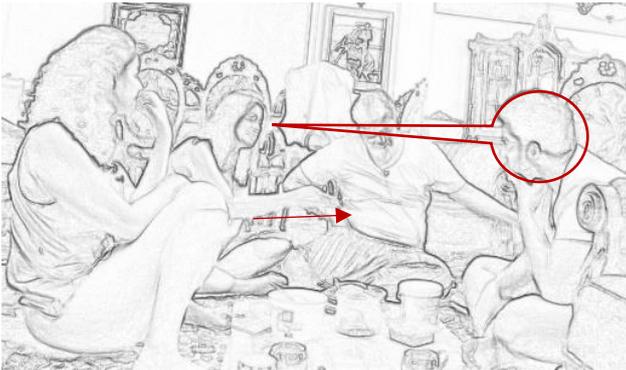
#2.14

```
03              (0.7)

04   HAL     shomâ *har  se-ta-tun+     ±bâyad bi-ây-n      ±
             2PL   each three-number-PL must  SBJV-come-PL
             all three of you must come
     mon                      -->+
     ami       -->*gaze at the pot-->
     ami                              ±move hand twrd pot±

05   HAL     ±kozâz*# [be-zan-in.
             tetanus    SBJV-take-2PL
             {and } take tetanus.

06   AMI          #  [barâ chi  be-zan-im.
                     for   what SBJV-take-1PL
                     why should we take {tetanus}.
     ami       -->*gaze at HAL-->
     ami     ±hold hand-->
     fig          #2.15
```
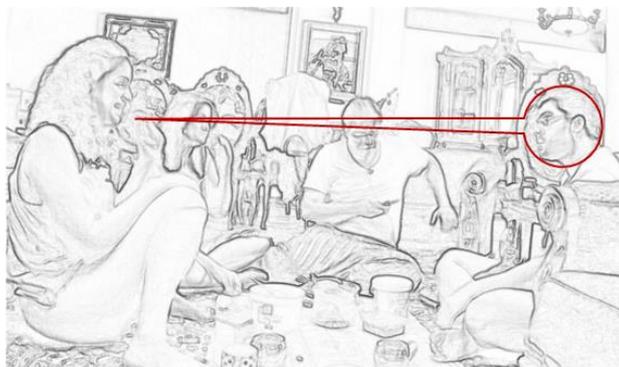

#2.15

```
07              (0.5)
```

```
08   HAL      Amir tarsu:: *[ây khodâ
              Name coward   INTJ God
              Amir coward Oh God

09   AMI                   *[bâz   barây chi  be-zan-im.
                            Again for   what SBJV-take-1PL
                            Why should we take {tetanus} again.
     ami                   *gaze at MON-->>

10            (0.7)

->   11  MON      #+be-de        man ±Amir.
                 IMP-give.2SG 1SG Name
                 Give {it to} me  Amir.
                 +stretches hand twrd pot-->>
     ami                          ±moves pot-->>
     fig      #2.16
```


#2.16
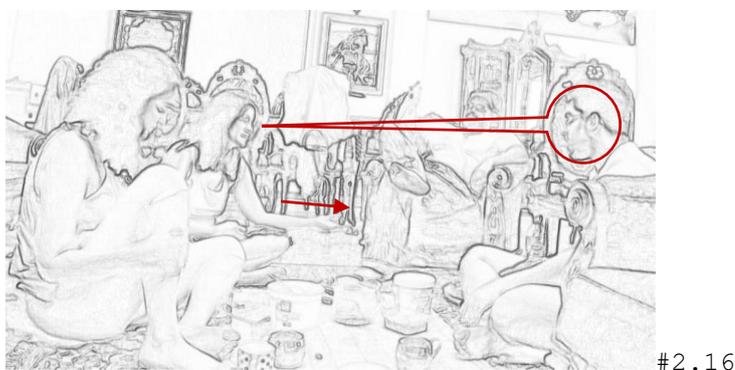
The modified repeat is produced by reordering one lexical component of the first imperative. The term of address which preceded the imperative verb in the first saying (line 2) is pushed to the end of the second imperative turn (line 11). In the first saying *Amir bede man* 'Amir give me', the address term preceding the imperative secures the imperative-recipient's gaze: the first imperative is issued in overlap with Amir's question addressed at Ebi (lines 1-2) and while Amir's gaze is directed on him (Figure 2.13). Immediately after Mona delivers the term of address, Amir's gaze is shifted towards her (line 2, Figure 2.14). It is only after securing Amir's gaze that Mona points towards the pot (line 2, Figure 2.14) and produces the rest of her TCU.

The first imperative turn does not include a grammatical object, and as explained in Chapter 1, Persian allows for a transitive verb without an object if the object can be inferred from the context. Given that they had not talked about the pot prior to this extract, the main way that the requested object becomes clear in this context is through Mona's pointing gesture. So the point complements the syntax of the imperative turn and functions similar to an embodied 'object' to the transitive verb.

Amir starts to show compliance: he shifts his gaze on the pot and moves his hand towards it: embodied preparation needed for transferring the requested object (line 4). However, upon receiving some information from Haleh, he suspends the transferring process (Raymond & Lerner, 2014). He stops his hand movement holds it near the pot, averts his gaze towards Haleh, and asks a question on the information (line 6, Figure 2.15). In the next slot, Haleh does not answer the question (lines 7-8). With no response provided from Haleh, Amir turns his gaze towards Mona and repeats his question this time addressing Mona (through his gaze, line 9) while the process of transferring the pot is still suspended.

With the compliance process suspended, Mona issues a subsequent imperative pursuing full compliance, an attempt to forward her own request CoA rather than Amir's question CoA (line 11, *bede man Amir* 'Give me Amir'). Contrary to the first imperative, the subsequent version is issued when Amir's gaze is already directed on Mona. Mona uses the same imperative verb phrase and a similar

pointing gesture to pursue full compliance (Figure 2.16), but the term of address is pushed back to the end of the TCU.

This address term in the subsequent version seems to have a different function than the one in the first version. Whereas the address term in the first saying seems to be designed to secure Amir's recipiency, the second seems to be oriented to Amir's accountability for keeping Mona's sequence on hold. More investigation is needed on the types and functions of Persian address terms, but using address terms when the recipient does not fully satisfy the relevance of a preferred response is observed in some other sequences in this collection as well (See Extract 2.11 above for a similar pattern). This extract was mainly presented to show an instance of a modified repeat by reordering.

Sometimes participants use a combination of operations – insertion, deletion, and reordering – that yields a modified repeat. For example, in Extract 2.11, the imperative *Amir âb beriz* 'Amir pour water' was changed to *beriz Amir* 'Pour Amir', so the operations of deletion and reordering were employed: the term of address is pushed back to the end of the imperative turn and the object of the verb is deleted. As mentioned in Chapter 1, this collection includes both full and modified repeats as means of pursuing a preferred response.

2.7. An Interim Summary

The sections above sketched the sequential structure of an imperative sequence: imperatives in the first position, a range of possible responses to them in the second position, and, in the third position, the imperative-speaker's potential orientations to resistance. Resistance was conceptualized as dispreferred and less than preferred responses. It was shown that one of these orientations is treating resistance as inadequate by pursuing a preferred response. Then different practices by which pursuing can be performed were exemplified.

The outlined practices of pursuit vary based on the operation they perform on the imperative turn. In incrementing, a grammatically fitting component expands the imperative TCU, and the speaker re-completes their imperative turn. While the literature on incrementing takes it to be an off-record practice of pursuit (e.g., Bolden et al., 2012), in the sequences in this study, the speaker can be overtly seen to be pursuing because the increments are employed after the recipient puts their resistance on display sometimes through overt rejection tokens. Addressing and accounting, I argued, can be among TCU level practices or beyond, depending on their prosodic and syntactic structure. Even when terms of address and account are designed as new TCUs, they can be still come across as appended to the imperative TCU. This is because contrary to imperatives in this thesis, terms of address and accounts do not by and in themselves get the coparticipant to do the directed action, rather, they are afforded that interactional function by virtue of their position after a prior imperative. In replacing and reformatting the whole of the imperative TCU changes through lexical and syntactic restructuring. In replacing, a prior imperative verb turns into a new semantically similar imperative. In reformatting, a new grammatical structure gets the participant to do the pursued action. So they operate beyond the TCU level. In repetition, the imperative verb and syntactic structure is kept intact, and the prior turn is recycled (in full or modified format). I argued that repeating a resisted action with the same verbal design indicates the speaker's treatment of the verbal design of their turn as adequate enough. Their orientations towards the recipient's resistance becomes clear in the multimodal design of their recycled turns and other actions that they perform alongside the pursuit. This is what the rest of this chapter is concerned with. We will see that depending on the accompanying actions, a pursuing party can hold the recipient at fault for not complying with an already adequately designed imperative and persist in advancing their CoA. Alternatively, they can back down and withdraw from the sequence without overtly holding the recipient accountable.

2.8. Persisting and Gradual Desisting

In analyzing pursuit, our analytic lens has so far been zoomed into the actual pursuit turns. In this section, I cast a wider analytic net to include other co-occurring practices or social actions that are frequently observed when an imperative-speaker pursues a preferred response. These co-occurrences matter since they can differentiate between two types of pursuit which I refer to as persisting and gradually desisting. These two pursuit types differ in respect to the stance the pursuing party puts on

display: a persisting stance displays that the imperative-speaker insists on furthering her CoA despite the recipient's resistance. A contrary stance is displayed by backing down and succumbing to the recipient's resistance. These stances project different trajectories for the imperative sequence. Persistence projects more sequence expansions while gradually desisting projects sequence closure. These stances are independent of the pursuit practices that were presented above. This means they can be identified in other practices of pursuit (i.e., incrementing, addressing, accounting, etc.) as well.

2.8.1. Persisting

With each subsequent imperative, the imperative-speaker renews the relevance of a preferred response, one that furthers the CoA projected by a prior imperative. While repeating the imperative, the speaker may laminate other practices or social actions in the pursuing action that show the pursuing party is not just pursuing but *persisting* in forwarding their CoA. Persistent pursuit projects more expansions to the sequence and in this way it differs from gradual desisting.

There is a wide range of different co-occurring practices and social actions which color a pursuit turn a shade of persistence, one group of which is those practices by which an imperative-speaker negatively sanctions the recipient's resistance. As elaborated in Chapter 1, negative sanctioning (equally termed as 'sanctioning' in this thesis) involves policing a coparticipant's conduct for violations of norms. Sanctioning holds an interactant accountable for such violations; it disapproves them and punishes the interactant for their nonconformity with acceptable conduct (Goffman, 1971/2010, p. 95). In this collection, a recipient's resistance against furthering a speaker's CoA is treated as the nonacceptable conduct. Sanctioning practices vary and can be as subtle as a frown or a click to extended turns of account solicitations and complaint formulations.

The following shows an example of persistence: the speaker pursues a preferred response and sanctions the recipient for her resistance. Farnaz and Amir (couple) are sitting on a sofa in the living room away from their home cinema system (Figure 2.18). The system's remote control is not within their reach and to turn the system on, one of them should physically get up and reach the system. Farnaz is reading a book and Amir is having some yogurt, both in silence. The extract begins with Amir's informing that he fancies listening to Enrique's *Ring My Bell* song, responsive to which Farnaz puts forward a solution for Amir, that he plays the song. This is when Amir issues the first imperative in focus: 'you yourself go and play it' (line 5). When Farnaz does not show immediate compliance, Amir employs a combination of bodily and verbal actions to sanction Farnaz's resistance. Then, he reissues the imperative (line 16). Thus, sanctioning accompanies the production of the subsequent version, and it is this combination that displays the speaker is *persisting* in furthering his CoA.

Figure 2.18. Participants in Extract 2.21. Farnaz (left) is reading a book. Amir (right) is having yogurt. To the left of them (behind the pillar), there is the home cinema system. The system's control is placed on the leftmost speaker.

Extract 2.21. RAY011a, 17.52

```
       01   AMI    .h âkh  âkh   cheghad   havas-e   âhang-e
                     INTJ INTJ how.much fancy-Ez  music-Ez

                   ring my belle    ₒenrikaro     kard-am
                   ring my bell-Ez Name-OBJ.mar do-1SG
                   oh how I've fancied Enrique's Ring My Bell
            far                      ₒgaze down at the book-->

       02   (1.0)

       03   FAR    biâ          be-zâre-sh
                   IMP-come.2SG IMP-play-3SG
                   come play it

       04   (1.2)

->     05   AMI    khob *khodet   bo-ro   play-e-sh      kon    dige
                   PRT   yourself IMP-go.2SG play-Ez-3SG do.2SG PRT
                   you, yourself, go play it
                       *gaze at the audio-player device-->

       06   AMI    faghat bâyad play-e-sh   kon-i *
                   only   must  play-Ez-3SG do.2SG
                   you must only press play {button}
                                           -->*

=>     07           *    (3.0)   *•+(2.0)
            ami   *gz at yogurt*fixated gaze at FAR-->>
            ami                 •frowns at FAR-->>
            ami                 +hold the spoon immobile near his mouth-->>

       08   FAR    chi? play-e-sh   kon-am?
                   what play-Ez-3SG do-1SG
                   what? {should I} play it?[15]

       09   (0.3)ₒ±(1.7)
            far    -->ₒgaze at AMI-->
            far       ±picks up the bookmark-->

       10   FAR    amir in   ketâbe   kheili bâhâl-e.
                   Name this book-ART very   cool-be.3SG
                   Amir this book is very cool.

       11   (0.6)

       12   FAR    be-khun
                   IMP-read.2SG
```

---

[15] In Persian, verbs have person suffixes. Additionally, Persian is a pro-drop language. This is not the case in English. So in line 8 there is no easy way to show the subject in English translation without the free morpheme 'I'. However, adding the pronoun 'I' to the English translation (e.g., I play it?) would probably add an element of focus on the pronoun 'I', which does not seem to exist in the Persian line. The Persian line could have been produced with the free morpheme *man* 'I' and in that case, a sense of focus on 'who should play the music' could probably have been conveyed. But the speaker has produced the repair in line 8 without that free morpheme which makes her repair target the action (playing) rather than the agent (who should play the music). In the English line, the pronoun 'I' is included just to make the English translation grammatical since the Persian line is grammatical, but it is put in curly brackets to show that that pronoun does not exist in the Persian line the way it appears in the English translation.

```
                 read {it}

      13         (0.4)

=>    14  AMI    be gushvâreye chapesh gereft
                 to earring-Ez left-Ez-3SG take.PST.3SG
                 she didn't effing care.

      15         (0.5)₀±
          far    -->₀
          far    -->±

->    16  AMI    ±<<f, h>bo-ro      play[e-sh   ko:::n>±
                 IMP-go.2SG play-Ez-3SG do.2SG
                 go play it
          ami                               -->
          far    ±puts bookmark on the page-----------±
```

Before analyzing the imperative-speaker's persistence, I first explain the recipient's actions that manifest resistance against completing the imperative-speaker's CoA: (a) a rather long silence during which she is engaged in her own CoA, reading her book, and (b) in initiating a new assessment sequence on the book rather than completing Amir's CoA. Starting with line 1, Amir's informing provides an opportunity for Farnaz to offer playing the music or to perform the action (Kendrick & Drew, 2016). Amir's informing, therefore, functions as an off-record recruitment. The opportunity is not taken up by Farnaz. Instead, she offers Amir a solution produced in an imperative format, suggesting that Amir play the music (line 3). So Amir's first off-record attempt to get Farnaz to play the music fails. He then goes on-record and produces the first imperative in focus (line 5). This imperative counters Farnaz's imperatives and hence reverts the flow of relevancy (Schegloff, 2007, pp. 16-19). Thus Amir upgrades his recruitment method from creating an opportunity for performing the action to creating an interactional slot where compliance or rejection becomes relevant (Kednrick & Drew, 2016). But he mitigates the action by framing it as a low contingent task (line 6, 'You must only press play') (Curl & Drew, 2008). In the next slot, Farnaz's response is expected, but she shows resistance against completing Amir's CoA by keeping her gaze down on the book and displaying engagement in her own reading activity.[16]

Next, Farnaz performs a combination of actions that has both compliance and resistance implications. Her compliance-implicative move is a repair initiation on Amir's imperative (line 8). Although this runs counter to what the literature on repair and affiliation has shown (e.g., other-initiated repair as preliminary to dispreferred responses, Kendrick 2015), the evidence here suggests that other-initiated repair can preface a response that has affiliative implications. By initiating a repair, Farnaz tacitly shows that her noncompliance (or delayed uptake) so far may be due to a lack of hearing or understanding and thus displays some initial orientation to a potential compliance. This is evidenced in her next move, when she reaches for her bookmark, which projects a halt in her reading activity, potentially to perform the directed action. (And indeed her reaching for the bookmark is the start of a CoA in which she puts the bookmark in the book, closes the book, puts the book away, gets up, moves towards the system, and eventually plays the music; all these happen after the presented extract). Her resistance is manifested in her other co-occurring actions that delay a full compliance: firstly, the repair initiation is performed as Farnaz is still reading her book, so she does not display full engagement in Amir's initiated CoA. Secondly, even when she finally raises her gaze towards Amir (line 9), she initiates a new sequence – an assessment on the book (line 10) and a suggestion that Amir should read the book (line 11). It is this

---

[16] If either of the participants is to perform playing the music, she or he must get up. So the fact that Farnaz is closer to the system (just by two steps or so) does not seem to justify Amir's recruiting her. In Persian culture, as probably in other cultures, there are instances where recruitment is done for the convenience of the recruiter (at least in the first glance). Sequences like this that contain recruitments for convenience should be the topic of future investigations.

combination of pull and push effect that conveys Farnaz's actions as having both conformity and nonconformity implications; hence, analyzed as a less than preferred response.

There is a multimodal Gestalt, including embodied and verbal practices, by which the imperative-speaker sanctions the recipient's resistance after the first imperative (Mondada, 2014 on multimodal Gestalts). The embodied package comes after Farnaz's initial display of noncompliance and runs till after the extract (lines 7 onwards for more than 11 seconds). Amir whose gaze was just on the yogurt he was eating, shifts his gaze towards Farnaz and holds his gaze at her in a 'fixated' manner even though Farnaz's gaze is down at her book (see Kidwell, 2005, on sanctioning looks); concurrently, he frowns at Farnaz (see Kaukomaa, Peräkylä, & Ruusuvuori, 2014, on frowns foreshadowing a problem), and above all, one embodiment which shows that Amir is actually doing a distinct social action with his gaze and frown is his sudden halt in his yogurt eating action. Amir, who had brought a spoon full of yogurt very close to his mouth (preparation phase before eating), holds the spoon just there immobile. Holding his yogurt-full spoon in that instable position has two functions: firstly, by virtue of its co-occurrence with the other two embodied actions (gaze and frown), it tacitly shows that those two bodily actions are indeed used to perform a *distinct* social action (i.e., negative sanctioning) which can even put the ongoing activity of yogurt-eating to a halt. Secondly, it, itself, is a part of the embodied Gestalt by which Amir negatively sanctions Farnaz's noncompliance: by holding his spoon just a few centimeters away from his mouth, he maintains his whole body totally still, and thus he tacitly shows that his gaze at Farnaz is not doing a simple 'looking', (awaiting a response, for example) rather it is fixated and hence it is doing 'staring'.

The sanctioning multimodal Gestalt also includes a verbal package: referring to Farnaz, the addressee, with a third person pronoun and an offensive expression (line 14, 'She didn't effing care'). This expression is produced after Farnaz evaluates the book and proposes that Amir should read it, and hence displays her continued engagement in her own CoA. Amir, who is still holding the spoon immobile near his mouth and continuing his fixated gaze and frown at Farnaz, firstly refers to Farnaz, his present coparticipant, with a third-person pronoun (i.e., 'she' rather than 'you'); a practice which is sometimes used in the context of complaint in Persian. Secondly, he resorts to a taboo idiomatic expression which implies that someone does not care about something at all (in this case, it implies that Farnaz does not care about Amir's imperative at all). Amir euphonizes the taboo expression by replacing the actual taboo word with 'earring' (line 14, the literal translation is 'She {Farnaz} diverted {my imperative} to her left earring' and one of the closest free translations can be 'She {Farnaz} didn't effing care about my imperative'). Amir's own orientation to the 'offensive' nature of the expression is displayed in the euphemism that he applies for the actual taboo word which is attempted to be captured in the translation 'effing' instead of the actual F-word (see Hoey et al, 2020 on the use of F-word in a similar sequential context). Therefore, Amir uses a combination of embodied (fixated gaze, frowning, and holding his body immobile) and verbal (a third person pronoun, addressing a present coparticipant, and a taboo expression) resources to negatively sanction Farnaz's initial noncompliance and later for a delayed compliance.

The imperative-speaker then re-issues the imperative in a modified format with prosodically loud, high, and elongated design features as he maintains his embodied sanctioning configuration (line 16). The co-occurrence of the prosodic and embodied features suggest that the upgraded prosodic design can be also analyzed as a part of the multimodal Gestalt by which Amir sanctions Farnaz's delayed compliance (more on multimodal upgrading and persistence in Chapter 3).

In short, a pursuit turn *renews* the relevance of a response; a pursuit turn post-resistance creates the relevance of only a *preferred* response; and finally, a pursuit turn post-resistance accompanying sanctioning practices *escalates* a mere clash in the compatibility of the two participants' courses of action to a point where more and more explicit practices are used to further the CoA projected by the imperative. The imperative-speaker's policing behavior shows that another noncomplying response is not expected in the newly offered response position since a prior round of resistance was just condemned. By designing a pursuit turn with sanctioning practices, the pursuing party puts a persisting stance on display. Such a stance projects further sequence expansion rather than closure.

Note that escalated pursuit is only one type of persistent pursuit. Even though in this thesis, I will only present examples of escalation, there are other more affiliative practices that can actualize a persistent pursuit.[17] These practices include producing positive assessment of the imperative-recipient (e.g., 'you're such a nice person) and pleading (e.g., 'I beg of you').

2.8.2. Gradual Desisting

A second type of pursuit identified in the corpus is gradual desisting: an incremental disengagement actualized in a process of backing down. Sometimes, an imperative-speaker pursues a preferred response by repeating the imperative after the recipient's resistance, but their co-occurring conduct or the design features of their pursuit turns project a contingent sequence closure and a likely withdrawal from the sequence. That is even though the mere repetition of the imperative after orientations to resistance actively pursues a preferred response, the design features and co-occurring actions project a forthcoming sequence closure and a potential withdrawal. In this pursuit type, rather than a straightforward withdrawal from the sequence – that is rather than abandoning the sequence after the first sign of resistance – withdrawing occurs in a gradual process. The process involves the imperative-speaker making some more 'attempts' in securing a preferred response but the attempts can be seen to be undergoing an attrition.

There is a variety of design features, actions, or practices by which a subsequent version projects a contingent withdrawal: withdrawing gaze (Rossano, 2012), retracting body part to home position (Schegloff, 1998), resuming a prior suspended activity (Raymond & Lerner, 2014), abandoning turns without reaching a syntactic completion (this thesis, Chapter 4), and speaking lower or quieter are among these practices and design features (Walker, 2012) (See Chapter 4 for a more comprehensive review). The following example shows some of these. Azita is visiting her friend Roshana (both late teenagers). Roshana had also invited another friend of theirs who did not accept the invitation. Azita makes a phone call to this third participant to talk about their university grades. Roshana is standing on the other side of the room, holding up her own mobile phone, presumably working with it (Figure 2.19). Just before this extract, Roshana, the imperative-speaker lowers her mobile phone. This shows a halt or an end to her ongoing CoA – checking her mobile phone. Then the extract begins with her first imperative issued at Azita: 'hang up the phone' (line 1). This is the imperative in focus. It is recycled after Azita's display of resistance, but the subsequent version is accompanied with co-occurring conduct that projects Roshana's withdrawal from the sequence. The lines in focus are indicated by single arrows (–>). Roshana's forward and backward bodily movements are indicated by double and triple arrows, respectively (=>; ≡>).

---

[17] Persistence pursuit is not necessarily disaffiliative (i.e., it is not necessarily escalated). There are other practices by which persisting may occur. Here are some schematic examples from the corpus.

- Orienting to seriousness:
    A: Do X
    B: Resistance
    A: *jedan* do X ('seriously' do X)

- Affective Pleading:
    A: Do X
    B: Resistance
    A: *jun-e man* do X ('if you care about my life' do X).

- Assessing the imperative-recipient positively:
    A: Do X
    B: Resistance
    A: Do X. *to ke inghad khobi* (Do X 'you're such a nice person').

Similar to sanctioning practices, through these orientations, too, the pursuing party *persists* in furthering their CoA; that is, the imperative sequence is likely to get expanded if more noncompliance occurs. However, unlike negative sanctions, these practices do not condemn noncompliance; rather, they display the pursuing party's affective orientations. Future research should investigate how precisely these practices may achieve compliance.

Figure 2.19. Participants in Extract 2.22. Azita (left) is talking to a third participant on the phone. Roshana (right) is holding her own mobile phone up, presumably working with it. Roshana's head is out of the camera shot.

Extract 2.22. RAY041e, 04.14

```
->    01   ROS    <<f>#gushi-o          ghat kon.
                  handset-OBJ.mar cut  do.IMP.2SG
                  hang up the phone.
             fig      #2.20
```


#2.20

```
      02   ROS    +be-gu         har  ki  injâ na-yumad-e\
                  IMP-say.2SG every one here NEG-come-PRF.3SG
                  tell her whoever that has not come here
=>                +walks towards AZI-->

      03          (0.2)

      04   ROS    †begu          har  ki  injâ na-yumad-e#
                  IMP-say.2SG every one here NEG-come-PRF.3SG
                  tell her whoever that has not come here
=>                †stretches hand towards AZI's phone-->
             fig                              #2.21
```


#2.21

80

```
05   ROS    ±alân-am   +†ne-mitun-e az rotbe-hâ khabar dâr    she>±#
            now-also NEG-can-3SG   of grade-PL news   having become.3SG
            can't know the grades now
      azi   ±lift her leg up----------------------------------±
≡>    ros          -->+walks backward-->
≡>    ros          -->†retracts hand-->

->   06   ROS    +<<p>>#ghat kon       gushi-o>±
            cut   do.IMP.2SG handset-OBJ.mar
            hang up the phone
≡>    ros   +stands away from AZI-->>
      fig          #2.22
```


#2.22

```
07          †*(1.2)
ros   -->†brings her own mobile up-->
         *gaze at her own mobile-->>

08   AZI      †âre#   ((Talking to MAH))
            yeah
            Yeah
ros   -->†hold her own mobile up-->>
fig          #2.23
```


#2.23

In this extract, the imperative-speaker, Roshana, performs a series of forward-moving actions toward the recipient, Azita, just after the original imperative and a series of backward movements preceding the subsequent version. The first imperative is issued after Roshana stops working with her mobile phone, projecting a new CoA (Figure 2.20). It is designed loud (captured by <<f> in the transcript). Then, as Roshana persists in getting Azita to further her (Roshana's) CoA (lines 2-4), she produces a combination of movements that project a grabbing action, targeted at Azita's mobile: firstly, she walks towards Azita (line 2-3) and then she stretches her hand towards the mobile as she continues approaching Azita (line 4, Figure 2.21). These forward-moving actions project a CoA in which Roshana grabs the phone and potentially hangs up the phone herself. So the movements can be seen to be at the service of advancing the CoA that the imperative has initiated.

Azita displays resistance against compliance. Firstly, although her gaze is directed at Roshana, displaying recipiency, she continues to hold the phone on her ear, doing listening to the third participant (Figure 2.20).[18] Secondly, right when Roshana is close enough to grab the phone, Azita lifts her leg up and stretches it towards Roshana, depicting a mock kicking gesture (line 5). Even though her leg could easily reach Roshana's body, Azita stops stretching her leg just before it touches Roshana's body, hence 'mock' (see Afshari Saleh, 2020; and chapter 6 in this thesis on 'mock aggression'). The kicking gesture is designed to figuratively 'push' Roshana back and put a stop to her reaching movement. Therefore, Azita's resistance is displayed in holding the phone on her ear (i.e., doing listening to the third participant) and her mock kicking by which she refuses to give up the phone to Roshana.

Following Azita's resistance, Roshana performs bodily movements that project a contingent withdrawal from the imperative sequence: she moves her body back as she retracts her hand, hence distancing herself from Azita (line 5). These distancing movements indicate that she is not going to grab the mobile anymore. They precede a prosodically quiet (captured by <<p> in the transcript) subsequent version as Roshana stands further away from Azita (line 6, Figure 2.22). Therefore, even though recycling the imperative, post-resistance, creates expectation for a preferred response, Roshana's co-occurring behavior indicates her lower commitment to pursuing compliance. This combination of pursuing a preferred response together with conduct that projects the pursuing party's contingent withdrawal depicts a process of gradual desisting: a process during which the imperative-speaker makes more attempts to forward their CoA, but the attempts are undergoing attrition. This type of pursuit projects a potential sequence closure. Indeed, following this, Roshana holds up her own mobile, shifts her gaze at it, and gets engaged with it (line 7, Figure 2.23; see Goodwin, 1981; Rossano, 2012; Schegloff, 1998; on bodily displays of disengagement). Azita also continues her talk with the third participant (line 8; Figure 2.23). By re-engaging in her previous CoA, Roshana tacitly treats the imperative sequence as closed.

Gradual desisting is a type of pursuit since the subsequent imperative still creates expectation for a preferred response; however, its co-occurring features puts the speaker's potential relinquishing stance on display. As we saw earlier, pursuit turns that display the speaker's persisting stance, project more sequence expansions; however, with gradual desisting, a sequence closure is more likely. Thus, the two types of pursuit differ in terms of the projected trajectory of the sequence. We explore these two pursuit types in more depth in the forthcoming chapters.

2.9.   Discussion

This chapter introduced the sequences that are explored in this thesis: pursuing compliance after the recipient's resistance. Firstly, an overview of the structure of these sequences was presented. We started with the imperative turn itself in the first position; then we continued with the second position where a range of different responses to the imperative can be produced and then we focused on the third position where the base imperative sequence may be expanded. We saw that response pursuit of the kind we are investigating in this thesis resides in sequence post-expansions. Next, some verbal practices of pursuit – incrementing, addressing, accounting, replacing, reformatting, and repeating – were briefly outlined according to the operation that they perform on the base first pair part. And then we aimed our attention to pursuit by repetition and distinguished between two types of pursuits – persisting and gradual desisting.

It is well-established that the sequential development of a sequence affords the possibility for each next speaker to display their stance towards a prior turn. A third position, for example, is the locus where a speaker can show orientations towards a second-positioned response as acceptable or not (Schegloff, 2007). Additionally, it is already shown that a mere response pursuit, after a coparticipant's (non)response, indicates the pursuing participant's treatment of the (non)response as nonacceptable (Bolden, et al., 2012; Davidson, 1984; Etelämäki & Couper-Kuhlen; Hoey, et al., 2020; Keel, 2015;

---

[18] The third participant is continuously talking to Azita from the other end of the line throughout the whole extract; however, only the interaction between Roshana and Azita is presented.

Peräkylä & Ruusuvuori; 2012; Pomerantz, 1984b; Stivers & Rossano, 2010). The findings here further our understanding on response pursuit a step forward: although pursuing tacitly treats the (non)response as nonacceptable or inadequate, the pursuing party can show two different stances towards this non-acceptability or inadequacy and these two stance types can, in turn, distinguish between two types of pursuit sequences with different trajectories.

In one type of pursuit, called persistence, the pursuing party treats the inadequacy of the (non)response as sanctionable, in which case, the pursuit turn is accompanied with sanctioning practices or actions that holds the recipient at fault for their (non)response. By pursuing compliance and at the same time sanctioning noncompliance, the pursuing party puts a persisting stance on display which projects more sequence-expansions in which the pursuing party makes further moves to further their CoA despite the recipient's resistance against it. This can potentially project an escalation in disaffiliation (Hoey, et al., 2020; Whitehead, et al., 2018).

In the second type of pursuit – gradual desisting – the pursuing party puts a relinquishing stance on display. Negative sanctioning rarely occurs in this type of pursuit. This stance is embodied in various design features, practices, or actions that project the pursuing party's withdrawal from the sequence: body retraction, gaze aversion, resuming a previous suspended action are among such practices and actions (Goodwin, 1981; Raymond & Lerner, 2014; Rossano, 2012; Schegloff, 1998). In this pursuit type, rather than a straightforward abandonment of the sequence after resistance, the pursuing party withdraws incrementally or in a process, so withdrawal is delayed. Even though the speaker is verbally engaged in pursuing; that is even though they recycle the imperative turn, their other co-occurring actions show their growing disengagement from the sequence. So in this type of pursuit, a contingent sequence closure is projected.

This chapter thus shows that sequences that are described under the umbrella term 'pursuit' in the literature do not have the same valence in terms of the projected trajectory of the sequence. In some, more sequence expansions are likely and in some a sequence closure. When pursuits occur in conflict or conflictual situations, the first type of pursuit projects escalation in disaffiliation and with the second type, a de-escalation and an exit from the disaffiliative encounter is projected. The projective potentials of these two types of pursuit have implications for the management of conflictual encounters.

For example, in Extract 2.22, an imperative-speaker commands the recipient to end her conversation with a third party on the phone. When the recipient resists, the imperative-speaker persists in ending the conversation by setting off to grab the mobile phone from the recipient, which in effect ends the recipient's conversation with the third party. The recipient deals with this display of persistence through a mock kicking gesture (designed to keep the imperative-speaker away from the phone, Afshari Saleh, 2020; this thesis, Chapter 6). The imperative-speaker then backs down, retracts her body, and stands relatively away from the recipient, and the recipient initiates to drops the 'protective' kicking. Ceasing the protective kicks indicates that although the imperative-speaker verbally pursues compliance, the recipient is not treating the act of pursuit as a physical 'threat', and this further shows the recipient's orientation towards the sequence as heading towards de-escalation. Thus, pursuing compliance in and by itself does not embody escalation; rather, it depends on the pursuing party's other co-occurring practices and actions.

Another piece of evidence that shows participants' emic orientation towards persistence as a pursuit type that projects escalation is third parties' interventions. A third party's intervention can be either sought or offered. In Extract 2.16, we saw an instance of sought intervention. In that extract, after multiple 'failed' attempts in getting a coparticipant to move her foot away, the participant finally slaps the coparticipant's foot. The co-occurring slap sanctions the coparticipant's refusal to change the positioning of her foot. The slap thus indicates the participant's persisting stance in getting the coparticipant's foot away and embodies a rather escalated attempt. Further escalation and persistence can potentially result in a full-blown fight. At this point a request for intervention is made. The

participant asks her mother to act as the 'judiciary' and resolve the rather heightened clash. So the participant herself orients to her act of persistence as a move which heads towards an escalated state such that an intervention of a third party is sought to settle the clash (see also Hoey, et al., 2020; Whitehead, et al., 2018 on turn design and escalation).

The next two chapters unpack these two pursuit types in more details. We will see that in addition to the co-occurring actions and practices, the multimodal design of a pursuit turn itself systematically differ while persisting as opposed to gradually desisting. In persistence, upgrading the multimodal design features of the pursuit turns is likely whereas in gradual desisting, multimodal downgrading is more frequent.

Chapter 3. Multimodal Upgrading in Pursuit Sequences

3.1. Introduction

Many domestic clashes have their origins in a 'minor' collision of pursuits: a simple wrong turn down the path that leads to escalated discord or a small spark that grows into a fire. But however they are ignited, many conflicts seem to have similar 'growth cycles'. The paths to an escalated state of affairs seem to have some common landmarks. Our interest in this chapter is in the process of escalation – the path that steers interactants into a situation in which more and more overt disaffiliative exchanges occur. More specifically, we examine one particular landmark on this path: *multimodal upgrading*. We continue to explore the projective potentials of the design of pursuit turns and seek to answer this question: what implications do the multimodal design features of a pursuit turn have for the structure of the sequence?

Consider an initial example where a mere clash of pursuit over who uses the phone grows into an escalated attempt for getting hold of the phone by one of the participants. The results of the National University Entrance Exam are announced. Azita and Roshana (friends) are discussing which universities would probably admit them. Roshana's mother is sitting in front of her, holding Roshana's phone and checking the details of her marks (Figure 3.1). Note how Roshana upgrades the design features of her subsequent imperative turn (line 13) when her mother resists submitting the phone twice (lines 2 and 11). Azita and Roshana's turns about different universities are not our focus.

Figure 3.1. Participants in Extract 3.1. Left to right: Azita, Roshana, and Roshana's mother. Roshana's mother is holding Roshana's phone.

Extract 3.1. RAY041b, 11.25 [L stands for left and R for right]

```
        01    AZI    +*†dâneshgah    do[†#la±ti  #]tor†bat   ††ne-mitun-i ber-i=
                      university   governmental city.name   NEG-can-2SG go
                      you can't go to Torbat's public university


->      02    ROS                       [†#be-±de,#]    †        †±
                                         IMP-give.2S
                                         give {it}
              mom    >>+holds ROS's phone with L hand -->
              mom    >>*gaze on cell phone-->
              ros      †.............†arm stretched†,,,,,,,†
              mom                       ±extends R fingers±
              fig                    #3.2      #3.3
```

85

#3.2

#3.3

((6 lines omitted in which AZI says a city does not have a public Uni))

```
     10   ROS     na †dolati-am      dar-e.†=
                  no  governmental-too have-3SG
                  no {it} has a public {university} too.
          ros         †.....................†

->   11   ROS     =#†mâm   be-    ⊤de† >         †⊤#
                  Mom    IMP-give.2SG
                  Mom give {it};
          ros      †arm stretched---†,,,,,,,,,,†
          mom                    ⊤retract body⊤
          fig      #3.4                          #3.5
```


#3.4

#3.5

```
     12          (2.0)

->   13   ROS     #†。<<f, h>khob+ be-de*       dige.> h.†
                          PAR   IMP-give.2SG    PAR
                          khob give {it} dige
          ros      †grabs phone------------------------†
          ros      。frowns-->>
          mom                 -->+
          mom                      -->*
          fig      #3.6
```


#3.6

```
     14          *•  (3.3)  *•
          mom    *gaze at ROS*
          ROS    •gaze at MOM•
```

86

Roshana addresses her mother with the imperative *bede* 'give {it}' three times (lines 2, 11, 13). The first two attempts face with her mother's embodied resistance against compliance. Following two rounds of resistance, the third imperative is issued with upgraded multimodal design features. In the first turn, the imperative *bede* is issued as Roshana extends her hand towards the phone (line 2, Figure 3.2). Mom resists by extending her fingers and 'ushering' Roshana's hand away from the phone (Figure 3.3). Then Roshana retracts her hand fully. In the second version, the term of address *mâm* precedes the imperative and again Roshana's hand is stretched towards the phone (line 11, Figure 3.4). This time Mom shows resistance by retracting her body and hence distancing the phone from Roshana's hand (her body retraction is designed fast and short. It is not easily captured in a still image). Again, Roshana fully retracts her hand (Figure 3.5).

The multimodal upgrade occurs in Roshana's third attempt (line 13): firstly, the particles *khob* and *dige* are inserted. *Khob* can function as an appositional token and *dige* can indicate that performing the directed action is already overdue. The turn is designed louder and higher in the speaker's pitch range compared to the two previous imperative turns. Instead of just stretching hand towards the phone, Roshana grabs the phone off her mother's hand (Figure 3.6) (see Hoey et al., 2020, p. 14, for a similar kind of upgrade by means of expletives). Thus even though her imperative makes a response relevant, her embodied action leaves relatively less opportunity for a response unless Mom also upgrades resistance. In this last subsequent version, Roshana furrows her brows by which she can be seen to be negatively sanctioning her mother's continued resistance (see Kaukomaa et al., 2014 on frowns in social interaction). Then the two participants hold a rather long gaze at one another in silence (line 14), as Roshana continues her frown.

The imperative in line 13 is, therefore, issued with simultaneous re-configuration of verbal, prosodic, and embodied features that convey an 'intensification' of a quality, quantity, or action. This simultaneous 'intensification' of various resources from different modalities is referred to as *multimodal upgrading*. I claim that multimodal upgrading of a pursuit turn projects more sequence expansions. Performing overt disaffiliative actions in these sequence expansions steers the trajectory of interaction towards escalation. In the rest of this chapter, I provide evidence for this claim, but before that a note on the selection of extracts for this chapter is worthy of a mention: As elaborated in Chapter 1, in full repeats, the verbal modality of the imperative turns is kept constant. Therefore, when multimodal upgrading occurs in full repeats, it only involves prosodic and embodied design features. Full repetitions are great sites to explore the prosodic variations of the subsequent versions compared to a prior since the turns are naturally controlled for their lexical structure (see also Curl, 2005). Therefore, although lexical upgrades also occur (as Extract 3.1 above shows), to better compare the prosodic features of the imperative turns, this chapter uses the 'full repeat' sub-collection.

In what follows, first a review of the relevant literature on up- and downgrading is presented (§3.2). In §3.3, I present different resources for upgrading a turn's design. First, two resource types that are already well-established in the CA literature – verbal and prosodic – are demonstrated with Persian examples, and then embodied upgrading, which is so far less discussed in the literature, is introduced. Then in §3.4, the phenomenon of interest – multimodal upgrading – is defined. The main analytic discussion comes in §3.5, where I elaborate the relationship between multimodal upgrading and escalation. Finally, §3.6 concludes the chapter by summarizing the findings.

3.2. Gradation in CA

In this section, up- or downgrades in CA are reviewed. The analysis in this chapter targets upgrades, and downgrades are addressed in Chapter 4, but the review here covers both phenomena since in most papers they are studied together. In interaction research, the terms upgrading and downgrading have been increasingly used as analytic concepts to characterize a participant's action, stance, or orientation, often as manifested in the design features of the speaker's turn, compared to a point of reference. In broadest terms, upgrade connotates 'more of something' and the reverse is true for downgrades.

The concept of grading – up or downgrading – in CA was popularized by Pomerantz (1984b) through her study on assessment sequences. Her paper examined a variety of ways that first assessments are responded to by co-interactants. Focusing on the turns' lexical design features, she showed three general ways: upgraded, downgraded, and same second assessments. Upgraded second assessments are the ones with a 'stronger' lexical component or an intensifier such as 'very' compared to a first assessment. She argued that such assessments engender strong agreement with the first assessor. On the contrary, downgrades are 'weaker' second assessments, again displayed in the lexical design of the turn, and are implicative of disagreement. The 'same' assessment "asserts the same evaluation as the prior speaker's evaluation" and they can do agreement or preface a disagreement (p. 66). The following show three examples on upgrading, downgrading, and same second assessment taken from Pomerantz's work (highlights added).

Extract 3.2. [SJ:II:28] (Pomerantz, 1984b, p. 65)

```
1    J:    T's- tsuh beautiful day out isn't it?
2    L:    Yeh it's just gorgeous
```

Extract 3.3. [GJ:1] (Pomerantz, 1984b, p. 68)

```
1    A:    She's a fox!
2    L:    Yeh, she's a pretty girl.
```

Extract 3.4. [J & J] (Pomerantz, 1984b, p. 67)

```
1    A:    Yeah I like it [( )
2    B:                   [I like it too but uhh hahheh it blows my mind.
```

As Pomerantz (1984b) explains, in the first example, 'gorgeous' is a semantically stronger adjective than 'beautiful', so it contributes to an upgraded design, doing agreement. While in the second example 'pretty' is a semantically 'weaker' adjective relative to 'fox', so it produces a downgraded second assessment, doing disagreement. And finally, in the third example, no re-scaling has occurred; the evaluative verb is lexically repeated. It engenders a 'same' second assessment and in this example, it prefaces disagreement.

Other studies then developed the concept of gradation based on the verbal features of a turn in various interactional contexts. Again, focusing on assessment sequences, Heritage and Raymond (2005) discuss various other ways by which first or second assessments are designed as up- or downgraded. To mention some of these verbal practices, evidential verbs and tag questions in first assessments are argued to downgrade the assessor's epistemic claims (e.g., "They're a luvly family now ar'n't they." p. 20). On the other hand, interrogatively formatted assessments upgrade the speaker's epistemic rights (e.g., "Isn't she a cute little thi:ng?" p. 21). In the context of repairs in tellings, increasing the precision of a person reference is shown to upgrade the credibility of the telling one (e.g., replacing the lexicon 'people' with a more specific term such as 'neighbors') (Lerner, et al., 2012). In getting someone to do something, telling a recipient what to do is considered an upgrade in the speaker's entitlement compared to asking them (Craven & Potter, 2010; Aronsson & Cekaite, 2011). Or more recently, Hoey et al. (2020) study the expletive *fuck* as a lexical resource for upgrading the turn-constructional design of a pursuit turn. They show that expletive insertion while pursuing an adequate response both sanctions the co-interactant's inadequate response and contributes to escalation of discord. In this way, they distinguish between upgrading the design features of a turn from the action that it achieves – escalation. The same distinction is also observed in this study.

A second influential line of research on gradation in social interaction is led by the phoneticians of talk in interaction. Pioneering work in this regard (e.g., Curl, 2005; Ogden, 2006, 2010) showed some

implications of up- or downgrading clusters of phonetic exponents compared to a prior. Ogden (2006) examines assessment sequences and lists the prosodic exponents by which upgraded assessments are achieved: "more dynamic pitch contours on accented items; pitch higher in the speaker's range than the first assessment; slower tempo; closer, tenser articulations" (p. 1762). And features that embody a prosodic downgrade are listed as "a narrower pitch span, and a concomitant lack of dynamic pitch movement; faster tempo; articulations with more open stricture; a decrease in loudness" (p. 1769). Similar to Pomerantz (1984b), he shows that upgraded second assessments do agreeing while downgraded and same assessments project a potential disagreement. Investigating repetitions in other-initiated repairs, Curl (2005) distinguishes two types of trouble resource based on the phonetic design of the repairs. Upgraded repairs indicate a trouble source that is 'fitted' "within the structure set up by the prior turn or sequence" whereas nonupgraded repairs are associated with trouble sources that are 'disjunct' from a prior turn (p. 8). CA has ever since witnessed more work on the phonetic gradation in various sequential environments (e.g., Couper-Kuhlen, 2012; Local, Auer, & Drew, 2010; Sikveland, 2019; and Zellers & Ogden, 2013; see Plug, 2014, for a critical review).

A more recent direction is taken by those scholars who have considered the embodied aspect of interaction in terms of graded design features (e.g., Keel, 2015; Hauser, 2019). In the context of self-repair, Hauser (2019) studies 'self-repeated gestures': gestures that are recognizable as a prior version performed by the same participant. He categorizes the subsequent gestures into attenuated and upgraded ones. Upgraded gestures – gestures with larger, more distinct, and/or more energetic movements – are produced when there has been some sort of problem with the initial gesture (p. 186). For example, for a pointing gesture, a problem may be that an initial version fails to show the recipient the target of the point. An attenuated gesture, on the other hand, does not deal with a problem in the initial gesture; rather it only 'comes along for the ride' when its lexical affiliate is repeated" (p. 184). Hauser does not call these downgraded gestures since they are not designed to fix a design issue in the prior gesture. He does not find any downgraded gestures in his data and concludes that "there may therefore be limits to the applicability of the concept of regrading to gesture, in that downgrading seems to be uncommon and that upgrading seems to be limited to dealing with problems" (p. 195).

This thesis develops this line of research by showing that up- or downgrading the design features of a turn can involve multiple modalities at once, in which case multimodal gradation occurs. We will focus only on upgrading. First various resources for upgrading a turn are exemplified through Persian data, and then multimodal upgrading is introduced in greater depth.

3.3. Resources for Upgrading

This section presents some examples for three types of resources that participants use to produce upgrades. The first two types are lexical and prosodic which are already established in the CA literature; nevertheless, two short examples are presented to examine these in the Persian imperative collection (§3.3.1-§3.3.2). The third resource for upgrading, which the CA literature has barely touched on is bodily design features (see Hauser, 2019; Keel, 2015), which §3.3.3 focuses on.

3.3.1. Lexical Upgrading

There are different ways by which an imperative turn can be lexically upgraded, for example, inserting adverbs or affective particles, replacing a component of the TCU with a semantically 'stronger' counterpart, repeating a component of the TCU within the same TCU, and so forth. Here are some simplified examples:

- Inserting adverb of urgency or affective particles

Example 1.   **be-frest**   **alan be-frest**
              IMP-send.2SG   now IMP-send.2SG
              send it.   *now send it.*

Example 2.   **be-de**   **be-de**   **dige**
                 IMP-give.2SG   IMP-give.2SG PRT
                 give.   *give* dige.

- Replacing a TCU components with a 'stronger' counterpart

Example 3.   **alân bi-â**   **fori**   **bi-â**
                 now IMP-come.2SG   urgently IMP-come.2SG
                 now come.   *urgently come.*

- Duplicating a component of the TCU

Example 4.   **bo-ro**   **bo-ro bo-ro**
                 IMP-go.2SG   IMP-go.2SG
                 go.   *go go.*

In the first extract of the chapter (Extract 3.1), we saw an example of lexical upgrade by inserting the particles *khob* and *dige.* The following extract shows upgrading by double-imperatives. Farnaz is addressing her husband, Amir, with imperatives *boro* 'go' (line 1). Amir who is expected to produce a response, displaying his understanding/commitment or at least its lack thereof, keeps silent as he continues having nuts and does not even turn his gaze towards Farnaz. All this is implicative of resistance (line 2). Farnaz then repeats the very same imperative albeit in a double format (line 3) (cf. 'doubles' in Curl, Local, & Walker, 2006).

Extract 3.2. RAY011a, 12.40

```
->    01    FAR    bo-ro.
                   IMP-go.2SG
                   go.
             ami   >>eats nuts>>

      02           (2.9)

->    03    FAR    bo-ro       bo-ro.
                   IMP-go.2SG  IMP-go.2SG
                   go go.
```

The single imperative in line 1 is lexically upgraded to a double format in the subsequent version (line 3). There is prosodic evidence for the two imperative verbs in line 3 forming a single TCU rather than two separately designed subsequent versions: (a) they are produced as a single intonational phrase (Vogel & Nespor, 2007) or breath group (Liberman, 1967) and (b) the pitch declines (Cohen & 't Hart, 1967) throughout the production of the line 3 which means there is no pitch reset at the beginning of the second *boro*. By upgrading the lexical design of her imperative turn after Amir's resistance-implicative nonresponse, Farnaz pursues a preferred response.

3.3.2. Prosodic Upgrading

Upgrading can also be performed through prosodic means. In the following extract, the register, the intensity, and the pitch span with which the subsequent imperative is produced are increased. The subsequent version is also produced with a tighter articulation setting. Taraneh has put a glass jar of oil in the microwave to warm up the oil. Anticipating a trouble, Mahsan advises her, through an imperative structure, to change the oil dish before warming it up. When the imperative fails to secure compliance, Mahsan repeats the same imperative turn with prosodically upgraded features.

```
       01    MAH    dafe-ye pish  hamchi    kard-in,     shekas.
                    time-EZ last  like.this do.PST-2PL  break.PST.3SG
                    you did the same last time, it broke.

       02           (0.3)

->     03    MAH    tu ye  zarf-e   dige    garm kon-in.
                    in one dish-EZ  another warm do.IMP-2PL
                    warm it up in another dish.

       04           (1.0)+(0.1)
             tar        +adjust the dish in the microwave-->

       05    TAR    ye::[k\
                    one
                    one

->     06    MAH        [<<f,h>tu ye  zarf-e   diga    garm k+o[n-in.>
                            in one dish-EZ  another warm do.IMP-2PL
                        warm it up in another dish.

       07    TAR                                         +  [yek +daraje
                                                           one  degree
                                                           one degree
             tar                                       -->+clse mcrvwe door+
```

Mahsan anticipates a trouble (see Kendrick & Drew, 2016 on anticipation and pre-emption of troubles), based on a past experience (line 1, 'the dish broke last time' which implies it will break this time as well). This anticipation provides Taraneh an opportunity to pre-empt the trouble by changing the oil dish before warming it up. When this opportunity is not taken by Taraneh (line 2), Mahsan produces the first imperative (line 3) to explicate a solution to the anticipated trouble, advising Taraneh to change the dish (see Kendrick, 2021 on methods of assistance). If Taraneh were to follow the advice, her next action would have been to take the jar out of the microwave; however, she adjusts the dish in the microwave (line 4). Even though her verbal turn is cut off (line 5), her embodied action itself shows that she is following her own course of action. Thus, Mahsan's imperative, too, fails to achieve the 'desired' outcome (line 4-5). This is when Mahsan produces the same imperative in a prosodically upgraded design, pursuing a preferred response (line 6), which is again rejected (line 7).

Relative to the first imperative (line 3), the second (line 6) is prosodically upgraded (Figure 3.7, Table 3.1). There is an increase in the pitch span of the TCUs from 7 ST in the first imperative turn compared to 12 ST in the second. The first starts around the speaker's overall average F_0 at 196 Hz (only 1.3 ST above the speaker's overall F_0 level; the speaker's average F_0 is calculated 181 Hz), but the subsequent imperative starts at 260 Hz (around 6 ST above her average F_0), rising to 300 Hz on the nuclear pitch accent of the intonational phrase (*ye* 'one'), which is around 9 ST above the speaker's average F_0. In the first imperative, the F_0 level drops to as low as 136 Hz at the end of the intonational phrase, but in the subsequent version (excluding *'konin'*, which has been produced in overlap with some non-speech noise), it drops to 147Hz. Therefore, both the top and bottom of the utterance's pitch levels increase in the second imperative turn. Overall, the average F_0 of the TCU raises from 167Hz in the first imperative to 221Hz in the second. In other words, the whole TCU is produced at a higher register (Table 3.1).
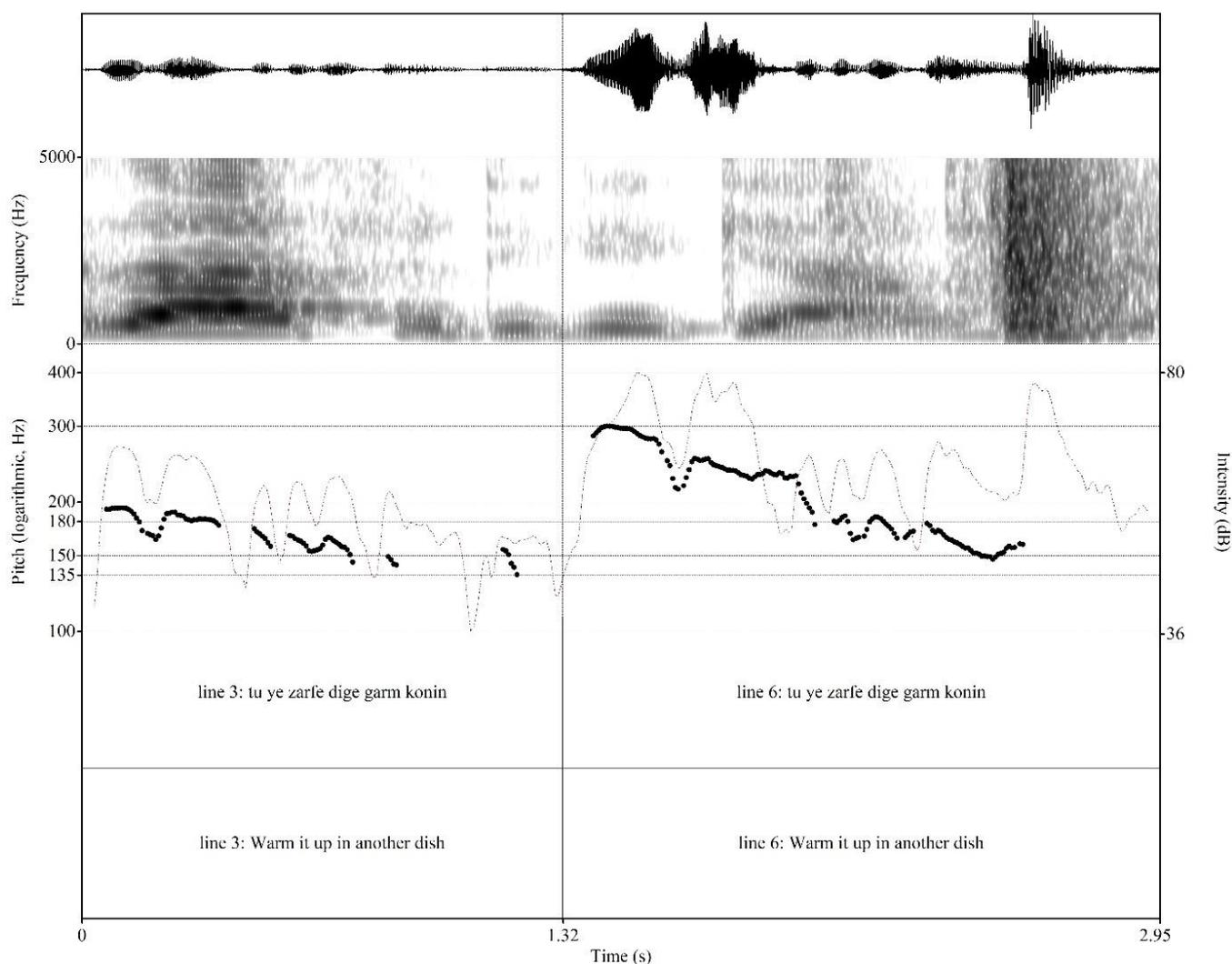
Figure 3.7. Prosodic upgrade in Extract 3.3. In the subsequent imperative turn (line 6) is prosodically upgraded compared to the initial version (line 3). In the left panel, the word *garm* 'warm' is produced creaky, and in the right panel, the word *konin* 'do' overlaps with the noise of the microwave door getting shut (hence no f0 trace for these words).

Table 3.1. Summary of the prosodic features of the imperatives in Extract 3.3.

Line	Average F_0 of the TCU (distance from speaker's overall F_0)	Span (ST)	Intensity (dB)	Articulation setting	Duration (ms)
3	167 (1.4 *below* the speaker's overall f0 level)	7	52	Lax articulation	132
6	221 (3.5 *above* the speaker's overall f0 level)	12	62.5	Tighter articulation	163

The average intensity also increases in the second production. The first imperative turn is only 52 dB whereas the second has an average intensity of 62.5 dB. The subsequent version is longer than the first (163 vs. 132 ms) which means that it is produced at a slower pace. All this goes along with tighter articulatory setting: in the first saying, the vowels have centralized features whereas in the second they do not; therefore, the first is heard more laxed (Table 3.1).

3.3.3. Embodied Upgrading

Now, I extend the concept of upgrading to the embodied domain.[19] First an example of embodied upgrading is presented and then some notes on the challenges around the gradation of bodily actions are discussed. The following extract shows embodied upgrading in service of requesting an object. Minuvash (left in Figure 3.8) is requesting a mobile phone from her sister, Nasima (right in Figure 3.8). Prior to the extract, Minuvash borrows her mother's cell phone to play with. Nasima asks for the cell phone to see some photos. In the extract, Minuvash produces three embodied requests to get the phone back, and she finally resorts to verbal means to do requesting. Embodied upgrading is achieved by performing more prominent bodily movements compared to a similar point of reference. This extract is especially selected since embodied upgrades are performed with no accompanying verbal components. It can thus show that embodied upgrading can be performed as an independent practice (with no accompanying verbal or prosodic means) to pursue a preferred response. Melisa (the two participants' mother) is in the kitchen, pouring tea for Nasima.

Extract 3.4. RAY014c, 07.19

```
01        (10.6)+(0.4)* (0.5) #+  (0.4)   #+
   min          +taps NAS's arm+strtch arm+
   nas              *gaze at MIN-->
   fig                   #3.8        #3.9
```


#3.8

#3.9

```
02   MEL   +mâmân       #+ nabât       be-ndâ#+zam?
           mom             sugar.candy SBJV-pour.1SG
           sweetie should I add sugar candy {to your tea}?
-> min   +bend fingers+lowers arm--------#+
   fig             #3.10,3.10'           #3.11
```


#3.10

#3.10' zoomed in

#3.11

```
03        +  (0.5)       #+              (0.6)        #+
-> min   +strtch arm frthr+gripping gesture as moves arm backwrd+
   fig              #3.12                          #3.13
```

---

[19] I acknowledge the fact that prosody, from a broader perspective, is an embodied phenomenon since this is our body (vocal tract) which produces prosody, but for practical purposes in this chapter I use the term 'embodiment', 'bodily' and 'physical' to refer to body movements or configurations (gestural or instrumental) which are not necessarily needed to produce vocal sounds.
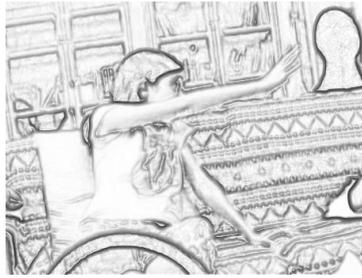
#3.12
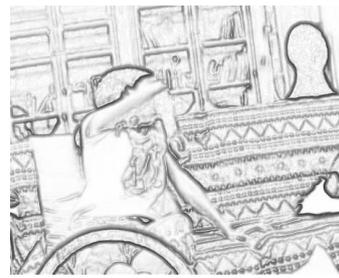
#3.13

```
04   NAS    +âre.
              yeah
             yeah.
     min    +hold-->

05          (0.6)+
     min     -->+

06   NAS    +<<p>be-gu>.
              IMP-say.2SG
              say {it}.
     min    +strtch arm frthr-->

07          (1.0)+#(0.2)#*     #
->   min     -->+gripping gesture and moving arm back-->
     nas             -->*
     fig        #3.14 #3.15 #3.16
```


#3.14

#3.15

#3.16

```
08          (0.6)+(0.8)+(1.6)+  (0.9)   +   (0.2)   #+  (0.9)
     min     -->+,,,,,+      +strtch arm+fingers click+
     fig                                    #3.17
```


#3.17

```
09   MIN    be-de         nasimâ
            IMP-give.2SG Name
            give it Nasima
```

Each embodied request is composed of three main phases: all or some of which get upgraded over the two embodied subsequent versions: (a) stretching arm toward the cell phone, (b) producing a 'gripping' gesture, and (c) retracting fingers/arm toward herself (see Figure 3.18 for the arm stretching phase).

After tapping Nasima on the arm and securing her gaze (line 1, Figure 3.8), Minuvash makes the first embodied request: she raises her arm to the level of phone in Nasima's hand and simultaneously stretches it towards the phone with her fingers open (Figure 3.9). Then she slightly bends her fingers towards herself, showing the direction that the phone should be passed (Figure 3.4), and retracts her hand back to the rest position (Figures 3.10 -3.10′ and 3.11). Thus this first version consists of arm-stretch and finger-retraction. As Kendrick (2020) notes, [t]he shape and orientation of a manual action can not only index an object but also specify a relevant next action" (p. 107). In this case, the two successive phases of arm-stretch and finger retraction constitute one single "gesticular unit" (Kendon, 1980, p. 212) and make relevant the transfer of the phone, which Nasima does not perform.

| Request 1 Line 2 | Request 2 Line 3 | Request 3 Lines 6-7 |
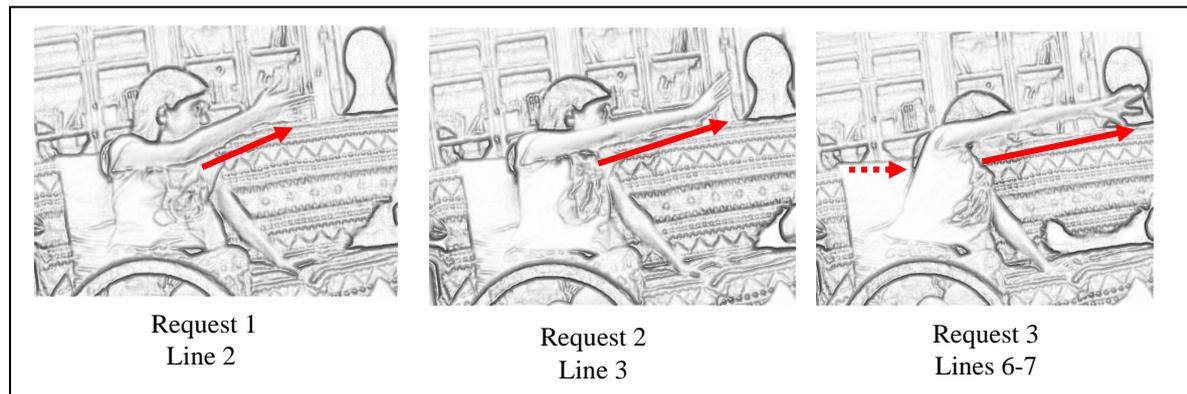
Figure 3.18. Embodied Upgrade in Extract 3.4. The arm-stretch phase of the embodied request. The participant stretches her arm successively further away from the rest position and towards the requested object.

Minuvash then produces a subsequent version: she raises her hand again and this time stretches her arm closer towards the phone/Nasima (line 3, Figure 3.12); successively, she embodies a gripping gesture – opening fingers and then closing them making a loose fist – (line 3, Figure 3.13). Then she bends her whole arm towards herself (Line 3, Figure 3.13). Again, these three phases are performed successively one after another in a single gesticular unit. Compared to the previous version, the upgrade is achieved in three aspects: arm movement further away from the rest position, inserting a gripping gesture, and bending the whole arm rather than just fingers. When there is again no response from Nasima (line 4-5), she further upgrades her embodiment for yet another time.

Overlapping with Nasima's turn in which she indicates that Minuvash should produce a verbal turn (line 6, 'Say {it'}), Minuvash makes a third embodied request again in upgraded design relative to the second version.[20] she stretches her arm even further away from the rest position (line 6, Figure 3.14), then she recycles the gripping gesture without producing any verbal turns (line 7, Figure 3.15). This is when Nasima averts her gaze away from Minuvash and hence stops doing recipiency (Figure 3.16). At this point, Minuvash stops the retraction phase. Instead, she raises and stretches her arm toward Nasima again and clicks her fingers, an embodied summonsing designed to secure Nasima's recipiency again (line 8, Figure 3.17). However, this fails to secure Nasima's recipiency. Finally, Minuvash produces the request verbally (line 9, 'give {it} Nasima').

In short, following each failure to achieve the pursued outcome, the performer upgrades her embodiment by stretching her hand further away from the home position, adding gripping gestures, and bending the arm instead of fingers to indicate the direction towards which the phone should be

---

[20] Nasima's turn in line 6 is produced with a lower volume compared to her previous turn (line 4) when she responded to her mother's question. This prosodic design makes it clear that her *begu* 'say it' in line 6 is addressed at Minuvash rather than her mother in the kitchen.

transferred. This is an example of what I refer to as embodied upgrading (see Hoey et al., 2020, pp. 12-13 for a similar example).

In broadest terms then in this project, embodied upgrades are bodily movements that involve more prominent bodily movements relative to a point of reference, for example, a more extended arm-stretch and the like. This conceptualization conveys that, like other scalar phenomena, embodied gradation is a relational concept. Given the diversity of bodily articulators, the more prominent conceptualization is by no means inclusive. For example, in this thesis 'gaze towards' or 'move towards' a recipient is recurrently accompanied with other prominent embodied and prosodic features, and hence they appear to belong to the same kind of 'upgrade' package. Nevertheless, gaze or movement towards an object/person does not fit the more prominent conceptualization. Also, there are situations in which, for example, 'movement away' occurs with other upgraded features. Consider an example from the corpus in which an interactant requests their pet cat that the co-interactant is holding. In the subsequent version, the requester moves her body away from the object of request – the cat – towards the home position while she directs her gaze towards the recipient and the cat. At the first glance, the body retraction may be seen as backing down (see Chapter 4 on embodied downgrades), until it becomes clear that the body retraction was in the service of achieving a better receptive position to receive the cat. The requester positions herself such that the cat can be put on her lap. The ecology of the interaction such as the spatial configuration between the two participants and the object of request (a cat – a living being – rather than a phone for example) all contribute to the design of the embodied actions. From this perspective, the 'movement away' in this context can be seen as an upgrade since it is in the service of achieving a physical situation in which the pursuit of the interactional outcome can occur more straightforwardly. From another perspective, however, conceptualizing bodily upgrades (or downgrades) in terms of their interactional outcome is in fact conflating design and action, which is from an analytical and methodological point of view problematic.

Therefore, there is an inherent complexity regarding an abstract generalized conceptualization of bodily upgrades. The complexity stems from the fact that analyzing an embodied action as an upgrade compared to a counterpart heavily relies on (a) considerations of what the bodily movements are designed to achieve at a particular moment in interaction, (b) other bodily articulators that are simultaneously operative, and (c) the ecology of interaction (see Mondada, 2018). Given all this, the list below only reflects some recurrent ways in which an embodied action is observed to co-occur with other upgraded prosodic and lexical features.

Direction of movement: extending body parts towards coparticipants or a relevant object or place. For example, an imperative-speaker may extend their stretched arm towards a requested object co-occurrent with a subsequent request for the object.

Contact or involved body parts: increasing the contact size or making more body parts involved in an action. For example, in chapter 1 (Extract 1.1) we saw that shielding the phone from being grabbed was done first through only the palm, and then through the whole forearm. Similarly, in the extract above (Extract. 3.4), the symbolic grip was first done only with fingers, and then the whole arm was used to make a similar action.

Force: the force by which the embodied action is performed can be increased. This can be judged by the visible outcome of the embodied action. For example, when directing a participant to 'come down', a pulling action may be exerted with bigger force in a subsequent version compared to a prior version.

Inserting new relevant bodily action: while producing a subsequent version, a relevant bodily action can be added to a prior version. For example, in chapter 1 (Extract 1.1), we saw that after two attempts of shielding the phone, the third attempt was accompanied with a 'pushing' action where the shielding party both shielded the phone from being grabbed and pushed the approaching party back.

In the rest of the chapter, some of these features are exemplified.

3.4.   Multimodal Upgrading

In producing a turn at talk, different design features involving multiple modalities (verbal, prosodic, and embodied) can sometimes undergo upgrading simultaneously in a temporally coordinated fashion, in which case, I argue, *multimodal upgrading* occurs (see e.g., Loehr, 2012 on body-prosody coordination). Multimodal upgrading, as conceptualized in this thesis, involves a coordinated integration of a turn's design features in different modalities which yields a unified whole. In the following extract, Melina, Safura's ten-year old daughter, is whispering a trouble in Safura's ear (lines 1, 3; Figure 3.13). In overlap, Faraneh, Safura's mother, sitting on the other side of the room, is addressing Safura with the imperative 'open your mouth' (lines 2) so that she tries throwing a pistachio into it from a distance. When Safura, who is visibly engaged in an interaction with her daughter, does not respond, Faraneh repeats her imperative (lines 4-5). Notice the co-occurring prosodic and embodied upgrades in each successive repetition compared to its previous version.

Figure 3.19. Participants in Extract 3.5. From left to right: Faraneh, Safura, Melina. Melina is whispering something in Safura's ear.

Extract 3.5. RAY040a, 08.12

```
    01    MEL    ±+<<whsp>ye zare az    un
                         one bit from that
                         one of one the
          far    ±peels a pistachio-->
          mel    +holds mouth near SAF's ear-->>


                 jarf-â#±[e tuye komod-e=                  ±#
                 dish-PL-of  in   cabinet-of
                 dishes in the cabinet of

->  02    FAR        #±[<<f>dahan-t-o         vâ   kon?>=±#
                         mouth-2SG-OBJ.MAR open do.IMP
                         open your mouth?
          mel
          far        -->±raises hand----------------------±
          fig        #3.20                                #3.21
```
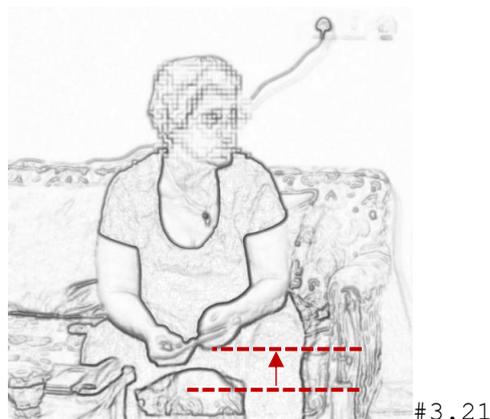
```
    03    MEL    =±[faran    kaj              ±#[shod-e>                        ±#
                    Name      wobbly              become-PRF.3SG
                  Faran has become wobbly a bit


->  04    FAR    =±[<<fh>dahan-t-o vâ kon?>±#                                  ±#
                        mouth-2SG-OBJ.MAR open do.IMP
                       open your mouth?


->  05    FAR    ±                            ±#[<<ff,hh>dahan-t-o vâ[kon?>±#
                                                  mouth-2SG-OBJ.MAR open do.IMP
                                                     open your mouth?


    06    SAF    ±                            ±#                    [chi::? ±#
                                                                      what
                                                                     What?

          far    ±beat hand raising-------±beat hand raising---------±
          fig                             #3.22                      #3.23
```

#3.22

#3.23
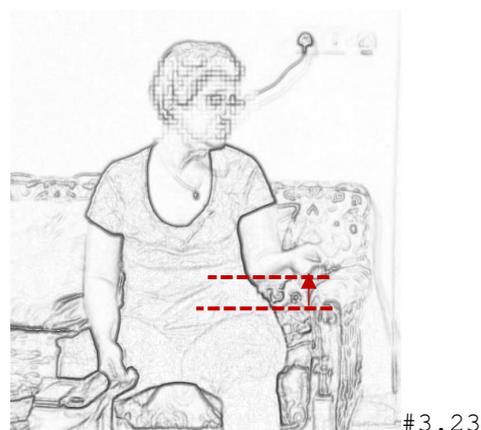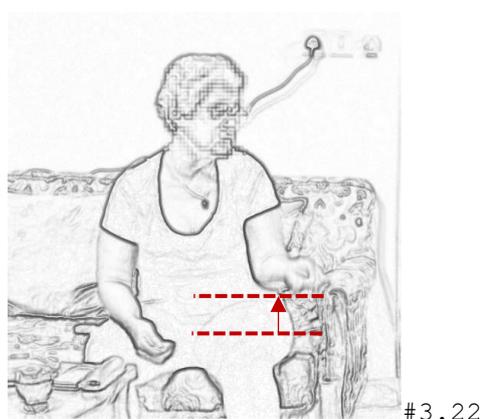
There is a fine coordination between the two modalities of interest – prosodic and embodied – in the imperative turns (lines 2, 4, and 5). This coordination is manifested in two aspects: (a) the congruent increase of the imperative-speaker's hand movement away from the home position and some of the prosodic features of the imperative turns and (b) the temporal alignment of the imperative-speaker's hand movement with the duration of imperative turns.

Let us first examine the congruent increase of the features. Firstly, the overall F_0 level, F_0 at the boundary tone, and the intensity of each successive imperative turn are increased relative to its prior version. Also, the speaker's hand is raised further away from the home position (Figure 3.24). Faraneh raises her hand during the production of the first version. That is, her hand, which is in home position peeling the pistachio just before issuing the first imperative (Figure 3.20), reaches higher in space right after the imperative turn is delivered (Figure 3.21). Prosodically, the TCU is designed rather loud (compared to Faraneh's other stretches of talk in the same recording, 77 dB), which can be due to its overlapping occurrence with a competing CoA in which the imperative-recipient, Safura, is visibly engaged. It starts around 4 ST above Faraneh's average F_0 level (see Table 3.2 for more details).

Table 3.2. Summary of prosodic features in Extract 3.5.

Line	Mean F_0 of the TCU (Difference from speaker's overall F_0)	Min F_0 (Difference from speaker's overall F_0)	Max F_0 (Difference from speaker's overall F_0)	Intensity (dB)	Pitch span (ST)	Duration (ms)
13	271 Hz (4.0 ST)	226 Hz (1.0 ST)	382 Hz (10 ST)	77	9.0	700
15	286 Hz (5.0 ST)	268 Hz (4.0 ST)	414 Hz (12 ST)	77	9.0	800
16	320 Hz (7.0 ST)	275 Hz (4.5 ST)	478 Hz (14 ST)	80	10.0	800
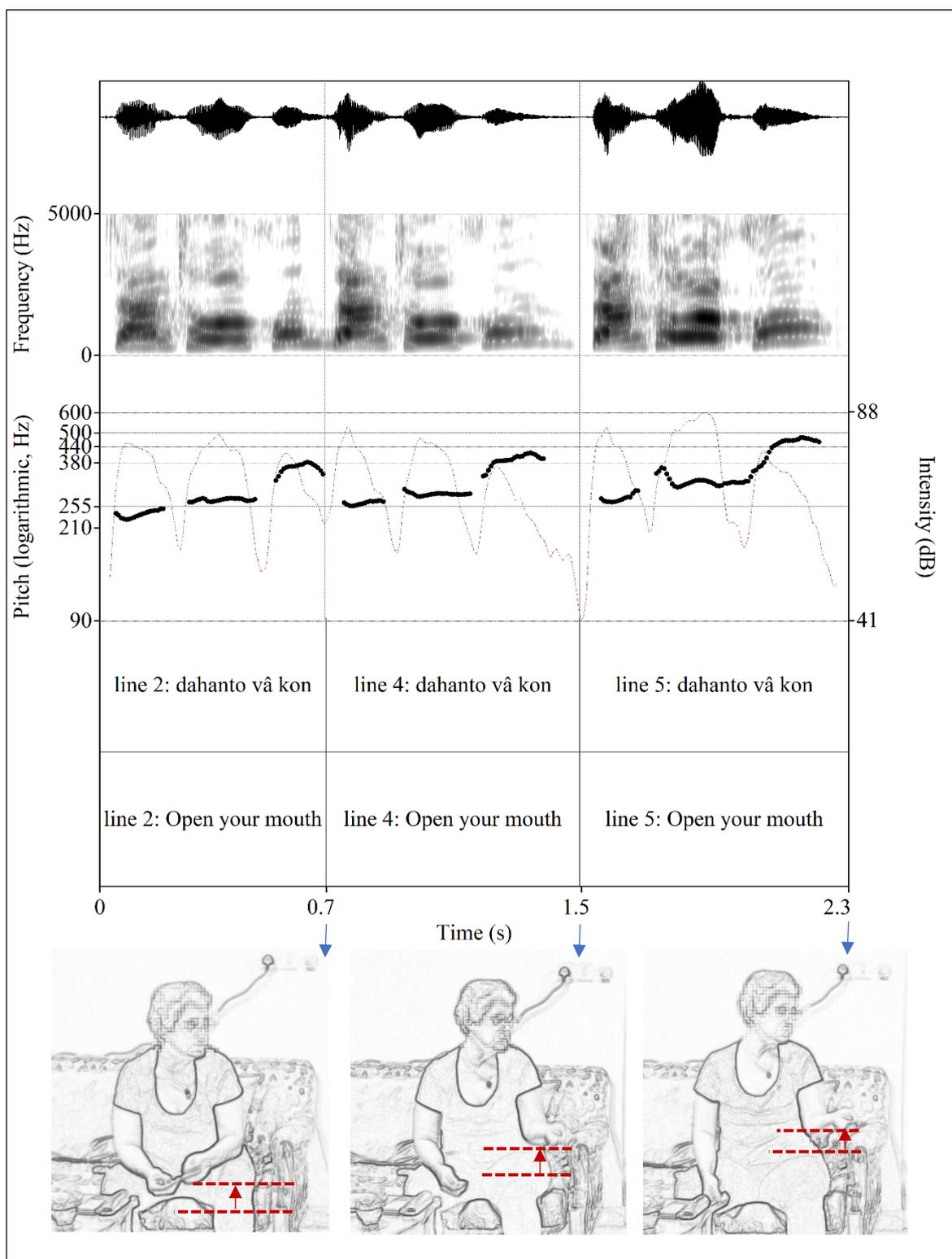
Figure 3.24. Multimodal upgrading in Extract 3.5.

The second imperative turn is produced with multimodal upgraded features (line 4). Faraneh speaks both higher in her register (overall, 5 ST above her average F_0) and reaches higher at the boundary tone (12 ST). The intensity does not change dramatically over the second production. Note that in the prior

imperative, she was already talking rather loudly. Simultaneously, she raises her hand higher with a beat movement. Her hand movement reaches a higher apex compared to the prior version (Figure 3.22).

The same pattern of multimodal upgrading is again observable in the third imperative turn: overall, Faraneh speaks higher in her register (7 ST) and at the boundary tone, too, she reaches higher compared to the prior versions (14 ST). This time, the overall intensity of the TCU also increases (80 dB). Her hand is also raised to a higher apex at the end of her imperative turn (Figure 3.23). Therefore, both prosodic and embodied features are modified with increasing measures.

The second aspect of the fine coordination between the two modalities lies in their temporal alignment. In this example, the temporal alignment itself has two facets: (a) the direction of the bodily actions that co-occurs with each imperative match the imperatives' pitch configuration and (b) the prosodic and embodied behavior are temporally coextensive, creating a 'cohesive gestalt' (Gibbon, 2009, p. 15). Before explaining these two aspects, first Faraneh's embodied conduct and the pitch contour of her imperative turns are described separately. The two modalities are explained separately only for the sake of clarity. This should not imply that the prosodic and embodied design features are treated as two independent resources. On the contrary, the kind of coordination that exists between the two suggests that these two resources in fact act as *one*. Perhaps, treating them as belonging to separate worlds is an artificial classification. They together form one Gestalt.

The bodily action co-occurring with each imperative consists of two stages: in the first stage Faraneh's hand moves down – dip – and in the second stage, her hand moves up – rise. The hand movement which co-occurs with the first imperative does not have the dipping phase; it starts at home position which is already relatively low in space. Through dipping and raising her hand, Faraneh displays a preparation phase of a throwing action. The still images in the transcript show the end of the 'rise' stages. The 'dip' stages are not shown to keep the transcript more readable.

A similar dip-rise pattern is also observable in the pitch contour of each imperative: at the beginning of each imperative turn, the pitch is reset to the speaker's overall F_0 (dip) and then it gradually rises towards a high boundary tone (rise). Again, the first imperative does not have a dip stage since Faraneh was silent before that. The imperatives are designed with a high boundary tone (H%; a rising pitch contour).[21] Although they are produced in succession with either no or only a micropause between them (line 2, 4, no gap and line 4, 5, only a micro-gap), Faraneh resets the pitch to her overall F_0 level at the start of each next turn (Figure 3.24).

Now, the temporal coordination between the prosodic and embodied modalities is observable in the concordant dipping and rising of the pitch and hand movements: when the pitch is reset to the average F_0 level, Faraneh lowers her hand close to home position and she raises her hand throughout the production of the imperative turn until she reaches the apex of her manual movement at the end of the intonational phrase.

The two modalities are also coextensive: Faraneh's hand movement reaches its apex as soon as the imperative TCU comes to a completion at which point she brings the rising movement to a halt. In the first two imperatives, as discussed above, the halt is followed by the dipping movement. In the third imperative, Faraneh just holds her hand raised for some time. Thus, her bodily conduct is temporally in tune with other modalities involved in the production of the imperative.

---

[21] Examples like this are counter evidence to the established argument in Persian intonation literature that Persian imperatives have a L% boundary tone. This collection shows both high and low boundary tones as possible intonational design for imperatives. The interactional contexts where each of these possibilities occurs merit a full-scale investigation.

Therefore, the fine coordination between the prosodic and embodied modalities is evident in two aspects: increasingly prominent design features in various modalities and temporal co-extension of some of the features. Overall, these design features together make the imperative turns sound and look more 'effortful'. This coordination can be observed in most extracts in this collection, and it is indeed this character that suggests these resources are in fact and in the first instance produced as one. To remain focused on the relationship between multimodal upgrading and escalation, in the rest of the extracts, I will not present details of the imperative turns' fine prosodic-embodied coordination.

3.5. Multimodal Upgrading and Escalation

With the phenomenon introduced, I now delve into the main analytic claim: when pursuing a preferred response, upgrading the multimodal design of the pursuit turn projects more sequence expansions. When these sequence expansions are actualized with overt disaffiliative actions, they project a full-blown escalated conflict (see also Hoey et al., 2020). Escalation is a process, a trajectory of interaction towards increasingly more overt disaffiliative actions. By escalation I mean resorting to more overt disaffiliative practices to advance a CoA that is incompatible with the coparticipant's CoA (See Drew & Walker, 2009; Hoey et al, 2020., Yu, et all, 2019). Multimodal upgrading can both shape an escalation trajectory and steer the interaction towards even more overt disaffiliative exchanges. This claim is made based on the observation of the frequent co-occurrence of negative sanctioning practices with multimodal upgrades. Through these practices, the pursuing party holds the recipient accountable for their noncompliance. This in turn can potentially lead the interaction towards more and more disaffiliative exchanges.[22] In this section, some evidence for this claim is presented.

Multimodal upgrading is only observed after the co-interactant's resistance against furthering the initiated CoA. As mentioned in Chapter 2, resistance in this thesis is conceptualized as a dispreferred response or a less-than-preferred response to an imperative turn (see §2.5-2.6). The following examples show multimodal upgrading after noncompliance (§3.5.1) and after pro-forma compliance (§3.5.2). To develop my overall argument the analysis of each extract shows three main points: (a) the phenomenon of multimodal upgrading in pursuit turns, (b) the sequential position of the phenomenon, and (c) some sanctioning practices that co-occur with the phenomenon.

3.5.1. Multimodal Upgrades after Noncompliance

A frequent sequential position where a subsequent imperative turn is produced with upgraded multimodal features is post-noncompliance: when the imperative-recipient does not complete the imperative-speaker's initiated course of action either through rejecting (explicitly or implicitly) or simply not performing the named action. In this collection, upgrading multimodal features of a subsequent imperatives after noncompliance is frequently accompanied with practices or actions which sanction the dispreferred response. Escalation is conveyed in two factors: firstly, re-doing action that is already resisted and secondly, using disaffiliative practices to hold the recipient accountable for resistance.

In the following extract, an embodied and a vocal practice (a sanctioning look and sighing) are used by the imperative-speaker to sanction the recipient's noncompliance as she pursues a preferred response. The extract shows that multimodal upgrades can be as subtle as a brow-raise coupled with upgraded prosodic features. Prior to the extract, Mahshid, Hanna's mother asks Hanna to take photos of a few documents. While Hanna scans the documents, she and Mahshid get engaged in a complaint sequence where Hanna complains about how Mahshid and other family members take poor photos whenever she (Hanna) needs digital copies of documents (lines 1-2 are the last lines of the complaint sequence). The first imperative *beshin* 'sit down' (lines 4) is issued by Hanna when Mahshid keeps standing next to

---

[22] In this collection, in sequences with multimodal upgrades where there is no orientation to accountability, orientations to urgency or what I term affective pleading is observable. Urgent and affective pursuits are not discussed in this thesis, but it is noteworthy that multimodal upgrades in these two types of pursuit do not necessarily project escalation.

her, projecting taking away the documents potentially before Hanna's scanning is complete. Note (a) how the multimodal design of the subsequent imperative (line 7) is upgraded after Mahshid's embodied and verbal resistance to complying and (b) how Hanna casts a sanctioning look and produces a sigh after Mahshid's further noncompliance (two sanctioning practices).

Extract 3.6. RAY052d, 13.57 [Left: Hanna; Right: Mashshid]

```
01    HAN    #shomâ aks    gereftan balad
              2.PL  photo taking    knower
      mah    >>stands next to HAN-->>

             ni-s[t-in hichkodum.±
             NEG-be.2PL
             none of you know how to take a photo

02    MAH       [balad  ni-st-im #±    bale bale.
                 knower NEG-be-1PL yes   yes
                 we don't know how yes yes.
      mah                          ±cut vegetable-->>
      fig                 #3.25
```


#3.25

```
03            (0.3)

->   04    HAN    *+be-shin+#
                   IMP-sit.2SG
                   sit down
                  *gaze down-->
                  +........+
             fig             #3.26
```


#3.26

```
05            +(0.6)
      han    +holds cell phone raised above the paper-->

06    MAH    be-zâr  ejâze    be-de      ur-e
             IMP-let permission IMP-give.2SG that-OBJ.MAR
```

```
                    jam        kon-am *[be-zâr-am s.\        *#
                    collection do.1SG  SBJV-put.1SG
                    let me allow me to collect that and put it\
```

```
->    07    HAN                        *[be-shin.            *#
                                        IMP-sit.2SG
                                        sit down.
            han                    -->*gz from corner of eye*
            fig                                          #3.27
```


#3.27

```
      08            (0.2)
```

```
      09    MAH   mi-khâ-m      be-+zâr-am   *sar-e+ jâ-sh.
                  PRG-want-1SG SBJV-put-1SG on       place-3SG
                  I want to put it away.
            han                  -->+taps-----------+lowers phone-->
            han                            *gz from the corner of eye-->
```

```
      10            (0.2)*+#
            han     -->*
            han     -->+
            fig          #3.28
```


#3.28

```
      11    HAN   +h.[:::]           + ((outbreath of 300 ms))
```

```
      12    MAH   + [gom]           +  mi-sh-e        bad;
                    lost               PRG-become-3SG then
                  {if not,} it'll be lost later;
            han   +put device on tble+
```

```
      13            (0.1)
```

```
14    HAN    vare-sh  dâr.
             prep-3SG IMP.have.2SG
             take it.

15    HAN    tamum    shod.
             finished become.PST.3SG
             it's over.
```

The first imperative is produced to pre-empt the imperative-recipient's possible disruption in the imperative speaker's ongoing action. After the complaint sequence (lines 1-2), the imperative, *beshin* 'sit down' (line 4), is issued when Mahshid, the recipient, who had previously been sitting and cutting some vegetables, resumes the cutting activity while standing next to Hana (line 2, Figure 3.25). Even though Mahshid is re-engaged in her previous activity (cutting vegetables), the fact that she does not go back to her previous 'stable' physical positioning (sitting on her own chair) projects her collecting the documents, laid in front of Hanna, as an imminent action. On the other hand, Hanna, the imperative-speaker, who has finished scanning one document, is not yet holding her mobile phone in a scanning position for the second document and, therefore, her scanning activity can potentially be seen as complete (Figure 3.25). This is when Hanna issues the first imperative 'sit down' (line 4). The imperative is designed to create more physical distance between Hanna and Mahshid, which in turn pushes back the potential imminence of Mahshid's collecting the documents.

The recipient shows resistance against complying: she does not complete the speaker's initiated CoA. Her resistance has two components: not only does she refuse to sit down (line 5), but she also requests permission to collect the documents (line 6), the very action that Hanna has been pre-empting. By so doing, she explicitly verbalizes that her next action is indeed collecting the documents. In overlap with Mahshid's turn, Hanna re-issues the imperative, pursuing a preferred response (line 7). Therefore, the subsequent version is produced post-noncompliance.

The subsequent imperative is designed multimodally upgraded relative to the first (Figure 3.29; Table 3.3). Prosodically, the first imperative (line 4) is produced around the speaker's average F_0 (overall only 1 ST above her average F_0 level), and quiet compared to the speaker's other stretches of talk in the same recording (only 58 dB). It is also designed with a narrow pitch span (only 3 ST wide) and short (only 300 ms long). The speaker's facial muscles look lax (Figure 3.26) and there is no hyper-articulation in the production of any of the sounds. Both vowels are produced with centralized qualities (i.e., they are not peripheral). The subsequent version, on the other hand, is produced higher in the speaker's pitch range (around 5 ST above the speaker's average F_0 level), loud (80 dB), wide-pitched (11 ST wide), and longer relative to the first imperative (640 ms long). The speaker's facial muscles look tense, and there is hyper-articulation especially in the production of the second vowel, [i]: the jaw is visibly close and the corners of the lips are drawn to sides such that the speaker's front teeth are fully visible (Figure 3.27).

Table 3.3. Summary of the multimodal design features of the imperatives in Extract 3.6.

Line	TCU mean f0 (difference from overall mean f0)	Pitch span (ST)	Intensity (dB)	Duration (ms)	Articulation	Embodiment
4	234 Hz (1.0 ST)	3	58	300	Lax	Gaze on the device
7	287 Hz (4.6 ST)	11	80	640	Tense	Gaze at MAH

Her co-occurring embodied conduct also changes: she raises her brows and moves the eyeballs towards left where Mahshid is standing while keeping her head orientation stable toward the document on the table (Figure 3.27, The filtering of the image blurs the fact that Hanna's gaze is directed towards Mahshid in this figure). This looking *out of the corner* of her eyes conveys a 'glaring' action at Mahshid. Of course, the ecological context where this embodiment is performed contributes to its occurrence: firstly,
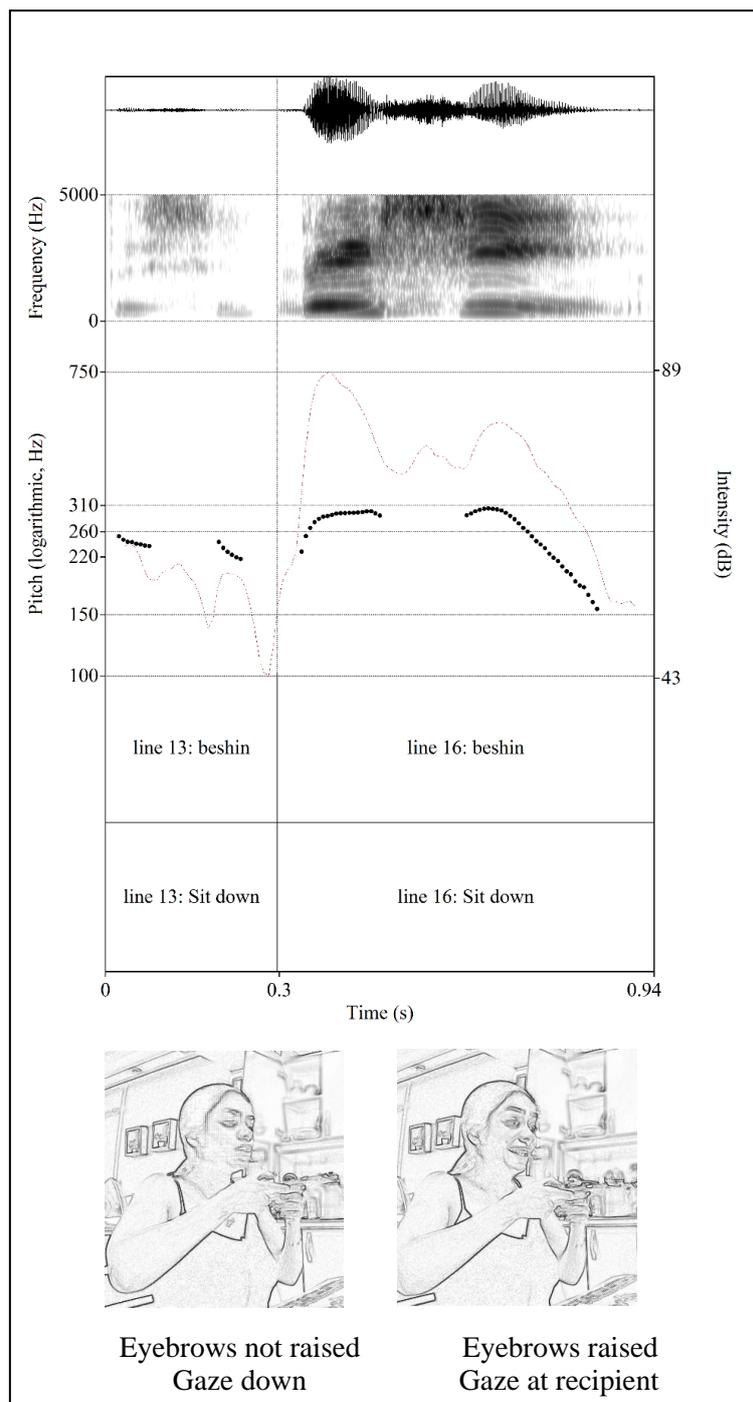
Figure 3.29. Multimodal upgrading in Extract 3.6.

Mahshid, the addressee of the glare, is standing at one side of Hanna (hence Hanna's eyeballs turning to the corner); and secondly, Hanna, who has just adjusted her hands' position over the document, is holding her body still for another round of scanning (Figure 3.27); therefore, moving her body (e.g., turning her whole body or even her head towards Mahshid) can dislocate her hands and disrupt her bodily preparation for the scanning action. In this context, the embodied upgrade is achieved by means of body parts which cause least disturbance to the progression of Hanna's ongoing activity (scanning) and those are the eyes and brows.

Compared to the first imperative, which was not accompanied with any 'extra' embodiment than the ones which were necessary for the scanning action (e.g., just hands above the documents and gaze down at the documents), this addition of a brow flash along with gazing at the recipient from the corner of the

105

eyes during the production of the subsequent version is analyzed as an embodied upgrade. As well as contributing to the upgraded multimodal Gestalt, Hanna's embodiment is the means by which she negatively sanctions Mahshid's dispreferred response.[23] With the combination of prosodic and embodied modification, Hanna achieves a multimodal upgrade in the production of the subsequent imperative.

Following the upgraded subsequent imperative (line 7), Mahshid continues to show noncompliance (line 9): she further refuses to sit and also changes her previous request for permission for collecting the documents to an announcement that she is indeed going to collect the documents (line 9). Thus, she, too, upgrades her noncompliance. Then, Hanna further sanctions Mahshid's continued noncompliance: again, she casts a sanctioning look by gazing at her from the corner of her eyes (Figure 3.28) (Kidwell, 2005) and produces a nasal outbreath of about 300 ms, hearable as sighing (Hoey, 2014). From the emic participant's perspective, Mahshid's own orientation to Hanna's conduct as practices for holding her accountable for her dispreferred responsive actions is clear in the account that she gives for her disprefered response (line 12).

Upgrading the multimodal design of a subsequent imperative turn after the recipient's noncompliance is produced along with practices or actions by which the speaker treats noncompliance as negatively sanctionable. Embodied practices for doing sanctioning (e.g., frowning, sighing, etc.) are sometimes produced concurrent with the delivery of a subsequent version (e.g., the brow flash and the first sanctioning gaze in the extract above, lines 7-8), in which case they are analyzed as a part of the upgraded multimodal Gestalt. Sometimes they are produced before or after a subsequent version. In this latter case, even though the preceding and following sanctioning bodily actions contribute to escalation, they are not analyzed as a part of the upgraded multimodal Gestalt (e.g., the sighing and the second instance of sanctioning gaze, lines 9-11, are not analyzed a part of multimodal upgrading). This is because multimodal upgrading, as defined in this thesis, requires temporally coordinated occurrence of design features in different modalities. Due to their frequent co-occurrence, multimodal upgrading and sanctioning practices together can be analyzed to belong to a 'package' for doing pursuing: offering the resisting party another chance to produce a preferred response while holding them accountable for their dispreferred response. The co-occurrence of these two – sanctioning and pursuing an already resisted response – projects sequence expansion and escalation. That is, it projects more post-expansions with progressively more explicit practices towards achieving a preferred response.

3.5.2. Multimodal upgrades after pro-forma compliance

Resistance can manifest itself in pro-forma acceptance. In agreement sequences, Schegloff (2007) notes, a pro-forma agreement, canonically actualized in "yes, but …" type of responses, functions to delay the actual dispreferred alternative (pp. 69-70; also Sacks & Schegloff, 1987). In the imperative sequences under scrutiny, a pro-forma compliance – a verbal compliance token (such as *bâshe* 'OK') followed by bodily noncompliance – has a similar function. It comes off as compliance at first, but the compliance is only in form, and it is followed by displays of noncompliance. Even though on the verbal level, the recipient commits themselves to further the initiated CoA, in practice they hinder the progressivity of the CoA and, therefore, on the practical level, they are dispreferred.

Similar to noncompliance, subsequent imperative after pro-forma compliance can be designed with upgraded multimodal design features and accompanied with sanctioning practices or actions. In this position, too, pursuit creates another response slot for the recipient to produce a preferred response. Again, the co-occurrence of upgrading and sanctioning while re-doing an action that has been already resisted tilts the sequence towards an escalated trajectory and projects a conflict if both participants continue to follow their incompatible courses of actions.

---

[23] Looking at someone out of the corner of one's eyes, in Persian, can be a conventional way of sanctioning. In fact, the idiomatic phrase *chap chap negah kardan* (gloss: left left look doing) which literally means 'to look at someone from the left corner of eyes' is often used to describe a reproaching stare/glare. Therefore, Hanna's eye movement towards her coparticipant is not just a mere gaze, but a sanctioning look (although Mahshid's physical positioning relative to Hanna contributes to this choice of sanctioning practice).

The following extract is a case in point. Saba (a university student) and Kian (Saba's husband) are discussing Saba's workload. Prior to the extract, Kian listed four tasks that Saba is required to do for her studies. The extract begins with him proposing yet another task (line 1). This is when Saba produces the imperative in focus *velam kon* 'Drop it' (line 2). The imperative both implicitly rejects Kian's proposal and initiates a new sequence which makes relevant for Kian to drop the proposal. The first subsequent version (line 4) is issued after Kian's noncompliance and the second subsequent version (line 10) after a pro-forma compliance. Note how Saba upgrades the multimodal design of her subsequent imperatives and how she negatively sanctions Kian's resistance against complying, using Persian interjection *sh* [ʃ] (lines 6, 8).

Figure 3.30. Participants in Extract 3.7. Saba (left) and Kian (right) are discussing Saba's workload.

Extract 3.7. RAY027d, 12.13

```
01   KIA    tâze  in  lâbelâ   be nazar-am shoro kon yavâsh yavâsh
            fresh this meanwhile to idea-1SG start do little little
            besides in the meantime I think gradually start to

            un maghâlehe tezeto darâr   bef+*[rest. dob\      +
            that paper   thesis extract send        cut.off.word
            write your thesis paper and send it. agai\

->  02   SAB                                  +*[vel-am#  kon.[(?)+
                                                free-1SG IMP.do.2SG
                                                drop it.
         sab                                  *furrows brows-->
         sab                                  +raises both hands--+
         fig                                          #3.31
```


#3.31

```
03   KIA                                              [dobâre
                                                       again
```

```
                          inam      +[mi\     #              +
                          this-too cut.off.word
                          again this one too will\

->   04   SAB              +[vel-[am# kon;       +
                          free-1SG IMP.do.2SG
                           drop it;

     05   KIA              +      [  #((click)) +
          sab             +raises hands higher+
          fig                      #3.32
```


#3.32

```
     06   SAB       •sh:::
                     INTJ
                     sh:::
                    •gaze at KIA-->>

     07             (0.9)

     08   SAB       sh[:::
                    INTJ
                    sh:::

     09   KIA       [bâshe. to+[:\     #              +
                    be.3SG 2SG
                    OK. you:\

->   10   SAB                   +[vel-am#  kon;       +
                               free-1SG IMP.do.2SG
                                drop it;
          sab                   +raises hands higher+
          fig                          #3.33
```


#3.33

```
11              (2.0)*(1.0)
    sab         -->*

12    KIA       khejâlat be-kesh.
                shame    IMP-pull.2SG
                you should be ashamed.
```

Let us first focus on the sequential positions of the imperatives and the vocal sanctioning practices. The first imperative is issued after Kian proposes an extra task for Saba's workload (line 1). This imperative is designed to not only reject Kian's proposal but also to shut down further talk about the proposal (line 2, 'Drop it'). It is produced in terminal overlap with the proposal (lines 1-2). Note the cut-off lexical component *dob* 'agai' in Kian's turn. A preferred response to this imperative would have been for Kian to stop proposing. But in his next turn, he recycles the cut-off verbal component (i.e., *dobâre* 'again') and continues his TCU. This shows that he is following his own CoA (line 3). By furthering his own CoA, Kian refuses to comply with Saba's imperative.

Following Kian's noncompliance, Saba issues the second imperative (line 4) and both participants address one another with vocal sanctioning interjections (lines 5-6, 8). The imperative is re-issued in interjacent overlap and prevents Kian's proposal turn from reaching a completion (line 4). It provides a new slot for a preferred response. In overlap with Saba's imperative, Kian produces a click, which in Persian is a conventional display of disapproval in some contexts (line 5). It can be heard to be sanctioning Saba's multiple overlapping imperatives which prevent Kian from completing his proposal turns. Saba, too, vocalizes a sanctioning interjection – an elongated [ʃ] *sh:::.* (line 6). This interjection is different from the 'shush' interjection. It has also an equivalent affricate version [tʃ:::], and it is another conventional vocal practice for displaying disapproval.[24] Almost a second passes during which Saba has maintained her gaze at Kian (line 7). Then she produces yet another similar interjection (line 8). With vocalizing the interjection *sh::* not just once but twice, Saba 'strongly' counter-sanctions Kian's continued resistance.

The third imperative is issued in the face of Kian's pro-forma compliance (lines 9-10). Kian produces a compliance token (line 9, *bahse* 'OK'), which has preferred implication; however, he then continues his turn with the pronoun *to* 'you' (line 9), referring to Saba. This projects that his next action is again related to Saba, and potentially her study tasks – a topic that Saba has been preventing from developing. Therefore, Kian verbally commits himself to dropping the proposal, but in practice by continuing to talk about Saba, he can be seen to be continuing the topic. This makes the *bahse* 'OK' only a pro-forma compliance token. Saba's own orientation towards Kian's response as not complying is evident in her issuing yet another subsequent version. The imperative is issued in post-onset position and achieves to cut off Kian's turn only after the first lexical component (which nevertheless shows that Kian's next TCU is going to be about Saba). This way Saba does not leave any chance for Kian to do *any* action about her (proposal or otherwise). This displays Saba's upgraded stance towards Kian's resistance as sanctionable conduct such that his action is hindered right after it is initiated. The two participants have mutual gaze at one another in silence for three seconds (line 11). Finally, Kian verbally sanctions Saba's persistence (line 12, *khejalat bekesh*). The verbal sanction, *khejalat bekesh*, is an imperative itself 'be ashamed' translated to 'you should be ashamed' in the transcript. This imperative structure is another conventional verbal practice to sanction a co-interactant for their social misconduct.

Now consider the multimodal design of the imperative turns. Each successive imperative is designed with upgraded multimodal features (Figure 3.34; Table 3.4). In terms of the prosodic design of the turns,

---

[24] As mentioned in Chapter 1, Persian is rich in terms of the variety of affective interjections and particles. The precise difference between a click and a '*sh::*' or its equivalent '*tsh:::*' and other similar sanctioning vocalizations should be the focus of future investigations.

the first imperative starts around the speaker's average F_0 level and falls low in her pitch range. The intensity is calculated 67 dB (not particularly loud relative to the speaker's other stretches of talk in the same recording), and on average, it is around 6.0 ST above the speaker's f0 level. The second imperative is prosodically upgraded. It is issued at a higher register: it starts much higher in the speaker's pitch range compared to the first imperative (at 333 Hz) and reaches a higher maximum (442 Hz). While the previous imperative fell low in the speaker's range (to 170 Hz level), this one falls to a higher level (300 Hz). This reduces the pitch span, which is not, in this extract, among the measures which contributes to the overall prosodic upgrade. The overall intensity is calculated lower than the previous version (61 dB) which can be in part due to the elongation of the 'hold' phase of the plosive [k]. Overall, this version is produced 11 ST above the speaker's F_0 level, which is an average of 5 ST higher than the first version. The third imperative is further upgraded: it starts even higher in the speaker's pitch range (at 388 Hz) compared to both previous versions; it rises to an even higher maximum (555 Hz), and it also falls to a higher level (at 362 Hz). Again, Saba uses a higher register to produce this imperative. Overall, the TCU is produced 15 ST above her average F_0. The intensity is also higher than both prior versions (an average of 71 dB).

Table 3.4. Summary of the multimodal design features of the imperatives in Extract 3.7.

Line	Mean F_0 of Imperatives (distance from overall F_0)	Intensity (dB)	Pitch span (ST)	Duration (ms)	Embodiment
02	297 Hz (6.0 ST)	67	15	900	Throws hands up
04	391 Hz (11 ST)	61	7	900	Throws hands higher
10	498 Hz (15 ST)	71	7	900	Throws hands highest

The co-occurring bodily conduct is also upgraded (Figure 3.34). With each imperative, Saba raises both her hands up, a canonical gesture which in this context connotes 'throwing the topic away' (lines 2, 4, 10). Both her hands are 'thrown' higher in air in each successive imperative, but her hands are not at the same level, so examine the manual upgrade focusing on her left hand. With the first imperative, the hand reaches only just below her chin (Figure 3.31); with the second imperative it reaches her nose level (Figure 3.32), while with the third it almost reaches her eye level (Figure 3.33).[25] There is also a frowning action starting with the first imperative (line 2). The brows are progressively more wrinkled in the production of each successive imperative (this aspect of bodily upgrade is harder to capture in the images due to the filtering effect). The furrows, conventionally associated with a negative stance in some contexts, contribute to the sanctioning action, explained just above. Saba unfurrows her brows in line 11 after Kian's noticeable silence which shows that he has potentially dropped the topic and complied. This is evidence that her frown was indeed oriented to Kian's proposals.

In this extract, upgrading can be seen at the turn-organizational level as well. The first imperative is issued in terminal overlap (lines 1-2), the second one in interjacent overlap (lines 3-4), and the third in post-transition onset overlap (lines 9-10). This indicates Saba's progressively 'decreasing tolerance' to be addressed with anything about her workload or even more generally about her.

---

[25] Her head slightly moves down, but the upgrade is still clear by comparing the position of her hands relative to the bed frame behind her (Figure 3.34).
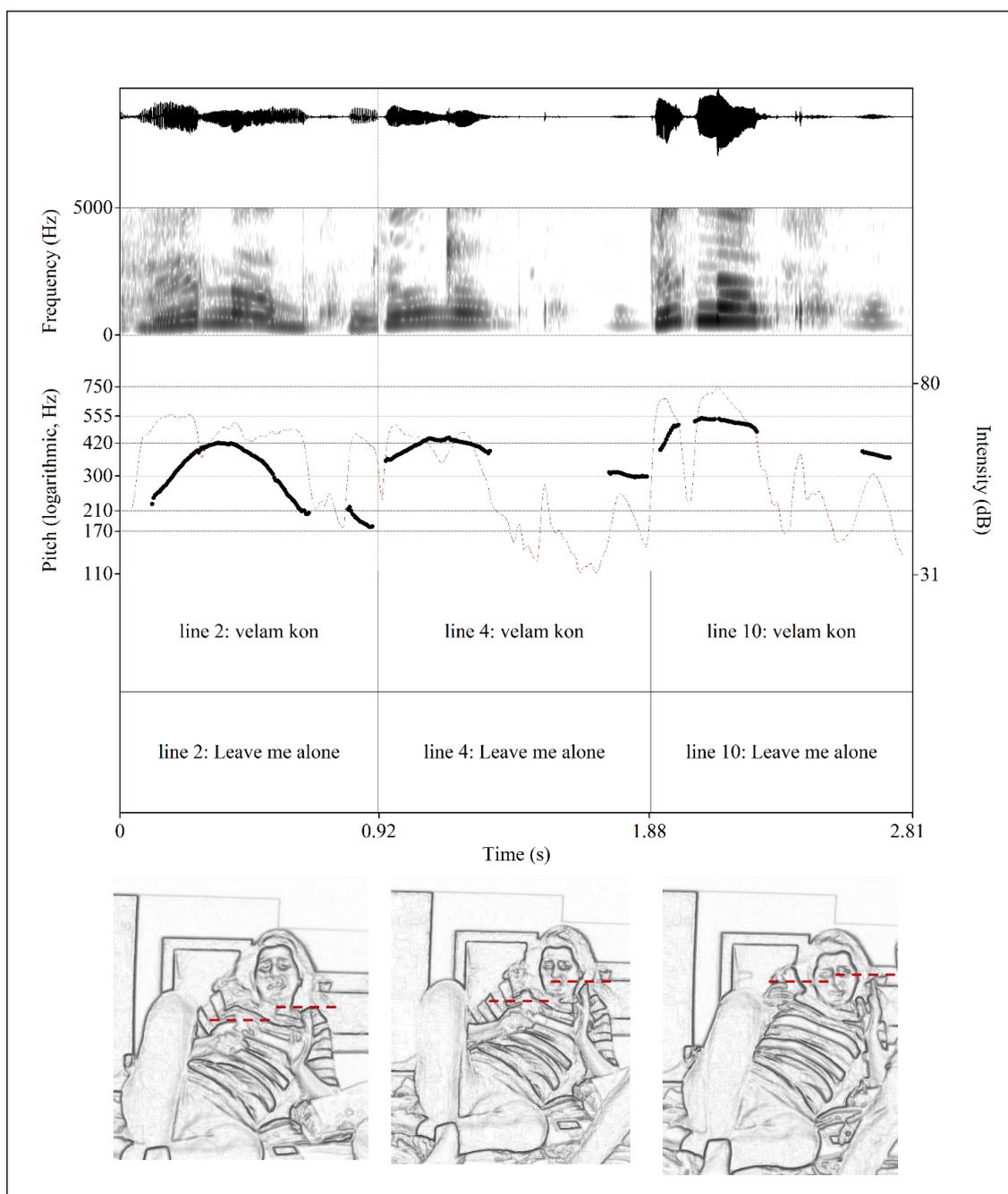
Figure 3. 34. Multimodal upgrading in Extract 3.7.

Therefore, multimodal upgrading is sometimes used after pro-forma compliance. In this sequential position, too, sanctioning practices are likely to join upgraded features. Multimodal upgrades can be heard and seen to be contributing to negative sanctioning (as in the case of furrows in this extract). Sanctioning and re-doing an action what is already resisted by the co-interactant, together, project sequence expansion and escalation in disaffiliation. By sanctioning, the imperative-speaker treats the recipient as being at fault for their ongoing resistance. In this extract, not just the imperative-speaker, but both parties used sanctioning practices, one targeting the ongoing resistance and the other targeting the continued persistence. And they used various means for this purpose: embodied (frown), vocal (click and interjection *sh::*), and verbal ('you should be ashamed'). This indicates that the sequence is on an escalation trajectory. The escalated status is also evident in Saba's minimum 'tolerance' towards any turn which may be about her, manifested in her cutting off Kian's turns.

111

3.6. Discussion

This chapter introduced a multimodal phenomenon termed as *multimodal upgrading*. Multimodal upgrading as conceptualized here involves a coordinated 'intensification' of various turn design features from different modalities: verbal, prosodic, and embodied. We examined multimodal upgrading in pursuit sequences. When a first imperative turn does not secure the recipient's compliance, the imperative-speaker may pursue compliance by repeating the initial imperative and at the same time upgrading the multimodal design features of the imperative. So in the examples in this chapter, upgrading is performed by the same-speaker of a prior imperative turn while re-doing the turn after the recipient's resistance.

It was claimed that multimodal upgrading coupled with negative sanctioning practices tilts the trajectory of interaction towards an escalated state of affairs. A methodological triangulation was applied to support this claim (Campbell & Fiske, 1959). That is, rather than directly associating multimodal upgrading with escalation, which could have resulted in a circular argument, an independent finding was incorporated to the picture. The independent finding was the frequent accompaniment of culturally established negative sanctioning practices with upgraded pursuits. The sequential positioning 'post-resistance' also consolidates the claim since after a coparticipant's dispreferred response, one relevant action, among a range of other possible actions, is to hold them accountable for their resistance. Re-doing the imperative turn provides the recipient a new opportunity for a responsive action. At the same time, holding the recipient at fault for their resistance against the initiated CoA shows that the imperative-speaker does not suffice to only re-doing the action; rather, she resorts to overt disaffiliative actions while pursuing compliance. It is precisely the re-doing of the initiating action accompanied with overt disaffiliative exchanges that manifests escalation. The frequent co-occurrence of multimodal upgrading with negative sanctioning practices can suggest that multimodal upgrades in this context are a part of the overall Gestalt that escalates disaffiliation.

In chapter 2, two types of stances that a pursuing party may adopt were introduced: persisting in forwarding the CoA and gradually desisting from the sequence. It was discussed that a persistent pursuit projects further sequence expansions while in gradual desisting, a sequence closure is likely. In this chapter, the concept of escalation was added to the picture. I distinguish between persistence and escalation. Persistence is conceptualized as a stance, displayed in the practices and design features that the pursuing party employs to further her CoA. Escalation, on the other hand, is the trajectory of the sequence in which disaffiliative actions are performed to further the CoA. Contrary to gradual desisting, a persistent pursuit displays that the speaker is insisting on making her CoA progress, and in this way, it projects more sequence expansions if the recipient does not comply. Escalation happens when these sequence expansions are materialized through increasingly more overt disaffiliative actions (there are other ways by which a pursuing party may insist on their CoA such as affiliative affective pleads, but in this thesis, we study only disaffiliative persistence).

This work develops the literature on 'gradation' (up- and downgrading) in CA (Pomerantz, 1984b; Hauser & Prior, 2019) in terms of form and function. Following enquiries on lexical gradation (Heritage & Raymond, 2005; Hoey et al., 2020; Lerner, et al, 2012; Pomerantz, 1984b), prosodic gradation (Couper-Kuhlen, 2012; Curl, 2006; Local, Auer, & Drew, 2010; Ogden, 2006; 2010; Sikveland, 2019; and Zellers & Ogden, 2013), and embodied gradation (Hauser, 2019; Keel, 2015), this chapter extends the scope of analysis to consider multiple modalities at once. Thus, the analysis here treats turn design features from various modalities as being inter-connected and shaping a unified whole. The concept of multimodal Gestalt is already established in multimodal research (e.g., Mondada, 2014). This chapter adds to this concept by showing that the whole Gestalt can be upgraded. Even though for the sake of analysis, we break the Gestalt into its constituents, the fact that all these constituents are upgraded at once can be evidence for their tight connection.

The findings support earlier work on the relationship between turn design features with sequence expansions and/or escalation. On sequence expansion, for example, Ogden (2010) shows that complaint turns with upgraded prosodic features seek an affiliative response while downgraded complaints project sequence closure (we will see a similar function for multimodal downgrading in the next chapter). Similarly Local and Walker's (2005) study on *holding* versus *trail-off* 'so' illustrates the two types of 'so' differ in terms of their sequential placement and prosodic design. Holding 'so' is reported to be designed as "relatively high in pitch, relatively loud, with final glottal closure", and it functions to project more on-topic talk by the 'so'-speaker (pp. 6-7). Their work is not on the gradation of turn design features; nevertheless, the phonetic cluster 'loud and high' and its relation to turn expansion is similar to the claimed affordance of upgraded multimodal Gestalts projecting sequence expansion (see Ogden, 2017, on the iconicity of high-loud versus low-quiet utterances).

As for the relationship between upgrading and escalation or pursuit, this chapter expands Hoey et al. (2020) and Clift's (2020) findings. In Hoey et al. (2020), the focus is in on expletive insertion, but they also show an example where other multimodal design features of a turn are upgraded in the process of escalation (pp. 12-15). In their example, design features from various modalities (such as what they call 'serious prosody') appear to be added one by one to successive pursuit turns. While this can be an example of multimodal upgrading, the instances in this collection show simultaneous upgrading of design features from multiple modalities all at once. Similarly, Clift (2020) shows that the bodily action 'palm up' used in a pursuit turn upgrades a verbal pursuit. Again, while the 'palm up' phenomenon may be characterized as multimodal upgrading, it does not show various design features of the pursuit turn to be upgraded all at once. Multimodal upgrading can then manifest itself in various forms. The type presented in this thesis reflects a coordinated version of it: various design features from multiple modalities are intensified in a temporally coordinated fashion. In the following chapter, we will see an opposite pattern to multimodal upgrading: multimodal downgrading in pursuit turns.

Chapter 4. Multimodal Downgrading in Pursuit Sequences

4.1. Introduction

It takes at least two parties for some discord to happen, whether it is a small clash, an escalated dispute, or a full-blown fight. And for it to reach an end at least one of the parties must disengage. This chapter is on how disengagement from conflictual episodes may materialize. We continue to investigate pursuit turns in imperative sequences in which the interactants' courses of actions clash and further explore the projective potentials of the design of these sequences. So we are still inspecting the implications that the multimodal design features of a pursuit turn have for the structure of the sequence. An opposite pattern to what we observed in Chapter 3 on multimodal upgrading and sequence expanding is introduced: *multimodal downgrading* projecting sequence ending.

An example of multimodal downgrading was already presented in Chapter 2 (§2.8.2, Extract 2.22) when I introduced the phenomenon gradual desisting. The aim there was to present one of the pursuit types in the collection. I argued that in gradual desisting, the pursuing party displays orientations to backing down, and this projects a contingent ending to the sequence. Three main types of these orientations were briefly mentioned: embodied, verbal, and prosodic. We saw that embodied orientations included resuming prior suspended activities, retracting body to home position, and withdrawing gaze, among others. Prosodic trail-offs and downgraded prosodic features were mentioned to be among prosodic features that project an upcoming ending, and verbal orientations, I noted, could actualize in lexical downgrades or an abandonment of a verbal turn while it is still syntactically incomplete.

In this chapter, we investigate these design features in more depth. I argue that these orientations are in fact parts of a larger package: a downgraded multimodal Gestalt, and even though for the sake of analysis, we will break this Gestalt into its constituents, their recurrent co-occurrence suggests that they are in fact produced as one. Let us examine this in an introductory example. Mona, Ebi, Amir, and Haleh have had their lunch (wife, husband, son, daughter, respectively). They are still sitting on the ground around the food cloth. Ebi has just started to move his body up to sit on an armchair behind him. Concurrently, he addresses Amir, who is sitting with his back towards the camera, to sit in his spot so that he faces the camera. Prior to the extract, he already issued the imperative 'come sit here' twice neither of which was responded by Amir. The extract starts as he uses double vocatives to secure Amir's recipiency and produces the same imperative (line 1, 'Amir Amir you come sit here'). Following Amir's resistance, he reissues the same imperative yet one more time but with downgraded multimodal features (line 9). Amir is silent in this extract. Except for the two imperative turns (lines 1, 2 and 9), the rest of the verbal turns belong to a side sequence between Ebi and Mona which is not our focus.[26] In line 2, *chiz* 'thing' refers to a research microphone. Ebi is not wearing a research microphone.

Extract 4.1. RAY048b, 08.34 [left: Ebi, right: Amir]
```
->   01   EBI    amir #*amir +to bi-â      injâ be-shin,=
                 Name   Name  2SG IMP-come here IMP-sit.2SG
                 Amir Amir you come sit here,
         ebi     >>+get up---+stops getting up-->
         ebi     >>•gaze at AMI-->
         ami          *gaze at EBI-->
         fig          #4.1
```

---

[26] A repairman was visiting their house that afternoon. In line 3, *in* 'this' refers to the repair work. In line 8, the pseudonym Yunesi refers to the repairman. Also, in overlap with Ebi's first imperative, Haleh produces two turns which seem to be self-talking. She expresses how sleepy she is. This self-talk is not transcribed.
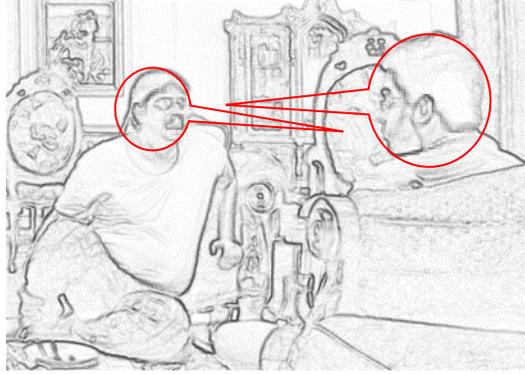
#4.1

```
02   EBI     =ke    toye dorbin bash-i,*#man ke [chiz na-dâr-am.  *•+
             that in   camera be-2SG   1SG PRT thing NEG-have-1SG
             so you're in the camera shot. I don't have the thing.

03   MON                            *          [in  cheghad ebi=*•+
                                                This how.long Name
                                                how long Ebi does this
     ami                            -->*gaze down----------------*
     ebi                                                         -->+
     ebi                                                         -->•
     fig                            #4.2
```
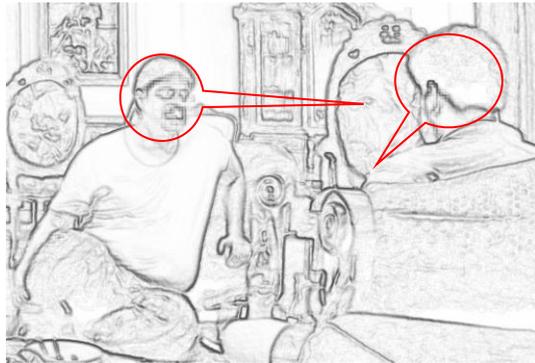

#4.2

```
04   MON     +kâr    dâr-e.
              action have-3SG
               take.
     ebi     +resumes getting up-->

05           (0.2)+
     ebi       -->+

06   MON     +e:::
     ebi     +sits on armchair-->>

07   EBI     chemidun-am cheghad.
             how.know-1SG how.long
             how would I know how long.

08   MON     [• y u    •nesi biâd;
             Name    SBJV-come.3SG
              when Yunesi comes;

->   09   EBI     [†•<<p>bi-â•    †injâ #*be-shin †≡tekye\*†>
                  IMP-come  here   IMP-sit.2SG leaning
                  come sit here, lea\ ((incomplete word))
```

115

```
ebi        •gz at AMI-•
ami                                    *gaze at EBI-------*
ebi        †..............†point----------†,,,,,,,†
ebi                                           ≡stretch left leg-->>
fig                        #4.3
```


#4.3

```
10    EBI    ±nemdunam            ±#
             NEG-know-1SG
             I don't know.
      ami    ±stretches right leg±
      fig                        #4.4
```


#4.4

The two imperative TCUs in focus are produced in lines 1 and 9. Line 1 is not Ebi's first imperative, but since it is the first in the presented extract, I refer to it as the first. Issued after the recipient's (Amir's) resistance, this imperative is multimodally upgraded. The upgraded Gestalt includes embodied, prosodic, and lexical features. Regarding Ebi's embodiment, firstly he suspends his ongoing action of getting up (lines 1-3; Figure 4.1-4.2), and hence shows a "superseding commitment" to furthering the CoA projected by his imperative (Raymond & Lerner, 2014, p. 237). Secondly, he maintains gaze on Amir throughout the production of the turn and the two follow-up accounts even after Amir averts his gaze away from him. In terms of prosody, he speaks rather loudly, and as for the lexical design, the imperative TCU is prefaced with two vocatives, and the first imperative (*biâ* 'come') is produced with a syntactic subject (*to* 'you). This tacitly identifies Amir, the imperative-recipient, as the agent who should be sitting in Ebi's spot.

The subsequent imperative TCU is produced after Amir's resistance. Amir's resistance is manifested in his gaze aversion when Ebi is still offering accounts for his imperatives (line 2) (Kidwell, 2006). This subsequent version is constructed with a downgraded multimodal Gestalt (line 9). The Gestalt has lexical, syntactic, prosodic, and embodied features. Firstly, the two vocatives and the subject of the imperative (*to* 'you') are dispensed with. Secondly, the TCU is abandoned before it reaches a syntactic completion: The phrasal verb *Tekyeh bede* 'lean' is abandoned in the middle of its production, yielding a syntactically incomplete turn. By doing so, Ebi shows reduced commitment to advancing his turn and by extension his CoA. Thirdly, the second imperative turn is produced quieter and faster compared to

the previous one. With the faster design, Ebi shows an orientation to being involved in the side sequence with Mona. The fast design and abandoning the turn together make it possible for Ebi to respond to Mona's question without delay when her turn is over. Hence, Ebi can be heard to be 'prioritizing' Mona's sequence over the imperative sequence. Finally, embodied downgrading can be seen in Ebi's gaze direction and leg movement. Whereas he maintained his gaze on Amir for the production of the whole turn for the first imperative, here his gaze is on Amir only during the production of the first lexical component of the TCU *biâ* 'come' (line 9). Also, he stretches his left leg to the spot which was emptied for Amir to sit (Figure 4.4). By stretching his leg to the spot where he is directing Amir to sit, he tacitly accepts that no compliance is forthcoming. It's only after this that Amir also stretches his right leg, signaling an upcoming re-positioning as unlikely (line 10).

Therefore, the multimodal constructional features of the first imperative TCU show Ebi's heightened engagement in the imperative sequence and project more sequence expansions (line 1), whereas the downgraded multimodal Gestalt in the subsequent imperative indicates a reduced engagement and projects a contingent disengagement and sequence closure (line 9). This is the main claim of the chapter. When multimodal downgrading follows a prior disaffiliative upgraded version, it functions to de-escalate the collision of pursuits.

It is the co-occurrence of multiple downgraded features that conveys downgrading various modalities as constituent parts of a larger Gestalt, one that projects an upcoming withdrawal from the course of action. I will present more details on the coordination between these modalities and more elaborate evidence for sequence ending later in the chapter. To get there, first I review some previous research on (dis)engagement and downgrading (§4.2 and §4.3, respectively). Then various resources that are recurrently seen to undergo downgrading are separately presented through examples from the Persian corpus in §4.4. Next, we examine the fine coordination among downgraded modalities (§4.5). Then evidence on my main claim on multimodal downgrading and sequence ending is presented in two main sequential environments: multimodal downgrading after compliance (§5.6) and after resistance (§4.7). In §4.8, multimodal downgrading and upgrading are compared, and finally, the chapter is concluded with a summary of the findings in §4.9.

4.2. Disengagement

What does it mean to withdraw or disengage from a course of action? Let us take Goffman's (1963) notes on *involvement* as our point of departure: "To be engaged in an occasioned activity means to sustain some kind of cognitive and affective engrossment in it, some mobilization of one's psychobiological resources; in short, it means to be involved in it" (p. 36). By extension, disengagement from an activity entails deactivation of those resources. Goffman notes that in the process of deployment of these resources, outward expressions and behavioral cues reveal participants' allocation of involvement.

Research in CA, focused on participants' observable conduct, has been fruitful in documenting some of these expressive cues on participants' engagement. Early notable contributions to this area were Goodwin's (1981) work on participants' gaze behavior and Schegloff's (1998) on body configurations. Goodwin (1981) shows detailed gaze patterns by speakers and recipients in various types of engagement frameworks: mutual engagement, mutual disengagement, entering disengagement, and so forth. Relevant to this chapter is his observation that even though recipients interpret a speaker's gaze withdrawal with reference to other intervening or ongoing activities, and not necessarily a sign of full disengagement, a speaker's gaze withdrawal can indicate "less than full engagement" in the ongoing course of action (p. 105). In a later work, Rossano (2012) shows that interactants' gaze behavior is sensitive to the sequence structure; that is, gaze withdrawal at points of possible sequence completion projects that the sequence is likely to end, contingent on the withdrawal of gaze by the other participant (in a dyadic interaction). In contrast, gaze maintenance at possible sequence completion projects sequence expansion.

Displaying and ranking involvement through embodied means, more specifically through body torque, is also noted by Schegloff (1998). Informed by Kendon's (1990) work, he defines body torque as "different or diverging orientations of the body segments above and below two major points of articulation – the waist and the neck" (p. 540). When a body is in torque, Schegloff notes, it projects

instability and change, potentially a return to the home position. Home position is "the position from which some limb or physical movement departed, and the return to which marks a possible ending to a spate or unit of activity" (p. 542). The argument in this chapter is in part built on his observation that a "release from torque away from the interaction and towards home position [happens when] the sequence is coming to a projected end" (p. 579). We will see that moving body parts to the home position is in fact a part of a multimodal Gestalt.

Further research on gaze and body posture (e.g., Heath, 1984; Ruusuvuori, 2001; Robinson, 1998), verbal, vocal, and bodily movements (e.g., Blythe et al., 2018; Szymanski, 1999), and lapses in conversations (e.g., Hoey, 2020) has also been conducted in relation to (dis)engagement. Heath (1984) shows how an interactant may use gaze and posture to display recipiency directly towards a co-interactant and that elicits speech or action from the co-interactant and encourages their involvement in the talk (p. 249). In the environment of group work, Szymanski (1999) notes verbal practices such as questions, noticing and announcing to be verbal means by which participants re-enter a state of engagement. Other visual and vocal cues can also prepare the scene for re-engaging in a task at hand (Szymanski, 1999). Additionally, Hoey (2020) elaborates how participants may treat lapses – "silences that develop when all conversationalists refrain from speaking at a moment when speaking was possible" – as a relevant place to disengage from talk (p. 2).

Some of these bodily displays are shown to be culture dependent. For example, Goodwin's (1981) study on English speakers shows that a coparticipant's gaze withdrawal can project their disengagement whereas Rossano, Brown, and Levinson's (2009) study shows that Tzeltal speakers in Mayan community in Mexico "tend not to gaze at speakers' faces while listening to them" (Rossano, 2013, p. 310).

A relatable line of research is on how participants calibrate dual involvements in simultaneously ongoing activities (Raymond & Lerner, 2014; Kamunen, 2019). Raymond & Lerner (2014) illustrate two practices – suspending and retarding – by which an interactant manages dual involvements in two courses of actions. In suspending, an interactant halts a course of action to deal with another interjecting action but retain their body in a "ready-to-resume position" to return to the interjected action. In retarding, the forward progress of the ongoing course of action is slowed by the interjecting action.

Prosodic features can also project a speaker's withdrawal from a sequence. For example, Couper-Kuhlen (2004) shows that a low pitched and quiet *Never mind* in complaints or troubles talk is used to exit the ongoing sequence. Local and Walker (2005) also show a similar function for tail-off *so* with low and quiet prosodic features. Similarly, doubles with a lower pitch, narrower span, and shorter duration in the second unit are shown to close sequences (Curl, Local & Walker 2006). Again, in complaint sequences, Ogden (2010) argues that low and quiet complaint turns work to close the sequence and disengage from the complaining action. Focusing on the final intonation patterns of 'formulations' in French, Persson (2013) argues that a final rise projects expansion relevance while a rise-fall pattern projects closing relevance. At the level of turn-taking, Walker (2012) notes that trail-off conjunctions have phonetic characteristics of "other designed-to-be- and treated-as-complete utterances" (p. 148). For example, there is no "audible or acoustic-graphically visible final glottal closure" or final assimilation. Rather, there are turn-ending pitch features (e.g., Szczepek Reed, 2004), such as "no disjunctive changes in pitch or loudness between the talk leading up to the conjunction and the conjunction itself" (p. 147). Walker notices that trail-off conjunctions at the end of a speaker-turn accompany visible embodied behavior, for example, looking down and moving to the next embodied action, from both participants by which they treat that the speaker's turn is coming to an end. In this chapter, we see that many of these expressive cues on (dis)engagement noted by different scholars can be in fact a part of a larger multimodal Gestalt, and together they project sequence ending.

4.3.   Downgrading

In chapter 3, the CA literature on grading was reviewed. Here, the mention of only a few selective studies on the interactional affordances of downgrading will suffice. Like upgrading, downgrading was first discussed by Pomerantz (1984b) in the analysis of second assessments. "A downgraded agreement

is an assessment of the same referent as had been assessed in the prior with scaled-down or weakened evaluation terms relative to the prior" (Pomerantz, 1984b, p. 62). Focused on lexical features, Pomerantz shows that a weakened lexical descriptor in second assessments engenders a weak agreement with the first assessment. In the following excerpt, 'pretty' is the weakened evaluative term compared to 'doll'.

Extract 4.1. [NB: VII:2] (Pomerantz, 1984b, p. 68)

```
1    E:     e- that Pat isn'she a do:[::ll?

2    M:                             [iYeh isn't she pretty,
```

Other types of lexical downgrades are reported to happen in using evidential verbs such as 'seem' in assessment turns or turn-final *or* in English by which speakers downgrade their epistemic stance (Drake, 2015; Heritage & Raymond, 2005). Similarly, 'I'm wondering' prefaces or accounts are shown to downgrade consultants' deontic stance in therapy sessions with their clients when making proposals (Ong, Barnes, Buus, 2021).

In the domain of prosody, downgrading can afford different interactional functions. For example, in second assessment turns, a prosodically downgraded agreement is shown to preface disagreement (Ogden, 2006; see also Couper-Kuhlen, 2012 on the prosodic design of disaffiliative turns in complaints). In other initiated repairs, Curl (2005) shows, downgraded repeats display *disjunct* repairs. The trouble source of such repairs is treated as "sequentially or interactionally inappropriate in relation to the prior turn and/or sequence" (p. 10). Also, in the context of interjected or aborted turns, repeating the turn with downgraded prosody is associated with retrieving that turn as opposed to redoing or resuscitating it (Local, Auer, & Drew, 2010).

This chapter shows that multimodally downgraded subsequent imperatives are used by a pursuing party as a resource to withdraw from the CoA. When they are produced after escalation, they function to de-escalate the disaffiliative situation. Downgraded imperatives either close the imperative sequence or project a contingent closure. The former occurs after the recipient's compliance, and the latter after their resistance. Before presenting evidence for these claims, first I show various resources by which an imperative may get downgraded, using the examples from the Persian corpus (§4.4). We then see the fine coordination between various modalities that can be involved in a downgraded multimodal Gestalt (§4.5). Next, common design features of upgraded versus downgraded multimodal Gestalts are compared (§4.6). Evidence on the relationship between multimodal downgrading and sequence closure after compliance and after resistance is presented in §4.7 and §4.8, respectively. This is followed by another comparison between multimodal upgrading and downgrading in terms of their interactional affordance (§4.9). Finally, a summary of the findings is presented in §4.10.

4.4. Resources for Downgrading

There are different resources which participants can use to produce downgrades, two of which that are already established in CA are lexical and prosodic. In this section, first, lexical and prosodic downgrading are reviewed through some examples from the current pursuit collection (§4.4.1 and §4.4.2). Then I extend the concept of 'downgrading' to the embodied domain (§4.4.3). It is worthy of mention that in the 'downgrade' subcollection, the downgraded turns are of two main types: downgraded imperatives relative to a prior upgraded counterpart, and downgrades relative to a 'neutrally'- or 'unmarked'-designed counterpart. Both types are similar in terms of their typical design features and interactional affordance; therefore, this chapter includes examples of both types.

4.4.1. Lexical downgrading

In pursuit sequences, an imperative can be lexically downgraded compared to its previous version. Lexical downgrading in a subsequent version can be achieved by using a verbal or vocal component which is semantically 'weaker' compared to a prior point of reference. For example, in the following extract, the imperative 'shut your mouth' (line 1) is replaced by the interjection *s:*, which is used to direct someone to be quiet ('sh::' in English). Prior to the extract, Azita starts talking about Roshana's dream. The first imperative is designed to prevent further talk about the dream.

```
->    01    ROS      dahan-t-o              be-band.
                     mouth-2SG-OBJ.MAR IMP-close.2SG
                     shut your mouth.

      02             (0.2)

      03    AZI      bezâ    khâb-e-sh-am     tarif [kon-am.
                     IMP-let dream-of-him-also explain do-1SG
                     Let me tell the dream too

->    04    ROS                                    [s:
                                                    s:
                                                    sh:
```

In line 1, an imperative format, with a rather 'offensive' lexical design, is used to stop Azita from talking. Azita does not comply. She displays explicit refusal by initiating a CoA in which she tells the dream. Roshana then produces a subsequent version to hinder Azita's CoA. The subsequent version is only the vocal interjection *s:* 'sh:', which is analyzed as a downgrade compared to the first version. Firstly, it is typically less offensive than the imperative *dahanto bebabnd* 'shut your mouth', and secondly, it is less precise. The imperative-speaker withdraws from the imperative sequence even though the recipient continues talking about the dream (not transcribed).

4.4.2. Prosodic downgrading

An imperative-speaker may downgrade the prosodic design of their turn by reducing some prosodic measures such as intensity, pitch span, average F_0. Non-measurable prosodic features such as articulatory setting can also be manipulated. For example, a more open articulatory setting compared to a tighter one can contribute to a sense of downgrade (see e.g., Ogden, 2006). The following example shows a downgrade from a prior prosodically 'unmarked' version. Melisa and Hamun are sitting on different armchairs on the opposite sides of the room. Melisa picks a game card pack to start playing with Hamun. Concurrently, she addresses Hamun with the imperative 'come here' (line 1), the acceptance of which makes it possible for both participants to access the cards. Hamun does not complete Melisa's initiated CoA: he does not move and continues his interaction with another participant (line 2 is addressed at a third participant). Melisa then recycles her imperative in a prosodically downgraded design (line 3).

Extract 4.3. RAY037a(2)-b, 14.32

```
->    01    MEL      +biâ       injâ*::,
                     IMP.come here
                      come here
                     +pick the card pack-->
            ham                  *gaze at card pack in MEL's hand-->>

      02    HAM      <p>na yeki dâsht-i> ((addressed at a third participant))
                     no one  have.PST-2SG
                     No you had one

->    03    MEL         +<l, p>biâ injâ,>
                            IMP.come here
                            come here
                     -->+opens the pack-->>
```

The subsequent imperative is prosodically downgraded relative to the first version (Figure 4.5, Table 4.1): the TCU's average F_0, pitch span, intensity, and the duration in the second version (line 3) are reduced compared to the first imperative (line 1). Whereas the first imperative is produced around the speaker's average F_0 level (on average 260 Hz vs. the speaker's overall average level, 250 Hz), the

subsequent version starts and ends lower in the speaker's pitch range. Overall, the second TCU is about 2 ST below the speaker's average F_0. The average intensity drops from 73 dB in the first to 62 in the second and the pitch span is reduced from 8 ST to 6 ST. Also, the subsequent imperative is produced shorter: in the first imperative, the last syllable is elongated (430 ms) whereas the second one is produced with less lengthening (280 ms). The combination of these features conveys a hearably downgraded design in the subsequent imperative. The recipient remains sitting in his own spot and the speaker temporarily abandons the sequence (not transcribed for brevity).

Table 4.1. Summary of the prosodic features of imperatives in Extract 4.3.

| Line	Average F_0 of the TCU (Hz)	Pitch Span (ST)	Intensity (dB)	Duration (ms)
1	260 (1.0 ST above speaker's average F_0)	8	73	950
3	220 (2.0 ST below speaker's average F_0)	6	62	730
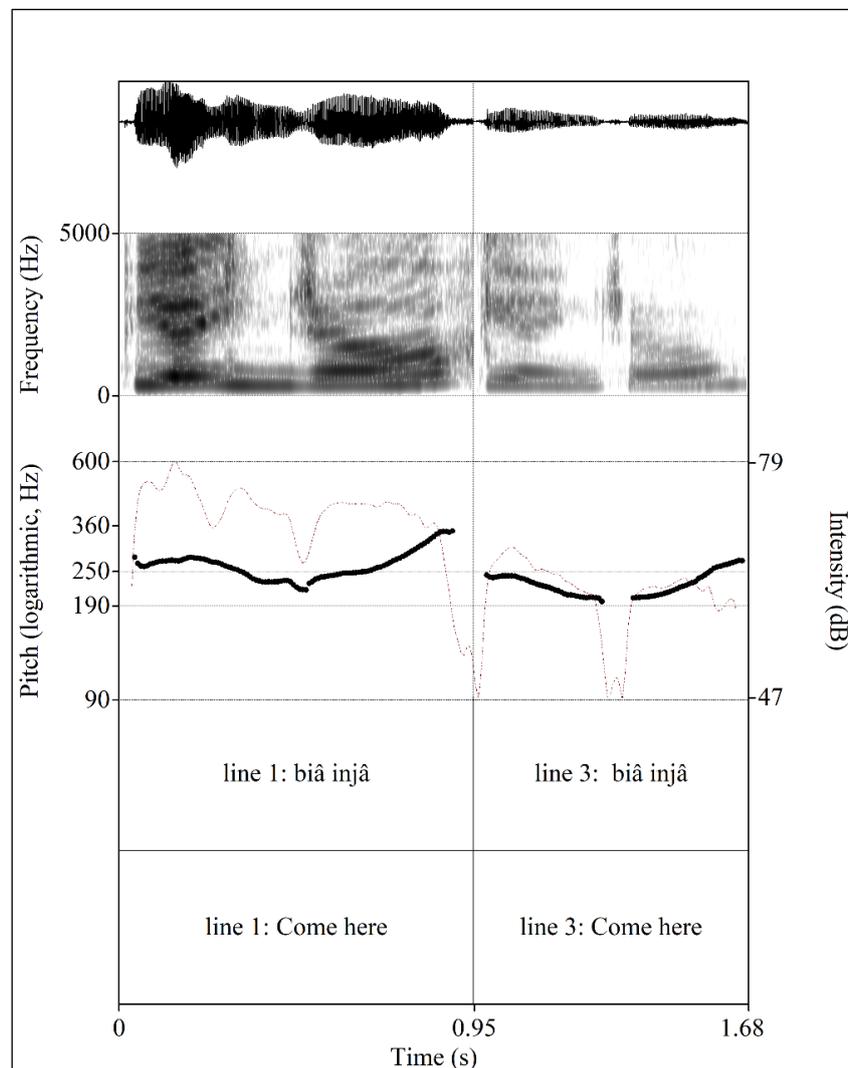
Figure 4.5. Prosodic downgrading in Extract 4.3.

4.4.3. Embodied downgrading

Similar to verbal and prosodic modalities, embodied actions can also be downgraded. As with embodied upgrading, conceptualizing embodied downgrades merely in terms of their physical properties is problematic (see Chapter 3, §3.3.3). This is because one particular physical property such as 'movement away from a co-interactant' may be analyzed differently in different situations. For example, although in this collection movements away from the coparticipants or a relevant object are recurrently seen to

contribute to a sense of downgrade (e.g., averting the direction of gaze away from the coparticipant, retracting hands or other body parts away, etc.), 'movement away' cannot be taken as an absolute feature for downgrading. To give an example from the corpus, a participant may move away from a resisting party and while doing so they slam the door hard disapproving of the coparticipant's resistance. This body movement away from the coparticipant is not oriented to as a downgrade by the coparticipant herself. The same is true for dispensing with or inserting bodily actions. In this collection, sometimes downgrading is achieved through dispensing with bodily actions that can be associated with disaffiliation (e.g., stopping a furrow addressed at the recipient) and in some other examples, it is achieved by adding embodiments that may have implications for affiliation and alignment (e.g., smiling or a downward nod conventionally doing acknowledgement).

Therefore, analyzing participants' bodily configuration in terms up- or downgrading depends on a pool of inter-related factors such as the involved bodily articulators (e.g., eyes, head, hands, torso, legs, etc.), the type of bodily conduct that is performed with these articulators (e.g., with eyes one may redirect their gaze, or they may wink, etc.), the CoA in which the embodied adjustments occur and the ecology of the interaction (Mondada, 2018), and most importantly, the interactional outcome of these bodily movements.

In broad strokes, embodied downgrades in this collection are bodily movements that contribute to a sense of backing down following and relative to a prior version. As in any other scalar phenomena, embodied downgrading is relational, and it should be analyzed in light of the larger social situation and interactional goals. The list below thus is intended to present only some recurrent features observed in their particular contexts, and it should not be treated as exhaustive or absolute measures for embodied downgrades.

Direction of movement: retracting body parts toward home position is one possible way for embodied downgrading to occur (but see Chapter 3, p. 12). For example, an imperative-speaker may retract their stretched arm away from a requested object.

Contact or involved body parts: sometimes reducing the contact size with the coparticipant's body/a relevant object or involving fewer body parts in an action achieve embodied downgrading. For example, an imperative-speaker, who touches a requested object with a full palm with the first imperative, reduces contact and touches the object with only fingertips in the subsequent version.

Force*: the force by which embodied action is performed can be reduced. This can be judged by the visible outcome of the embodied action. For example, when the force of a pushing action exerted on the coparticipant is reduced, the coparticipant's body is seen to be moved less.

Space: increasing distance between interactants or an interactant and a relevant object. This can be achieved by retracting the whole body or body parts from its previous position. For example, a participant who has walked towards an object, concurrent to requesting the object, walks back as she produces a subsequent version.

Dispensing with a prior bodily movement: while producing a subsequent version, an embodied action which was originally performed with an imperative, can be dispensed with. For example, tapping on a requested object simultaneous to a prior imperative is dispensed with in the production of a subsequent version.

Performing aligning bodily movement: this involves adding bodily actions that display the imperative-speaker's alignment with the recipient's CoA. For example, following the recipient's resistance to give her mobile phone to the speaker, the speaker may pick up her own mobile phone projecting working with it as she produces a subsequent imperative.

The extracts in the remainder of this chapter each show at least one of these ways. As we saw in chapter 3, embodied actions can be resources for getting a participant to do something, without having to accompany other multimodal resources. To re-establish this, this section presents an example in which a participant's embodied action, on its own, displays the ongoing relevance of the coparticipant's response. Nasima has been showing pictures of a singer to Melisa, her mother (and a third participant, not in the transcript). She finds the pictures on the Internet, gives Melisa the mobile phone and after

Melisa and the third participant see it, she takes the phone back and finds the next picture. This process has been going on for a while. In the extract, Nasima is holding her phone when Melisa verbally requests to see the next picture. Concurrent with the request, Melisa stretches her fingers towards the phone and starts moving them back and forth. The embodied downgrade is actualized first by dispensing with the back-and-forth movement, then by reducing the extension of the fingers, and finally a full retraction of the hands.

Extract 4.4. RAY030c-d, 19.12 [left: Melisa; right: Nasima]

```
      01           ±(1.0)
         nas       >>holds cell phone-->>
         nas       ±thumb goes up and down in a scrolling movement-->

->    02   MEL      + be-bin-am#
                       SBJV-see-1SG
                     Let me see {it}.
                    +fingers strtched twrd NAS-->
         fig                  #4.6
```


#4.6

```
      03           (2.0) †(0.5)±†
         nas                 -->±
         mel            †fingers back&forth†

      04   NAS      bas-e        dige+
                    enough-be.3SG PAR
                    It's enough
         mel                     -->+

      05   NAS      +kheili negâsh kard-in=#
                     very   look-3SG-2PL.PST
                    You watched her too much
         mel        +fingers laxed held twrd NAS-->
         fig                            #4.7
```


#4.7

```
      06   NAS      =((giggles))+
         mel               -->+
```

123

```
07      MEL     +hasudi-t      shod?#
                jealousy-2SG become.PST
                Are you jealous {of her/the singer}?
                +retracts hands-->>
        fig                            #4.8
```



```
                                            #4.8
```

This extract shows the process of downgrading an embodied action by gradually retracting relevant body parts: (a) the extension of fingers is reduced; (b) a back-and-forth movement is omitted; (c) space between hands and the requested object is increased by moving hands back to home position (Figure 4.9). Nasima is holding her cell phone with her gaze down at it and her thumb moving up and down as in a scrolling movement, which can potentially display her search for a next photo (line 1). Melisa then reaches out (Kendrick, 2020, p. 107) to request the phone (line 2). This request is designed with multimodal features: through the verbal TCU 'Let me see {the photo}[27]', Melisa makes relevant the transfer of Nasima's cell phone. There is also an embodied component to this request. Melisa extends her fingers towards the phone (line 2, Figure 4.6) and holds them extended for 2.0 seconds, then she adds a back-and-forth movement to the extended fingers (line 3). Both hand-stretch and the back-and-forth movement of the fingers display the ongoing relevance of the transfer of the cell phone. Since the transfer of the phone requires passing it from Nasima's hands to Melisa's, stretching hands towards the requested object symbolically facilitates the transfer; therefore, Melisa's embodied action is recognizable as an embodied request (see Kendrick, 2020, p. 107 on the orientation and shape of manual actions specifying a relevant next action). Nasima refuses to complete Melisa's initiated CoA by giving an account for a rejection (line 4).


| Fingers stretched | Fingers laxed | Hands retracted |
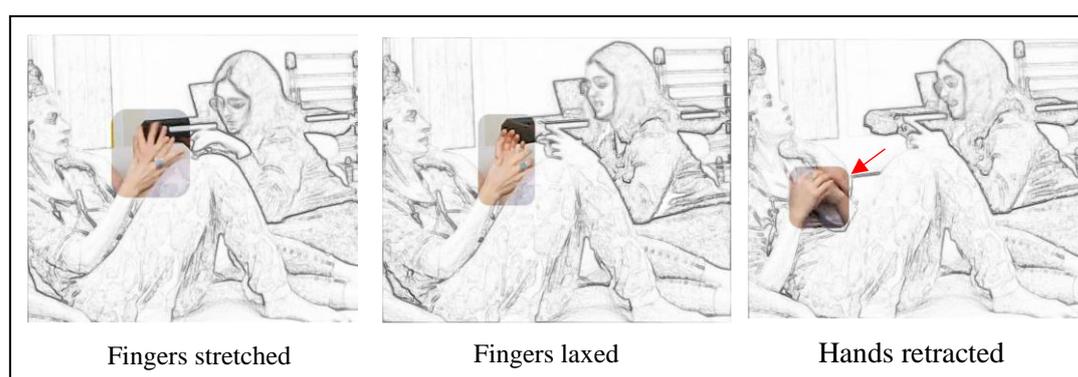
Figure 4.9. Embodied downgrading in Extract 4.4. The incremental retraction of fingers and hands indicates the process of embodied downgrading.

After the recipient's implicit rejection, the request-issuer downgrades the embodied design of her request by relaxing – reducing the extension of – her fingers and stopping the back-and-forth movement,

---

[27] Bebinam in Persian is grammatically a subjunctive, literally meaning 'so that I see (it)', but in this context, it is heard as a truncated imperative: 'give it to me so that I see (it)' or 'Let me see (it)'.

but she continues to hold her hands up towards the cell phone (line 5, Figure 4.7). This combination of continuing to hold her open palms up *and* at the same time reducing the extension of fingers and dispensing with the back-and-forth movement is analyzed as the embodied downgrade. The downgrade is solely performed by means of bodily movements with no other modalities involved. Melisa's 'unstable' hand position, i.e., holding her hands up, shows that the CoA for which she originally moved her hands up is not yet over (see Schegloff, 1998 on unstable body positions and Raymond & Lerner, 2014, on ongoing relevance of an action). Therefore, even though the embodiment is downgraded and there is no verbal accompaniment, it still makes the transfer of the cell phone as Nasmia's next expected action.

Following the recipient's continued resistance, the requester fully retracts her hands. Nasima gives a second account for rejection (line 5). As Melisa continues to hold her hands up, Nasima giggles (line 6), potentially orienting to her own disaffiliative action. At this point, Melisa fully retracts her hands and initiates a new sequence enquiring an account for Nasima's rejection (line 7, Figure 4.8). Therefore, following the recipient's resistance to complying with the request, the requester first downgrades her embodied conduct and then fully retracts the involved body parts.

Rather than immediately retracting hands to home position after the recipient's first rejection, the requester withdraws hands in steps and by so doing, she delays the withdrawal process – gradually desisting (see Raymond & Lerner, 2014 on retarding). Through her embodied downgrade, she displays backing down, gradually disengages from the request sequence, and then she initiates a new sequence.

4.5.   Multimodal Downgrading

Now the phenomenon of interest – multimodal downgrading – is presented in more depth. As with multimodal upgrading (Chapter 3), it is the congruent and temporally coordinated integration of various modalities which accredits multimodal downgrading as a single phenomenon, rather than simply an overlapping occurrence of two separate interactional events. The following example is presented to show such integration. The participants are consulting a book to find an answer to a question. Safura is holding the book. Before this extract, she reads some lines of the book aloud and criticizes it for not giving them any answers. Line 1 is the continuation of her criticism – a negative assessment of the book (i.e., too wordy). This assessment also functions as an account for her inability to find an answer in the book. Partially overlapping with this, Kati produces the first line of interest, directing Safura to give her the book ('Give {it to} me', line 2). It is this imperative which first gets upgraded (line 2) and then later multimodally downgraded (line 6). Thus this extract shows an instance of downgrade from an upgrade. The focus is on the multimodal downgraded design of line 6 compared to line 2; however, the upgrade is also explained to provide a better ground for comparison.

Figure 4.10. Participants in Extract 4.5. Left to right: Kati, Faraneh, Safura. Safura criticizes the book she is holding. Kati defends the book. Faraneh takes side with Safura.

Extract 4.5. RAY040e-f, 18.26

```
01   SAF   hamash#+[(.) neveshte dare+
           all-of-it   writing   have.3SG
           There's too much text.
     kat   >>.....+
     saf   >>gaze at the book-->>
```

```
->    02    KAT          #+[b±ede#±man.±<<f,h>bede#±man;>±<<ff,hh>bede#±man,>±
                         give   1SG        give   1SG              give 1SG
                          Give {it to} me. Give {it to} me; Give {it to} me,
                         +arm extended—->
                          ±bend±,,,,±bend fingrs±,,,,,± bend fingers±,,,,,±
              fig        #4.11    #4.12           #4.13              #4.14
```
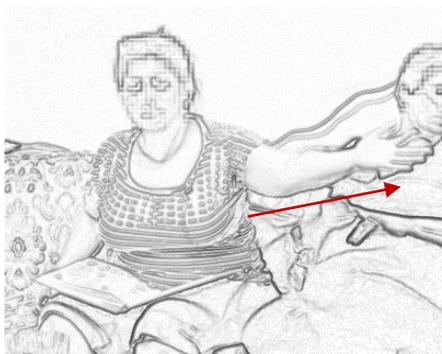

#4.11

#4.12   #4.13   #4.14

```
      03    KAT      balad ni-st-in=
                     guide NEG-be-2PL
                     You {pl} don't know how {to interpret the text}.

      04    FAR      [na:: (.) na na
                      no       no no
                      No (.) no no

      05    KAT      =[kheili-am khob javâb dâd
                      very-too good answer give.3SG
                      It gave an answer very well.

->    06    KAT      +<<l,p>be-d-e>+#
                      IMP-give-2SG
                       Give {it}.
              kat    -->retracts arm--+holds arm-->>
              fig                      #4.15
```
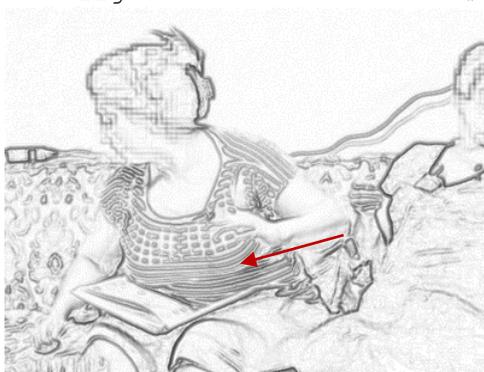

#4.15

The first line in focus (line 2) is produced in overlap with Safura's accounts for her inability to find an answer in the book (line 1). Kati issues the imperative 'Give {it to} me' three times, with each successive one latching to the prior (line 2). The imperatives project a CoA in which Kati herself searches for an answer in the book. Safura does not produce any response. She keeps her gaze down on the book, which displays her engagement in her own CoA – reading the book. Kati then holds both Safura and Faraneh accountable for not knowing how to use the book (line 3), which Faraneh disagrees with (line 4). In overlap, she defends the book (line 5) and then she pursues compliance through repeating the same imperative in a modified format with downgraded multimodal design features (our second line of interest, line 6).

There is a congruent and temporally coordinated relationship between the prosodic and embodied modalities. This coordinated relationship is observable in the symmetrical increase of prosodic measures and the size (length) of extension of the speaker's arm throughout line 2 as opposed to the symmetrical decrease of the same modalities in line 6 (Figure 4.16, Table 4.2).

Let us start with the first three imperatives in line 2. Each version is produced as an intonational unit. Each successive version is prosodically upgraded relative to its previous version: it is designed higher in the speaker's pitch range and louder compared to its previous counterpart (see Table 2 for the details). There is an opposite pattern to a natural pitch declination which typically occurs throughout the production of one breath-group (see Liberman, 1967 on breath group). Therefore, even though these versions are latching, the second and the third production are heard to be upgraded.

The speaker's bodily movement is coordinated with the verbal and prosodic design of the imperatives (line 2). Throughout the production of the first three versions, the imperative-speaker's arm is extended and her head is oriented to front and slightly right, where the recipient is sitting (line 2). Temporally, Kati initiates the preparation of her arm extension towards the requested object such that by the beginning of the first imperative, she reaches the apex of the extension. Then she embodies three beckoning gestures, one with each imperative (Figure 4.11). The beckoning gestures are temporally coordinated with the prosodic design of the imperatives: Kati bends her fingers towards herself with the production of the accented syllable *be* – the imperative-maker morpheme – and then opens them with the rest of the imperative (Figures 4.12, 4.13, and 4.14). In this context, the combination of these bodily movements – extending the arm towards the object and bending fingers towards oneself – makes relevant the transfer of the object in addition to the verbal imperative (see Kendrick, 2020, p. 107).

Table 4.2. Summary of the multimodal features of the imperatives in Extract 4.5.

Line	Average f0 of the rejection TCUs (Hz)	Intensity (dB)	Span (ST)	Duration (ms)	Bodily actions
2a	291 (3.3 ST above speaker's average f0)	72	6	470	Arm extension
2b	324 (5.1 ST above speaker's average f0)	73	3	370	Arm extension
2c	419 (9.6 ST above speaker's average f0)	76	4	240	Arm extension
6	260 (1.3 ST below speaker's average f0)	63	3	320	Arm retraction

Then the fourth rendition is multimodally downgraded. The downgraded Gestalt consists of verbal, embodied, and prosodic resources (line 6). Firstly, the object pronoun *man* 'me' is dispensed with. In the two prior upgraded versions, Kati had kept the object pronoun 'me' even though they were produced in a sequence-subsequent position (Schegloff, 1996b), the person to whom the book should be passed is clear from the context, and Persian allows for dropping such objects. By omitting the object 'me' in line 6, which is again in a sequence-subsequent position, she reduces the specificity of her turn, compared to the previous versions, and makes the imperative verbally downgraded. Secondly, the intensity and the average F_0 of the TCU is decreased (Table 4.2).

The figure contains the following labels:

- Frequency (Hz) axis: 5000, 0
- Pitch (logarithmic, Hz) axis: 600, 450, 300, 240, 100
- Intensity (dB) axis: 79, 57
- line 2: bede man bede man bede man
- line 6: bede
- bede man | bede man | bede man | bede
- line 2: Give {it to} me | Give {it to} me | Give (it to} me | line 6: Give {it}
- beat | beat | beat
- Time (s): 0, 0.47, 0.845, 1.09, 1.40
- Arm extension throughout line 2
- Arm retraction in line 6
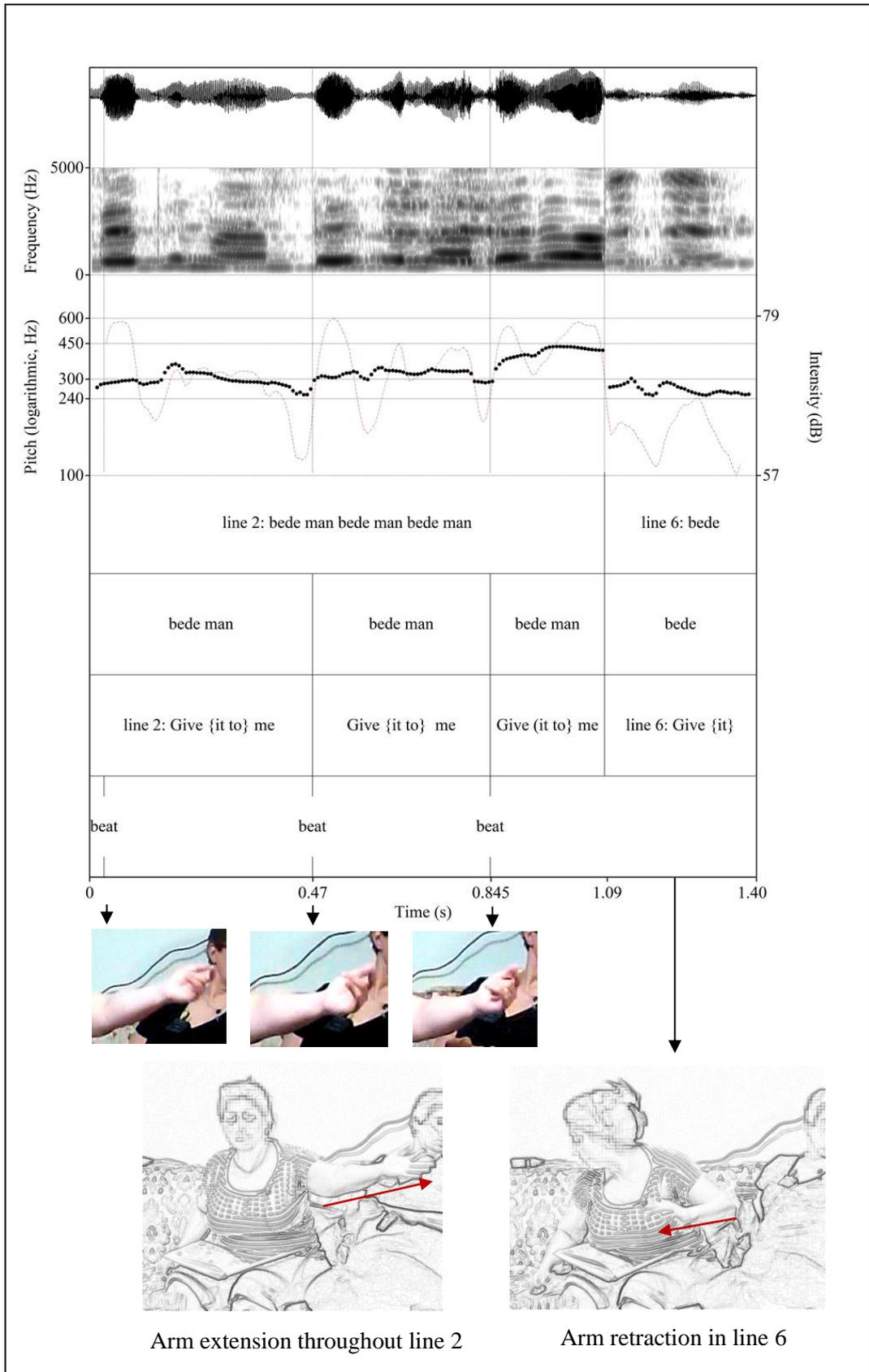
Figure 4.16. Multimodal downgrading in Extract 4.5. Multimodal downgrading (line 6) involves a drop in intensity and the overall pitch and also arm-retraction and head aversion. Compare it with the previous versions (line2) where the arm is fully extended, head is oriented to front and slightly right, where the recipient is sitting. The overall pitch level and intensity successively increase as well.

128

And finally, the embodied downgrading is manifested in two moves: arm retraction and head aversion. Her arm which was kept fully extended over the production of the previous versions, gets retracted (Figure 4.15). What makes Kati's imperative temporally related to her bodily action is duration of her arm retraction: she retracts her arm for as long as the imperative is being produced. Then she halts the retraction process. Only later, she fully retracts her arm (not transcribed). She also averts her head towards left, away from the recipient. After this extract, Kati puts an object that she has been holding in her left hand on the left side of her body. So this head aversion to the left already projects her involvement in her next CoA and a contingent ending of the imperative sequence.

In short, this extract shows the directional congruence and temporal coordination among the three modalities involved. It also illustrates the opposite patterns (extension vs. retraction) often observed in multimodal upgrading (line 2) versus multimodal downgrading (line 6). That is, concurrent to holding her arm fully extended, the directive-issuer successively upgrades the prosodic design of the imperatives. Inversely, the reduction of the extension size of the arm co-occurs with downgraded prosodic measures. Thus, the two modalities are congruent in the increase versus decrease of measurable features. Furthermore, the fine-tuned temporal coordination is manifested in (a) the timely beckoning gestures over the successive imperatives in line 2 and (b) performing the retraction movement only for as long as the subsequent version is under production in line 6. This directional and temporal alignment between the modalities yield a multimodal Gestalt which had different affordances in various sequential positions. In this chapter, two of these affordances in two main sequential positions in pursuit sequences are discussed: closing the sequence (§4.7) and projecting a sequence closure (§4.8), but before that a brief comparison on how various features work in multimodal upgrading and downgrading is presented.

4.6. Features in Multimodal Downgrading and Upgrading

Having seen some of the features that contribute to a downgraded multimodal Gestalt, we can now compare upgrading and downgrading in terms of the parameters within various modalities. As mentioned earlier, abstracting away from examples and generalizing features may make some of them look ostensibly contradictory, for example see the row on the rate of articulation (see also a§3.3.3 and §4.4.3 on other potential contradictions); therefore, I re-emphasize on my previous point that these features must be examined within their contexts.

Table 4.3. Summary of some features in upgraded vs. downgraded multimodal Gestalts.

Modalities	Features	Upgrade	Downgrade
Verbal	semantics of lexical or vocal components	Stronger	Weaker
	affective components or adverbs of urgency	Stronger	Weaker
	syntax of verbal TCUs	Tend to be syntactically complete	May be left syntactically incomplete
Prosodic	intensity, average F_0, and pitch span	Increased	Decreased
	articulation setting	More closed	More open
	rate of articulation	Increased or decreased	Increased or decreased
Embodied	direction or orientations of the articulators	Moved or averted towards	Moved or averted away
	force, speed, contact size	Increased	Decreased
	relevant bodily actions (e.g., tapping, pointing)	Added	Omitted
	aligning bodily actions	Omitted	Added

In the rest of the chapter, we explore some of the interactional affordances that a downgraded multimodal Gestalt can achieve in two main sequential positions: after compliance (§4.7) and after resistance (§4.8). It will be shown that a downgraded imperative after compliance closes the sequence and hence it is not an initiating imperative; rather, it is a sequence closing imperative. After resistance, on the other hand, a downgraded imperative pursues a preferred response but at the same time projects a contingent sequence closure.

4.7. Multimodal Downgrading and Closing the Sequence after Compliance

We now turn to one of the interactional affordances of the practice: closing the pursuit sequence. An imperative-speaker may repeat the imperative even after the imperative-recipient has complied with a prior version. As discussed in Chapter 2, compliance – the preferred response – is actualized when an imperative-recipient performs the named action and advances the initiated CoA. The question arises why reissue an imperative if the directed action has already been performed. I argue that in a post-compliance position, the subsequent version functions as a sequence closing third. Although imperatives are typically associated with initiating actions, this section shows that they may be used in a third position to close the sequence and treat the coparticipant's compliance as adequate. A subsequent version, after compliance, is often accompanied with confirmatory gestures such as a downward nod in Persian or confirmatory tokens such as *âre* 'yes'. In a third position, after an interactant's compliance, these gestural and verbal actions are typically used to do acknowledgement. Their co-occurrence with the imperative suggests that the imperative is also a design feature of the acknowledging action. Therefore, a multimodally downgraded imperative post-compliance does not pursue a preferred response; rather it functions to close the imperative sequence, acknowledges the recipient's compliance, and displays the speaker's orientation toward the recipient's response as adequate. In this collection, a subsequent version after compliance is never observed to be designed with upgraded features.

The following extract is a case in point.[28] Amir (AMI) and Farnaz (NAZ) are guests at Faraneh's (FAR) house and are helping her to mix different kinds of rice. Faraneh (Amir's mother) should leave for work if any of her clients rings. If that happens, the activity of rice-mixing will come to a halt since Faraneh leads the activity. The imperative in focus is 'Mix' by which Faraneh directs Amir to start mixing the rice. Line 11 is analyzed to be locally the first since it is the first one issued after the rice which needs mixing is completely poured into the bowl. Note Faraneh's multimodally upgraded imperatives before Amir's compliance (lines 14, 16) and her multimodally downgraded version after his compliance, which comes with a confirmatory downward nod (line 18). This extract is another example of downgrade from an upgrade. The focus is on the multimodal downgrade, but the upgrades are also discussed to better contrast the environment in which upgrading versus downgrading occurs.

Extract 4.6. RAY020a, 04.30

```
01    FAR    +amir zud   bash
              name quick be.IMP.2SG
              Amir hurry
          >>+pours rice into a big bowl-->

02    FAR    alân moshtari zang mi-zan-e     *mi-g-e      biâ
              now client    ring PRG-hit-3SG PRG-say-3SG come.IMP.2SG
              a client will ring now and will ask me to go
       ami                                  *lean twrd rice bowl-->


03           (1.1)

04    AMI    âkh  jun
              intj life
              Hooray
```

```
05          (0.6)

06   FAR    mige        âkh  jun [ heheheh ((sniffs))
            PRG-say-3SG intj life
            He says hooray

07   NAZ                           [((laughs))

08          (0.2)

09   AMI    dusti  mus•ti  kesi  na-dâr-i    zang       be-zane
            friend nonword anyone NEG-have-2SG ring       SBJV-hit
            Don't you have a friend or someone to call
                      •gaze at NAZ-->

10   AMI    .h +[be-ge       khânum [faran\•
               SBJV-say      missus  name
            .h and say 'Ms. Faran\'

->   11  FAR   +[biâ *       sa\   [ da \•   ghâti kon
               Come.IMP.2SG  incmplte words\ mix   do.IMP.2SG
               Here sa\ da\ Mix


12   NAZ      +     *              [(( gi•ggles  ))
         far   -->+
         ami        -->* leans back-->
         ami                           •gaze at the bowl-->

13          +(0.4)
         far   +moves bowl toward AMI-->

->   14  FAR   <<f↑>ghâti ko[n+>
               mix   do.IMP.2SG
               Mix

15   AMI              •*[ha+m lâzem     nist       hatman    bere
                       PAR necessary NEG-be.3SG definitely go
                       She doesn't necessarily have to go
         far               -->+
         ami           -->*moves hand toward rice bags-->
         ami           -->•

            [posh\*
             behind
             behind
             -->*

->   16  FAR   +[<< ff↑↑>ghâti kon>+
                     mix   do.IMP.2SG
                     Mix
               +point toward bowl--+

17          *(0.4)
         ami   *leans forward and starts mixing-->>

=>   18  FAR   +<<pp,ll>ghâti kon>+
                     mix   do.IMP.2SG
                     Mix
               +downward nod------+
```

131

```
19          (0.4)

20    AMI   lâzem    nist    hatman    bere po:shte chiz:
            necessary NEG-be.3SG definitely go   behind stuff
            She doesn't necessarily have to go behind the um:

21    AMI   faghat hamin bege da\ man poshte dar-am
            only   this  say  da\ 1SG behind door-be/1SG
            If she can only say that I'm behind the door

            [badesh ke           mâmân raft=
            then    relative.word Mom   go
            then when Mom is gone

22    NAZ   [((laughs till the end of the extract))

23    AMI   =mâ mirim esterâhat mikonim
            1PL go    rest      do
            We'll go and rest
```

Before delving into the multimodal design features of the imperatives of interest, first an analysis of the sequence and the recipient's responsive actions is presented. There are two parallel sequences which compete with one another: a proposal, which is initiated by Amir and is addressed at Farnaz, and the imperative sequence, initiated by Faraneh, addressed at Amir. First the proposal sequence is launched by Amir after Faraneh directs him to speed up his engagement in the mixing activity (lines 1-2). Amir produces a disaffiliative responsive action to Faraneh's directive (line 4), which is treated as laughable by Faraneh herself and Farnaz (lines 6-7). He then initiates the proposal sequence, addressed at Farnaz – a humorous suggestion for Farnaz to plot a scenario to make Faraneh leave (lines 9-10). This proposal sequence continues to compete with Faraneh's imperative sequence. Before Amir brings his proposal TCU to a completion point, Faraneh launches the imperative sequence in overlap (line 11). Faraneh's persistence in producing the imperative despite the fact that Amir continues his turn shows that this imperative is already competing with Amir's proposal.

The first imperative in focus is issued as Faraneh balances the rice pile in the bowl, and hence can be seen to be preparing the rice to be mixed (line 11). This imperative receives a noncomplying response (embodiment in line 12). A complying response to Faraneh's imperative would have been for Amir to stretch his arm towards the rice bowl and then to start the manual action of mixing. This embodied action would have furthered Faraneh's CoA, but Amir does not perform either of these actions. Instead, he leans back, making more space between himself and the rice bowl.

Concurrent to the second imperative, Faraneh moves the rice bowl closer to Amir (lines 13-14). That is, Amir leans back, a display of non-compliance, and Faraneh moves the rice bowl toward him, pursuing compliance. Faraneh's embodiment can be analyzed as a facilitating movement since with Amir leaning back, the mixing activity may be relatively more physically demanding. By moving the bowl near Amir, Faraneh is in effect ruling out 'distance' as an account for non-compliance. In this way, Faraneh's embodied action is facilitating compliance, which makes noncompliance potentially more blameworthy.

Following this imperative, Amir produces actions that have implications for compliance; nevertheless, his actions do not project a completion of Faraneh's CoA (line 15). He moves his hand towards one of the rice bags. On the one hand, this projects his involvement in the physical activity of rice-mixing (although verbally he resumes his parallel CoA, i.e., suggesting what Farnaz's friend needs or does not need to do for the plot to be successful, line 15). On the other hand, for the pursued outcome of the imperative 'mix' to be obtained, his hand should be moving toward the rice bowl, not the rice bags. Hence although his hand movement projects some engagement in the activity, the movement is not in

the right direction for Faraneh's initiated CoA to be completed, so it does not project the completion of the imperative sequence. This bodily action, therefore, is less than preferred. Amir's projected physical involvement in the activity can be seen as less than full compliance. This is when Faraneh produces another subsequent version (line 16).

The third imperative co-occurs with the imperative-speaker's head point towards the rice bowl (line 16). Through her head-point, Faraneh points Amir, whose hand is reaching for the rice bags, towards the rice bowl. Thus, this imperative is oriented to Amir's response as not acceptable, and it is designed to correct Amir's method of compliance. This imperative is issued in interjacent overlap with Amir's turn in line 15 (Drew, 2009). For line 15 to be syntactically complete, it requires one more word to indicate what Farnaz's friends does not necessarily need to do, but for a pragmatic completion, it still requires at least another clause to communicate what Farnaz's friend needs to do. Amir's turn is pragmatically complete only at the end of line 23. Faraneh's third imperative (line 16) makes Amir abandon his turn (note the abrupt cut-off in the middle of his turn in line 15). This indicates Faraneh's persistence in forwarding her own CoA despite Amir's involvement in his own. Amir then leans towards the bowl and starts mixing rice and hence shows compliance (line 17): he satisfies the conditional relevance of the imperative by performing 'mixing' and furthering Faraneh's CoA. This time, Faraneh both wins the floor and secures Amir's compliance. Even though Amir is now engaged in Faraneh's CoA – mixing rice – Faraneh repeats 'Mix' (line 18), but in a multimodally downgraded design.

We now turn to the multimodal design of the four imperatives, three of which are issued before the recipient's compliance and one after his compliance. The subsequent versions before compliance are successively upgraded and the one after compliance is downgraded. Prosodically, the first imperative has an average intensity of 73 dB and a pitch span of 8 ST. (Speaker's overall average F_0 is 210Hz and the average bottom and top level of her pitch range are 90 and 600Hz, respectively.) The second rendition is prosodically upgraded: it is designed louder (76 dB), and it reaches a higher pitch ceiling, producing a wider pitch span (12 ST). It is also heard higher in the speaker's register. Finally, the third version is even more upgraded: it is produced louder (84 dB), and it reaches a higher pitch maximum and a lower minimum, producing a wider span (20 ST). The rather slight drop in the average F_0 of the TCU is due to the imperative's fall to a lower level at the end of the pitch contour. Nevertheless, since it reaches a higher maximum compared to the previous version, it is still heard higher in the speaker's register (Figure 4.17, Table 4.4).

Table 4.4. Summary of the multimodal features of the imperatives in Extract 4.6.

Line	Position	Average F_0 of the TCU (distance from overall F0)	Pitch Span (ST)	Intensity (dB)	Duration (ms)	Bodily Actions
11	Pre-compliance	293.7 Hz (5.8 ST)	8	73	600	Prepares rice
14	Pre-compliance	374.9 Hz (10.0 ST)	12	76	460	Pushes bowl
16	Pre-compliance	330.2 Hz (7.8 ST)	20	84	560	Head point
18	Post-compliance	173.7 Hz (3.2 ST)	1	55	370	Downward nod

The co-occurring bodily actions before compliance are also upgraded. The preparatory action (i.e., balancing the rice pile in the bowl) changes to a facilitating action during the production of the second imperative (i.e., reducing distance between the recipient and the bowl) and that changes to a correcting action with the production of the third (i.e., pointing the recipient to the direction of the bowl). The bodily actions co-occurring with the second and third are oriented to the recipient's non- or less than full compliance, and since the directed CoA is both facilitated and corrected for the recipient, noncompliance will be logically more blameworthy, and in this sense the embodied aspect is also upgraded.
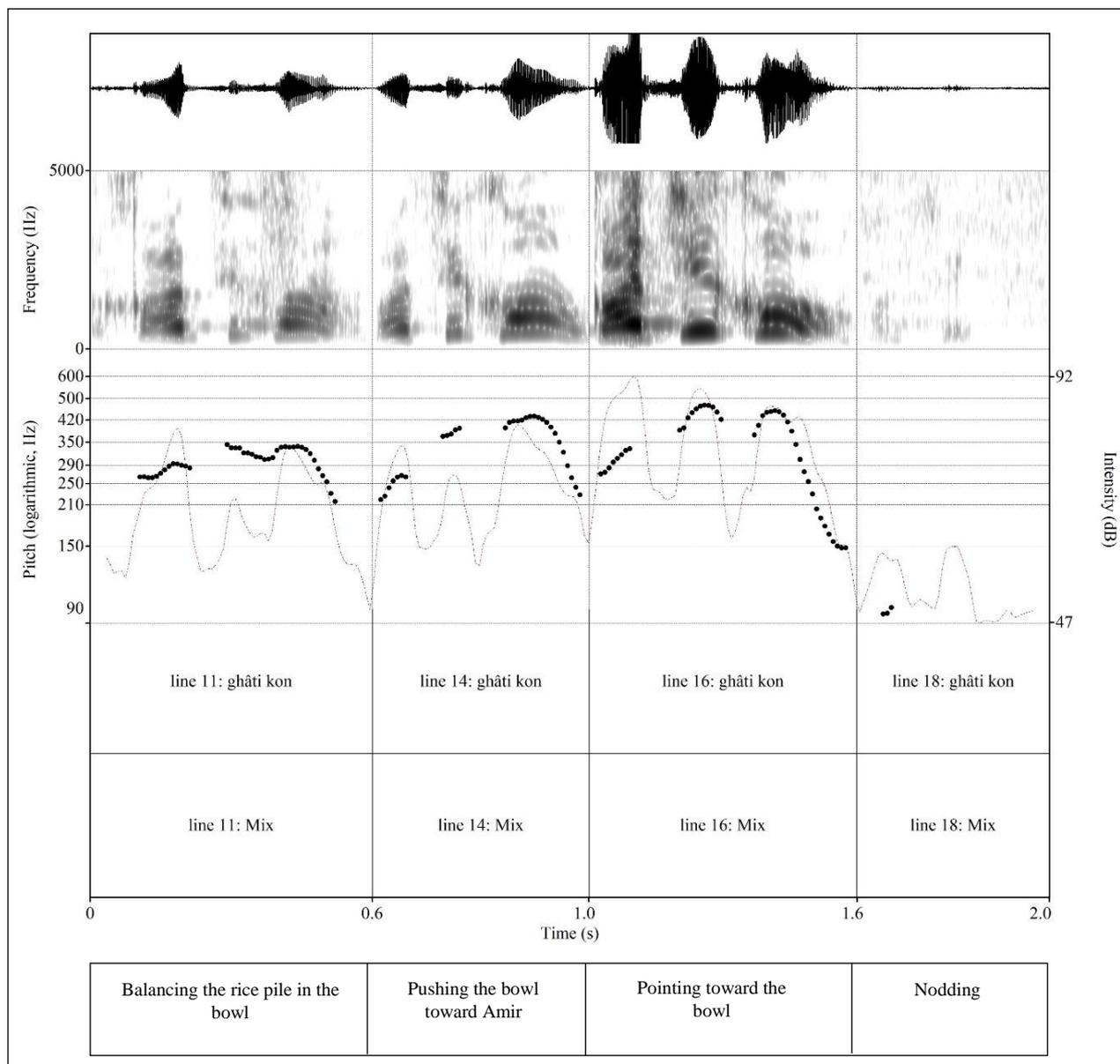
Figure 4.17. Multimodal upgrading versus downgrading in Extract 4.6. Multimodal downgrading comes after compliance.

The main analytic focus in this extract occurs after the recipient's compliance, when the imperative-speaker reissues the same imperative that she previously used to pursue compliance but with a multimodal downgrade design. The imperative is prosodically very quiet (55 dB), narrow (1 ST), and low in the speaker's pitch range. Regarding the bodily conduct, the imperative-speaker dispenses with the previous facilitating and correcting movements. Instead, she produces a downward nod as she issues this last version. In Persian, a downward nod is frequently used as a confirmation gesture. Whereas the embodied actions before compliance worked to hold the recipient accountable for his misconduct, the downward nod is oriented to the adequacy of the recipient's conduct. The co-occurrence of this gesture with the prosodically downgraded version is evidence for the claim that a downgraded imperative post-compliance is oriented to the adequacy of the recipient's response.

This last imperative (line 18) is produced in a third position, after a preferred response to the imperative that initiated a sequence. Note that the directive 'Mix', as an initiating action, is not relevant anymore since the recipient is already advancing the imperative-speaker's CoA. Hence the imperative must be

used to do something other than pursuing a preferred response. It is the combination of three factors that reveals its function: its sequential position – post-compliance where some acknowledgement or confirmation is potentially apt; its downgraded prosody; and the simultaneous gestural confirmation. These three convey that this last imperative is used to close the sequence. Thus, it is a sequence closing imperative even though the imperative grammatical format is typically associated with actions that usually initiate a sequence (e.g., requests). This re-establishes the independence of form and function in interaction (Walker, 2014).

4.8. Multimodal Downgrading and a Contingent Sequence Closure after Resistance

We saw that, after the recipient's compliance, a subsequently positioned imperative with downgraded multimodal design closes the imperative sequence. A downgraded subsequent imperative has a different interactional affordance when it is produced after the recipient's resistance against complying: it projects the imperative-speaker's disengagement from the imperative sequence and a contingent sequence closure. As explained in Chapter 2, resistance to complying is manifested in the recipient's dispreferred response – refusal to further the CoA projected by the imperative – or in responsive actions which embody less than a preferred response. The following sections show some examples.

4.8.1. Multimodal downgrading after a dispreferred response

In previous chapters, two ways by which an imperative-speaker may orient to the recipient's resistance were discussed: abandoning the sequence (i.e., straightforwardly withdrawing from the sequence without expanding the sequence) and pursuing a preferred response (Chapter 2). We saw that pursuit turns may be designed with upgraded features. When upgraded pursuit turns are accompanied with sanctioning practices, escalation in discord is projected (Chapter 3). We also briefly examined sequences in which pursuit turns project a contingent withdrawal. Here, we take a more detailed look at this third possibility: when the imperative-speaker pursues a preferred response but with downgraded subsequent imperatives. In this collection, the downgraded subsequent versions are rarely produced with sanctioning practices, and they project a contingent withdrawal from the initiated CoA. The imperative sequence is thus projected "as limited in extension" (Schegloff, 1998, p. 579), and withdrawal from the sequence is delayed (see Raymond & Lerner, 2014 on retarding in another context and Rossano, 2012 on gaze and sequence closure).

The following example comes from the same rice-mixing recording as in Extract 4.5 above. Faraneh's family are helping her to mix different kinds of rice. Sanam, Faraneh's daughter is sitting outside camera shot so that her image is not recorded. She has been mostly silent since the beginning of the recording.[29] Before this extract, Faraneh asked Amir, her son, to collect the rice grains, spilt around the mixing bowl. At the beginning of the extract, Amir is engaged with collecting rice grains with his gaze down. In lines 1-2, Sanam complains to Amir about him missing her bodily signals by which she has been quietly directing Amir to bring her a rice bag. Then, Faraneh, the imperative-speaker, addresses Amir with the first imperative in focus, *pâsho* 'get up'. After Amir's noncompliance, Faraneh issues the same imperative but multimodally downgraded. Facing Amir's further noncompliance, she abandons the sequence, gets up and passes Sanam a rice bag herself.

Extract 4.7. RAY020b, 05.02

```
01  SAN   amir hey         be-het imâ       eshâre mi-kon-am
          Name continuously to-2SG body.signal point  PRG-do-1SG
          Amir I keep pointing to you {I keep sending you bodily signals}
          >>collects rice from around the bowl-->>

02  SAN   mi-gam      yeki az in   berenj-â ro     bi-âr    injâ dige
          PRG-say-1SG one  of this rice-PL  OBJ.MAR IMP-bring here PAR
          I'm saying bring one of these rice {bags} here dige
```

---

[29] Earlier Sanam told her family that she would try not to talk since she is sitting outside the camera shot.

```
->   03  FAR  [*+<<ff↑↑>pâ+[sho::::>   +#
                    get.up.IMP
                    Get up
               *gaze at AMI-->
               +..........+taps on bag+
          fig                       #4.18
```


FAR taps on the rice bag.
FAR's gaze is on AMI.

FAR

AMI

Rice bags

#4.18
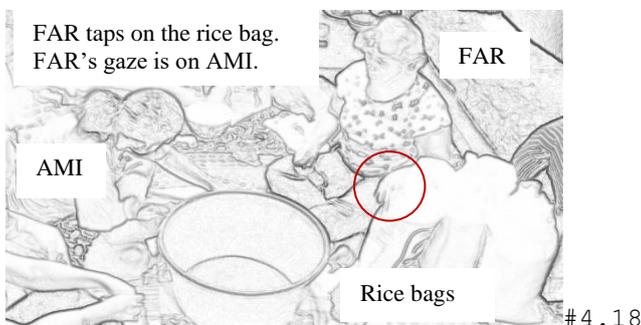
```
04  NAZ³⁰ [<<ff>he:   [heheh.h>

05  SAN          [<<giggles>ne-[*+mi-fahm-i?>
                              NEG-PRG-understand-2SG
                              Do you not get it?

06  FAR                      [*+ehhehe
                        -->* gaze at a rice bag
                        +............-->

07  FAR  [<<pp>ne-mi-fahm-e.>*+((sniff))#
              NEG-PRG-understand-3SG
          He doesn't get it.

08  AMI  [ehhehehe .h       *+          #
    far               -->*gaze at AMI-->
    far               -->+grips a rice bag handle-->
    fig                        #4.19
```


FAR grips the rice bag handle.

#4.19
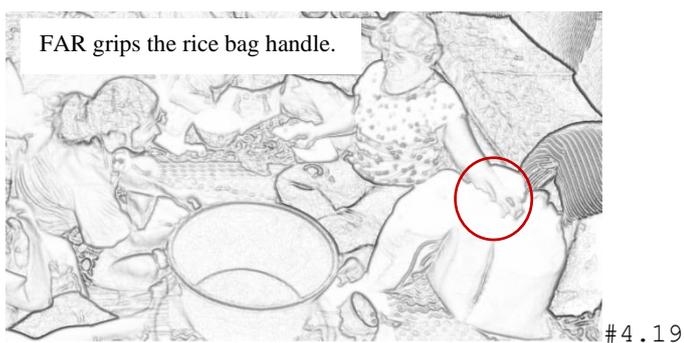
```
->   09  FAR  <<f↑>pâ[sho:>+
                  get.up.IMP
                  Get up
                     -->+

10  AMI          [mi-g-am*+    in   chi  mi-g-e
                  PRG-say-1SG  this what PRG-say-3SG
                  I was wondering what this {she=Sanam} is saying
```

---

³⁰ Farnaz (NAZ) is Amir's fiancée.

136

```
far                  -->*
far                    +shakes off rice dust-->


      ±injuri   ±injuri   ±<<gl>mi-[kon-e>+he
      this.way  this.way         PRG-do-3SG
      {by} acting like this
```

11 NAZ   **[hehehe+**

12 FAR   **[ehhehe+**
```
far                              -->+get up/give SAN bag-->>
```


#4.20

Before examining the design features of the imperative turns, the sequential positions where the imperatives are produced are analyzed. The first imperative (line 3) is issued after the recipient is treated accountable for missing Sanam's embodied directives (lines 1-2). Sanam's turn has lexical components which displays her orientation towards Amir's not passing a rice bag as sanctionable. Firstly, the adverb *hey* 'continuously', which is translated to the verb 'keep' in the translation line, shows that Sanam has been making embodied requests for a rice bag more than once (line 1). Secondly, the particle *dige* in Persian is sometimes used to treat a recipient accountable for an interactional failure (line 2). It is in this environment that the first 'Get up' is produced by Faraneh. The imperative in effect directs Amir to pass a rice bag to Sanam. This is because the action of 'getting up' is the first movement which Amir should make if he is to pass a rice bag to Sanam. In this context, directing Amir to perform the first step of the action implies doing the whole action of passing Sanam a rice bag.

The second imperative is produced after the recipient's display of noncompliance. A complying response – one that advances Faraneh's CoA – would have been for Amir to halt his ongoing CoA of collecting rice and to get up, but he continues to collect rice, showing no orientation to getting up. His engagement in his own CoA in the slot where the directed action was normatively expected to be performed (lines 4-8) is analyzed as a dispreferred next action. After everyone treats Amir's lack of understanding of Sanam's embodied directives as laughable (lines 4-8), Faraneh pursues a preferred response by repeating the same imperative (line 9). Again, this imperative does not receive a complying response. Nevertheless, Faraneh does not re-issue the imperative. Instead, she herself gets up and passes Sanam a bag of rice (lines 10-12, Figure 4.20).

The subsequent imperative is multimodally downgraded relative to the first (Figure 4.21; Table 4.5). In terms of prosodic design, the first imperative (line 3) is produced loud (83 dB), high in the speaker's pitch range (overall 13 ST above her average f0), and with an elongated second vowel. (Speaker's overall average f0 is 210Hz and the average floor and ceiling are 90 and 600Hz, respectively.) The rather loud, high, and long imperative is oriented to Amir's accountability for not having responded to Sanam's embodied requests. The second imperative (line 9) is shorter in duration, lower in intensity (79 dB) and pitch range (overall 9 ST above her average f0).

Table 4.5. Summary of the multimodal features of the imperatives in Extract 4.7.

| Line	Average f0 of the rejection TCUs (Hz)	Intensity (dB)	Span (ST)	Duration (ms)	Embodiment
3	414 (11.0 ST above speaker's average f0)	83	13	530	Tapping on bag
9	360 (9.0 ST above speaker's average f0)	79	13	400	No tapping
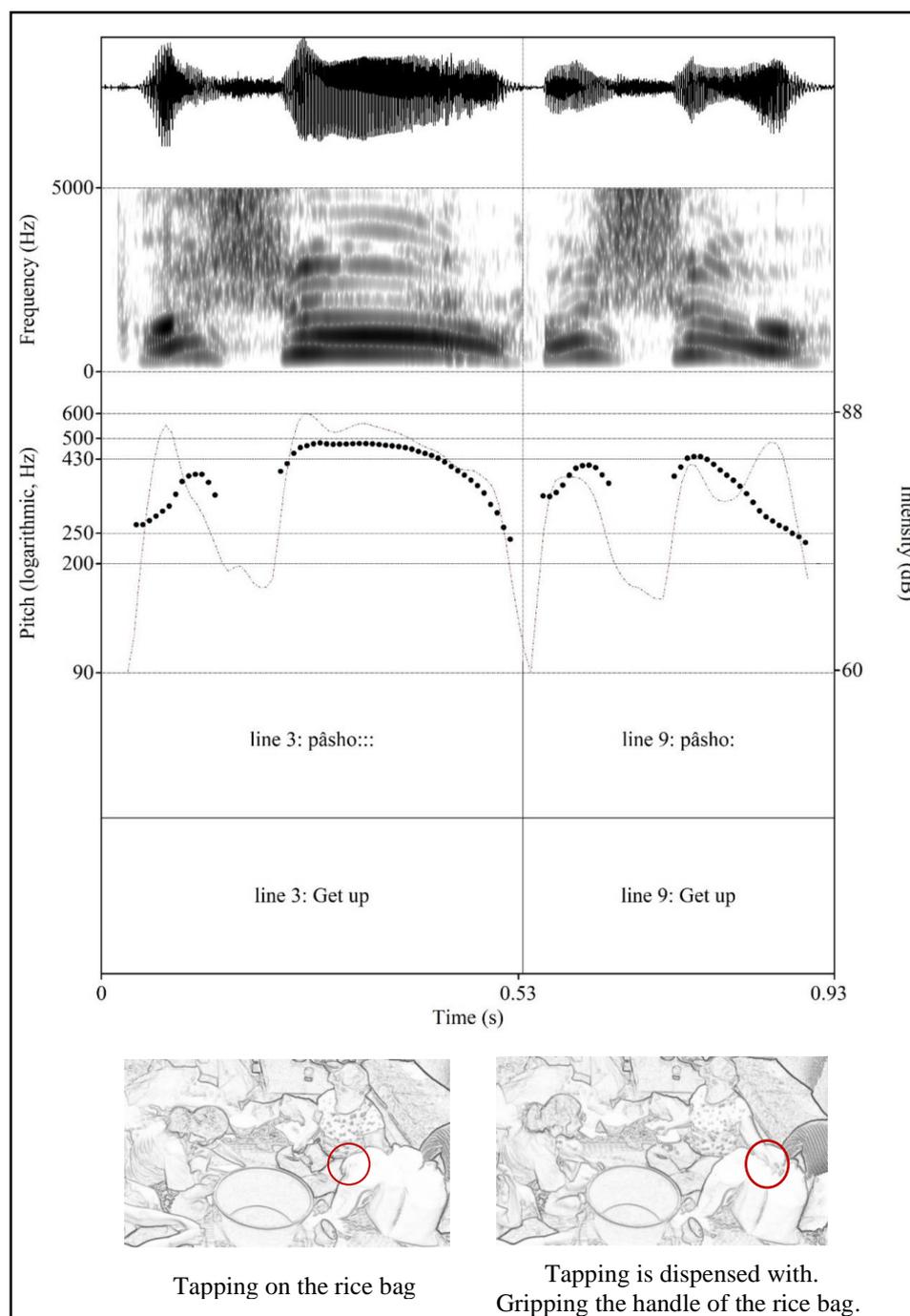
Figure 4.21. *Multimodal downgrading in Extract 4.7.*

The bodily actions also differ. Simultaneously with the first imperative, Faraneh taps on one of the rice bags, which should be passed to Sanam, twice (Figure 4.18). In the second version, the tapping action is dispensed with; rather, Faraneh grips the handle of one of the rice bags as she slightly straightens her

138

back (Figure 4.19). These are embodied actions which would be relevant if she herself is to perform the directed action. Thus Faraneh's bodily conduct co-occurring with the second imperative already projects her own getting up (straightening back) and passing the rice bag (gripping the bag's handle).

Through the downgraded second imperative, the imperative-speaker incrementally disengages from the sequence. The incremental process is especially clear in the participant's bodily conduct. As mentioned earlier, with the first imperative, a tapping action is performed. Tapping in this context can be analyzed as a kind of deictic gesture, one that refers to the rice bag that should be passed by Amir. Besides the imperative, which is addressed at Amir, the deictic tapping shows that, at this stage, Faraneh is treating *Amir* as the performer of passing the rice bag. With the second imperative, on the other hand, hand-gripping and back-straightening actions are performed. Even though the imperative is still addressed at Amir, these two bodily actions already show a shift in Faraneh's orientations about the agent who is going to perform the passing of the rice bag. Indeed, after the imperative, Faraneh shakes the rice dust off her clothes – a further preparation to the action of getting up – and she herself gets up and passes the rice bag. Although multimodally downgraded subsequent versions still pursue a preferred response, they project the pursuing participant's withdrawal from the sequence and contribute to the speaker's gradual disengagement.

In the following, an example of downgrading after a less than preferred response is analyzed.

4.8.2. Multimodal downgrading after less than preferred responses

As illustrated in Chapter 2, when an imperative form is used to get a coparticipant to do something, it can receive a continuum of responses with noncompliance at one end and compliance at the other. Compliance is a preferred response: it completes the imperative sequence by advancing the initiated CoA. Noncompliance is the dispreferred alternative; it hinders the progression of the initiated CoA. There are a range of responses between these two extremes and these middle ground responses can be termed as less than complying or in other words less than preferred. Anything but 'fully' complying responses have some degree of dispreferred features. In Chapter 3, we saw that after such responses, an imperative-speaker may pursue a preferred response through upgraded subsequent imperatives, which contribute to persisting and project an escalation in disaffiliation. This section shows that a contingent withdrawal from the sequence, achieved through downgraded design features, is another possibility after these responses. In this way, the speaker disengages from the sequence through a process.

Here is an example. Melisa and Minuvash are playing a card game (Figure 4.22). In each turn, each player takes two cards from the card deck and plays three cards. It is now Minuvash's turn. She has only played one of her three moves when Melisa reaches to take cards from the card deck. This is when Minuvash produces the first imperative (line 5, 'Wait'). Melisa slightly raises her hand and breaks contact with the card deck, displaying compliance, but at the same time she continues to hold her hand above the card deck, which shows orientation to continuing her own CoA. Minuvash then reissues the same imperative but in downgraded multimodal design (line 9).

Figure 4.22. Participants in Extract 4.8. Melisa (left) and Minuvash (right) are playing a card game.

```
01   MEL   khob. ki-e?
           OK    who-be.3SG
           OK.   Whose turn is it?

02         (0.1)

03   MIN   ((lip smack)) .hhhhhhh n::obate:: +ban::::da-s::::.hhhh+
                                    turn-of      slave-be.3SG
                         .hhhhhhh it's::: m::::y t::urn::::.hhhhh
                                                  +puts a card down---+

04         *         (0.5)#
     mel   *moves hand twrds card deck-->
     fig               #4.23
```


#4.23

```
->   05   MIN   ((lip smack))   +sa*b+    kon#
                                 patience do.IMP.2SG
                                 wait
          mel                  -->*touches card deck-->
                               +....+clenches MEL's wrist-->
          fig                              #4.24
```


#4.24

```
06         (0.3)

07   MIN   yek::*harkat+ raft-am *+#.
           one   move    go.PST-1SG
           I played one card.
     mel    -->*raises hand-----*
     min            -->+unclenches+
     fig                      #4.25
```

#4.25

```
08          *+          (0.3)          *+
      mel   *hold hnd above card deck*
      min    +hold hnd above MEL's hnd+


->    09   MIN   +*#<<pp,↓,dim> sab      kon#      sab   +  kon>      +#
                              patience do.IMP.2SG patience do.IMP.2SG
                  Wait Wait
            mel   *moves hand back-->
            min   +moves hand parallel to MEL's-----------+retracts hnd+
            fig    #4.26                    #4.27                    #4.28
```


#4.26

#4.27

#4.28

```
10            *+(1.0)#
      mel   -->*hand on lap-->>
      min       +hand fully retracted-->>
      fig          #4.29
```


#4.29

The imperative-recipient (Melisa) is about to perform an 'illegal' action and the first imperative is issued to pre-empt the projected violation before it occurs. In line 1, Melisa requests information on whose turn it is to play. Soon after Minuvash responds that it is her turn and plays one card (line 3), Melisa reaches for the card deck (line 4, Figure 4.23). This projects a CoA in which she takes cards from the deck and shows Melisa's understanding of Minuvash's turn as over. Melisa's hand stretch towards the card deck is the 'illegal' action since Minuvash has played only one of her three moves. Minuvash issues the first imperative of interest, 'Wait' (line 5). Concurrently, she moves her hand towards Melisa's in an accelerated design (see Raymond & Lerner, 2014 on accelerating). By the time Minuvash's hand reaches Melisa's, Melisa's hand has already come in touch with the card deck. Minuvash clenches Melisa's wrist as she issues the first imperative (line 5-6, Figure 4.24) and then accounts for her imperative (line 7).

The subsequent imperative is issued after the recipient's not 'full' compliance. If Melisa is to 'fully' comply, her next action must be to (a) break contact with the card deck, and (b) retract her hand. She raises her hand and thus breaks contact with the cards (line 7, Figure 4.25), but she does not retract it; rather, she holds in a "ready-to-resume position" – above the card deck – which makes her appear ready to touch the cards again (line 8; Raymond & Lerner, 2014; p. 238). By holding her hand in that unstable position, she shows orientation to her 'illegal' CoA as still ongoing (Schegloff, 1998). Thus she is not fully compliant and her response is less than preferred. Minuvash's responsive actions reveal her orientation towards Melisa's actions as complying but not fully: after Melisa stops touching the card deck, Minuvash unclenches Melisa's wrist (line 7). Thereby, she displays her understanding of Melisa's bodily movement as heading towards compliance (see Kent, 2012 on incipient compliance). However, Minuvash does not retract her hand. She, too, holds her hand above Melisa's hand for as long as Melisa holds her hand above the card deck (line 8). The position of Minuvash's hand indicates her preparedness to re-enforce the rule if needed. This sustained embodied configuration in that unstable position from her, too, shows her orientation to the imperative sequence as not yet closed (Raymond & Lerner, 2014; Schegloff, 1998). It is in this context that the subsequent imperative is issued (line 9). Concurrent with the subsequent version, Melisa retracts her hand towards home position (Figure 4.26). Minuvash first moves her hand in parallel with Melisa's, which appears to be an 'ushering' action (Figure 4.27). When Melisa's hand almost reaches the home position, Minuvash also retracts her hand (Figures 4.28 and 4.29).

Compared to the first imperative, the subsequent version is designed with downgraded multimodal features (line 5 vs. line 9; Figure 4.26; Table 4.6). Prosodically, the first version is produced with an average intensity of 71 dB, and it starts rather high in the speaker's pitch range (5.5 ST above her average f0 level. The subsequent version is in the format of double imperatives *sab kon s\ * 'Wait W\', which progressively gets lower and lower in intensity such that, of the second *sab kon* 'wait', only the fricative is hearable. This makes the second version be heard as truncated: the production of the second imperative turn is abandoned before it reaches a completion. The overall intensity of the subsequent version is 62 dB, which is greatly reduced compared to the prior version and it starts only around the speaker's average f0 (at 290 Hz). The pitch span is also reduced from 9.4 ST in the first to 5.0 ST in the second.

Table 4.6. Summary of the multimodal features of the imperative turns in Extract 4.8.

Line	Average f0 of the rejection TCUs (Hz)	Intensity (dB)	Span (ST)	Duration (ms)	Embodiment
5	400 (5.5 ST above speaker's average f0)	71	9.4	300	Clenching
9	290 (at the speaker's average f0)	62	5.0	450	Ushering

The double format in the subsequent version is not hearable as an upgrade feature due to the lexically truncated and prosodically barely hearable design of its second unit. Rather, the double format seems to be oriented to the duration and the process that it takes for Melisa to fully comply: the subsequent version is issued as the recipient starts to move her hand back. Almost simultaneous with the second imperative unit in the subsequent version, Melisa retracts hand and then Minuvash abandons the production of the second unit. Thus the imperative-speaker continues the production of her imperatives until there is observable evidence that the recipient is fully complying. The double format is oriented to the 'processual' nature of the recipient's compliance rather than an upgrade in directing.
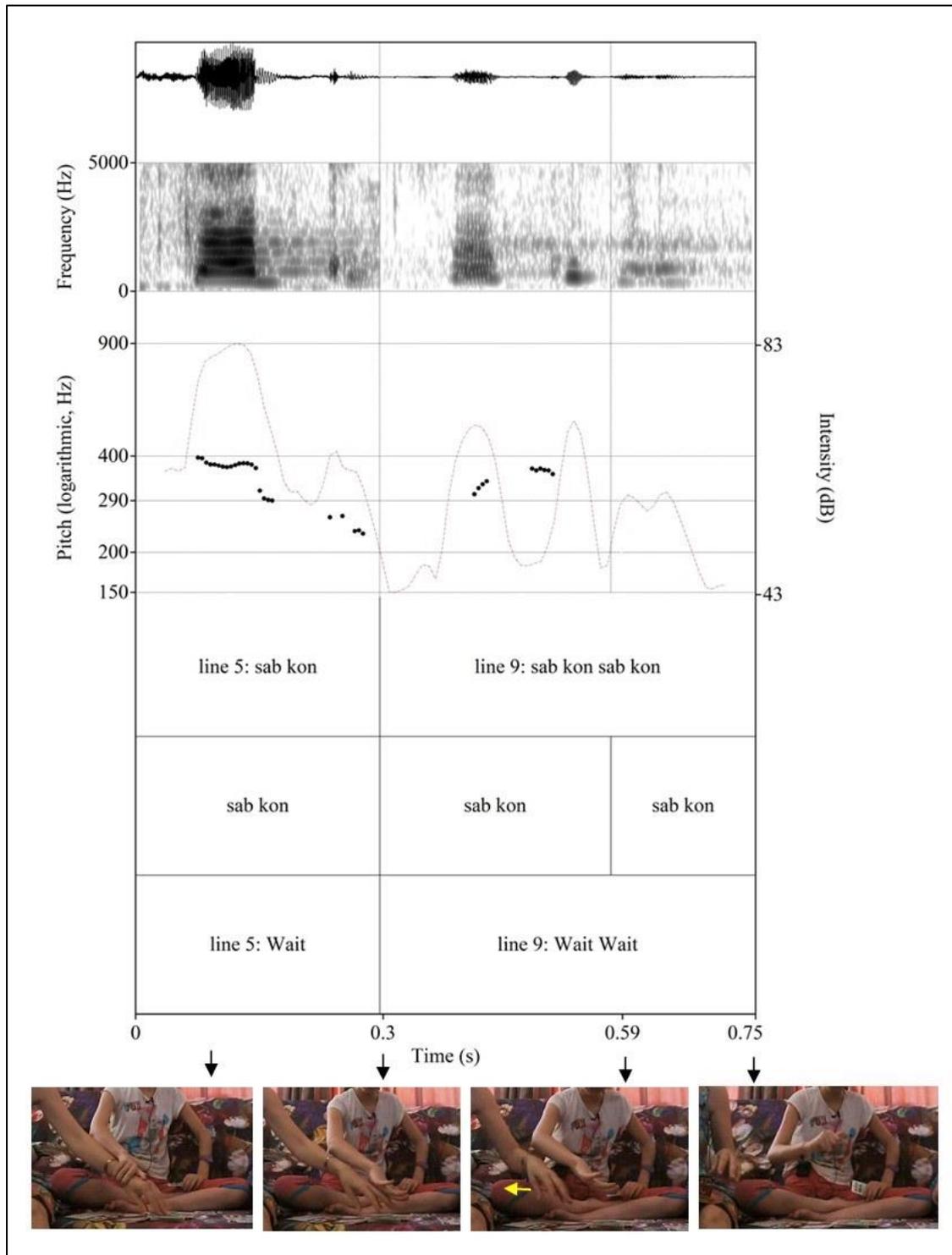
Figure 4.26. Multimodal downgrading in Extract 4.8. Co-occurring with the original imperative, the imperative-speaker's hand touches the recipient's wrist in a clenching movement. After the recipient's less than preferred response, the imperative-speaker breaks contact with the recipient's hand. As the recipient displays orientations to producing a preferred response, the imperative-speaker allows more and more space between her hand and the recipient's wrist.

The imperative-speaker's bodily conduct is also downgraded. The embodied downgrade is actualized through tactile and proximity features. There are five stages involved: full contact, break contact, hold hand near the recipient, increase distance, full retraction. Firstly, the first imperative is issued as the imperative-speaker clenches the recipient's hand; therefore, the speaker has made physical contact with the recipient to pre-empt the recipient's projected action (Figure 4.20). Secondly, she unclenches and breaks physical contact with the recipient's hand as the recipient shows some degree of compliance (Figure 4.21). Thirdly, she holds her hand above and close to the recipient's hand when the recipient's bodily conduct projects potential noncompliance (Figure 4.22). Fourthly, co-occurring with the subsequent version, she performs an 'ushering' movement by which she appears to guide the recipient's hand away from the card deck (Figure 4.23). As the recipient retracts her hand, she allows more and more space between her hand and the recipient's (Figure 4.24). Finally, she completely retracts her hand (Figures 4.25).

Like downgrading after dispreferred responses, the practice after less than preferred responses projects a contingent sequence closure and contributes to the speaker's gradual backing down from the sequence. Again, the incremental nature of the speaker's withdrawal is evident in her embodied conduct. For example, the speaker starts with clenching, which is a relatively 'strong' means for making the recipient perform the directed action and after going through a hold stage, the clench turns into 'ushering', which is a relatively 'weaker' way of making the recipient perform the directed action. By reissuing an imperative with downgraded features after a less than preferred response, the speaker makes it public that there is still work needed to get done to achieve 'full' compliance. And again there are no sanctioning practices observed. This negative observation is another piece of evidence that the speaker is not holding the recipient responsible for their less than preferred response and thus the sequence may as well be abandoned.

4.9. Interactional Affordances of Multimodal Upgrades Compared to Downgrades

Having seen some interactional affordances of downgrading the multimodal design of a subsequent imperative, we can now draw on the findings of Chapter 3 to compare the interactional affordance of an upgraded pursuit turn with a downgraded one in two main sequential positions: after compliance and after resistance (Table 4.7).

Table 4.7. Different interactional affordances of upgrading versus downgrading a pursuit turn.

Design / Position	Multimodal Upgrade	Multimodal Downgrade
After Compliance	Ø (not observed in the data)	Closes the sequence
After Resistance	- Projects more sequence expansions. - Orientations to persistence. - If it is accompanied with negative sanctioning practices, it shapes and projects escalation.	- Projects a contingent sequence closure. - Orientations to desisting and withdrawal. - If it occurs after escalation, it contributes to and projects de-escalation.

After compliance: Multimodal upgrading is never observed after a recipient's compliance. If a subsequent version is produced after the recipient's compliance, it is always downgraded. In a post-compliance position, the first pair part has already received a preferred second pair part; the sequence is already completed with a preferred response and, therefore, pursuing is not a relevant action anymore. A subsequent version in this position does not make any response relevant; it is not an initiating action. Rather, it is a SCT, by which the imperative-speaker explicitly closes the sequence. It is often

accompanied with verbal or gestural productions that do acknowledging. By virtue of its co-occurrence with these acknowledgment devices, a SCT imperative also contributes to acknowledging the recipient's response as adequate.

After resistance: After resistance, both upgrading and downgrading are observed. Multimodal upgrading occurs often when there are orientations to sequence expansion and forwarding the CoA projected by the imperative despite the recipient's resistance. With multimodal downgrading, on the other hand, there are orientations to ending the sequence and a contingent sequence closure is projected. The imperative-speaker is seen to be withdrawing from the sequence in a gradual process and so withdrawal is delayed.

These orientations and projections are evidenced by the co-occurrence of practices that negatively sanction the recipient's resistance with multimodal upgrades and the scarce co-occurrence of such practices with multimodal downgrades. In Chapter 3, some sanctioning practices from the pool of resources that persisting participants use were shown. These practices often accompany upgraded subsequent versions to hold the recipients at fault for their resistance (rather than downgraded versions). To name only a few from the corpus: verbally, addressing the resisting party with negative assessment terms (e.g., you moron); phonetically, Persian vocalizations such as elongated [tʃ] or [ʃ]; and also embodiments such as frowning, fixated gaze, and what I refer to as 'mock aggression'. I will dedicate Chapter 6 to investigating this last resource, recurrently observed with multimodal upgrading in persistence sequences. In this collection, these practices rarely co-occur with multimodal downgrades. This negative observation can imply that imperative-speakers, themselves, rarely treat their recipient's resistance as misconduct. Thus, they may withdraw from the sequence. In short, upgraded subsequent versions when co-occurring with sanctioning practices project escalation in disaffiliation whereas downgraded ones project a contingent withdrawal.

4.10. Discussion

In this chapter, the phenomenon multimodal downgrading and some of its interactional affordances in pursuit sequences were discussed. The findings contribute to four research areas in CA: conflict management (e.g., Antaki, 1994), action formation (e.g., Levinson, 2013), sequence structure (e.g., Schegloff, 2007), and preference organization (e.g., Clayman, 2002; Heritage, 1984; Pomerantz & Heritage, 2013; Pillet-Shore, 2017). It was reconfirmed that disengagement from a sequence can occur gradually and in a process. In other words, to be involved in an interaction does not have a categorical nature. There is a continuum of engagement and participants have cues to rank their involvement on the continuum (e.g., Raymond & Lerner, 2004; Schegloff, 1998). Goffman (1981) introduced the concept of 'participation framework' and argued that a *speaker-hearer* relationship in interaction is too naïve; rather, that there are different roles that speakers or hearers can assume. Later, Goodwin (2004) criticized such a static view to participants and showed that "[s]peakers attend to hearers as active coparticipants and systematically modify their talk as it is emerging so as to take into account what their hearers are doing" (p. 223). This work re-established the fact that participants may dismantle a participation framework in pieces. They employ cues to indicate how engaged they are and where the interaction is heading towards, escalation or withdrawal. This has implications for conflict management: multimodal downgrading can be an action design used to exit a conflictual encounter.

Regarding action formation, a subsequent imperative is produced after the recipient's response or sometimes its noticeable lack thereof. Adjusting the multimodal design of a subsequent imperative can display the imperative-speaker's fine orientations towards the recipient's (non)response (Raymond & Lerner, 2014; Schegloff, 1998), and it also has implications for action formation: the same imperative format, which has been just used to get a coparticipant to do something, can be re-used to acknowledge the interactant's response when it is produced after compliance and its multimodal design features are changed.

As for sequence structure, we saw that after the recipient's resistance, downgrading the multimodal design of a pursuit turn projects a contingent closure and the pursuing party's potential withdrawal from the sequence. CA has documented some verbal, prosodic, and embodied resources that project a forthcoming sequence closure. Sequence closing thirds or pre-closing tokens such as OK in English are among verbal resources (Schegloff, 2007; Schegloff & Sacks, 1973). Some phoneticians of talk in interaction have shown some prosodic patterns that project the completion of turns or sequences (e.g., Couper-Kuhlen, 2004; Curl, Local & Walker, 2006; Local & Walker, 2005; Ogden, 2010; Persson, 2013; and Walker, 2012). Research focused on embodied behaviour has also shown some bodily resources that signal a contingent sequence closure, for example, gaze withdrawal (Rossano, 2012) and body torque (Schegloff, 1998). This chapter firstly added one more verbal resource for closing down an imperative sequence: a post-compliance subsequent imperative itself when designed with downgraded features. Secondly, we saw that these various resources often shown separately in different investigations can in fact belong to a multimodal Gestalt with each resource being but one constituent of a larger Gestalt (Mondada, 2104). The recurrent co-occurrence of bodily features such as gaze withdrawal and releasing torque with prosodic and verbal downgrades suggest that these bodily features can be a part of a downgraded Gestalt. This is evidenced by what these resources collectively achieve – backing down as opposed to persisting. Needless to say, the function of each of these resources is context dependent. For example, an imperative-*recipient*'s gaze aversion can be a sign of *upgrading* resistance to performing the directed action.

We saw that actions are often produced in a process and the processual nature of action production has implications for the preference organization. Organization of preference in CA has considered a rejection in response to a request as a dispreferred action and acceptance as the preferred alternative. But these responsive actions can be done with a variety of constructional features. Later contributions such as Schegloff (1988) and Kendrick and Torreria's (2015) distinguished between preference when it is used to analyze actions (e.g., in response to a request, rejection is typically the dispreferred action while granting is the preferred alternative) and when it is used to analyze turn design (e.g., hesitation particles and delays contribute to a dispreferred format). CA has also considered contradictory design features of the actions. For example, "yes, but…" type of responses are considered as only 'pro-forma' acceptance. But ultimately in analyzing an *action* (not the design) of a turn as preferred to dispreferred, it seems that it traditionally relies on the ultimate outcome of the turn; that is, if the turn is in the service of rejection, the action is analyzed as dispreferred and vice versa. But this kind of approach may dismiss the fact that many times the performance of a social action occurs in a process that involves multiple steps for an ultimate recognizable action to be conceived. Participants themselves may orient to how these constituent steps may differ in terms of their preference status.

Consider those sequences in which the recipient produces a less-than-preferred response. In these sequences, a second pair part is performed through a combination of movements or actions. Some of these actions or movements project the completion of the CoA initiated by the imperative while some others project the sequence to be left incomplete. For example, in Extract 4.5, a case was shown in which an imperative-speaker directs the recipient to mix rice. A preferred second pair part would require the recipient to (a) raise his hand, (b) stretch his arm towards the rice bowl, and (c) mix the rice (Extract 4.7). The participant raises his hand, and this projects a preferred response since he begins to display engagement in the activity of rice-mixing. However, he then stretches his arm towards a wrong direction, and this projects a dispreferred response. Another example was presented in Extract 4.8 in which a preferred response required the recipient to (a) break contact with game cards, (b) raise her hand, and (c) move it back to home position. The imperative-recipient broke contact with the cards and raised her hand. These movements project completion of the imperative CoA. But then she did not retract her hand which projects that she may potentially touch the cards again.

In these examples, it takes a process for a preferred response to be produced, and at each stage within the process of the formation of the ultimate action, the participant may perform actions with different

valences in terms of preference. (These embodied examples may be analogous with 'yes, but…' responses. See Kendrick & Torreira, 2015 on qualified responses, and Robinson, 2020 on conditional response). From the emic perspective of the participants, even though these stages are connected, participants orient to them bit by bit and coordinate their responses accordingly. For example, in Extract 4.8, we saw that after the recipient's breaking contact with the cards, the imperative-speaker unclenched her wrist, but when the recipient held her hand above the card deck, she too held her hand in parallel, orienting to the fact that a 'fully' preferred response is not yet produced. Although CA considers the complex design features of a turn, when it comes to analyzing *action* as preferred or dispreferred, it has been so far mainly focused on the *ultimate outcome* of the action. However, this approach may not work for all social actions. As we saw coparticipants orient to and align their behavior with each step as their interactant advances in the process of producing an action.

Some studies on up- or downgraded turn design have considered another category where the design features of a turn remain virtually the same relative to a prior turn (e.g., Pomerantz, 1984b; Ogden, 2006). Ogden (2006), for example, argues that second assessments with similar prosodic features as the first behave similar to prosodically downgraded second assessments in that they both embody a weak agreement with the first assessor. In this regard, two observations are worthy of mention about the sequences in focus in this thesis. Firstly, a post-resistance subsequent imperative with virtually the same multimodal design features is rare in this collection. In a couple of cases where they do occur, there are orientations towards either problems with understanding/hearing or persisting. Repeating an imperative with virtually the same prosodic, embodied, and verbal features can treat the recipient's resistance as not acknowledgeable. So even though it is subsequently-positioned, the subsequent rendition has the design features of the first imperative copied in it, as if it is the first time it is being issued. This functions to disregard the recipient's response, and thus such design can have implications for persistence. Therefore, whereas in other studies, 'same' and downgraded' design have been grouped together, for the sequences in this study, grouping 'same' with 'upgraded' design may be a possibility since both contribute to persisting. Due to their rare occurrence in this collection, they were not analyzed, but future research can explore these subsequent versions in more depth.

Up to this point in the thesis, we have examined the structure of a pursuit sequence (Chapter 2). Based on the design features of a pursuit turn, two types of pursuit were distinguished (persisting and gradually desisting). In Chapter 3, we saw that multimodal upgrading is more associated with persisting and sequence expansions, in this chapter, we learnt that multimodal downgrading can be projective of a contingent withdrawal. In the following chapter, we explore pursuit turns and multimodal upgrading and downgrading in terms of the pursuing party's agentive stance

Chapter 5. Multimodal Gradation and Agency

5.1. Introduction

When the pursuits of two parties collide, one key concern is who 'gets her way' and who 'gives way'. Who gets the mobile phone that is fought over and how does she get it? Who sits in the coveted spot in front of the TV and how does he get to sit there? Who finally has to leave the comfort of the sofa to turn off a leaking tap and how is he made to do so? What is at stake is the question of *agency*: whose words prevail and, equally importantly, what do the two parties do to 'fight' their way out. This chapter tackles agency in pursuit sequences. We shall have a closer look at the design of a pursuit turn in light of the agentive stance of the pursuing party. We continue to focus on multimodal gradation and see the implications of various design features of a pursuit turn for the agentive stance of a pursuing party.

Let us cast a brief look at an initial example which only draws a general sketch of how the concept of agency is used in this thesis. Melisa and Nasima (mother-daughter) are having a conversation about some Eastern and Western cultures as Nasima's hand is on the family's cat. The cat starts to move forward on the sofa, projecting a forthcoming jump off the sofa, but Nasima grips the cat firmer, slowing the progressivity of its jumping CoA. This makes the cat appear as though it is 'struggling' to move forward. Witnessing this, Melisa issues the imperative 'let it go' once, but Nasima's hand continues to grip the cat. Melisa then reissues the imperative and pulls Nasima's hand away.

Extract 5.1. RAY014c, 12.37

```
01   NAS   ghashang     *+yek (0.2)  •+ ((cough))
           nice            one
           totally one westernized person
     cat               *moves forward-->
     mel                              •gaze at cat-->>
     nas   >>hand on cat+moves frwrd+grip the cat-->

           gharbzade  be [tamâm  m[anâ
           westernized to whole meaning
           in every sense of the word

02   MEL                   [±mâmân.            ±
                             Mom
                            sweetie.
                           ±reaches NAS's arm±

->   03   MEL                   [±vel-esh   kon.±#
                                 free-3SG  do.IMP.2SG
                                 let it go.
                                ±hold NAS's arm±
     fig                                          #5.1
```

#5.1

```
->    04    MEL      ±<<f,h>vel-esh +kon;> *=#
                         free-3SG do.IMP.2SG
                         let it go.
                     ±pull NAS's arm down-->
           nas                 -->+
           cat                           -->*
           fig                                  #5.2
```


#5.2

```
      05    MEL      =*mi-khâd  *± be-r-e±
                     PRG-want.3SG SBJV-go-3SG
                     it wants to go
                           -->±,,,,,,,±
           cat    *jumps dwn*
```

Here the two colliding CoAs are freeing the cat (Melisa's CoA) and holding onto it (Nasima's CoA) (Melisa can be seen to be acting on behalf of the cat, so another CoA is the cat's, but for practical purposes we only consider Melisa and Nasima's CoAs). With the first imperative, Melisa provides Nasima a slot to perform a responsive action, two of which are compliance and noncompliance (line 3). In the response slot, Nasima performs bodily movements that are indicative of noncompliance (the dipreferred option). At this point, Melisa has two main choices: to abandon the sequence or to pursue compliance. This chapter shows that these two options have different implications for the pursuing party's agentive stance. Whereas abandoning tacitly lets Nasima's CoA prevail, pursuing is in effect another attempt by Melisa to further her own CoA. In the clash over freeing against holding the cat, Melisa takes a next turn to pursue compliance and, therefore, she makes another attempt or performs an additional action to achieve her own interactional goal. Performing actions to further one's CoA displays an interactant's agentive stance. The stance manifests itself in the agent's observable conduct, and it is upgraded when *more* actions are performed to achieve the pursued interactional outcomes. Therefore, by pursuing compliance Melisa displays upgraded agentive stance than if she were to abandon the sequence.

It is not only her pursuing action that embodies her upgraded agentive stance. The design of her pursuit also contributes to it. Even though on the verbal level, her subsequent imperative is offering Nasima another responsive slot, her bodily action leaves a diminished chance for Nasima's independent response: Melisa pulls Nasima's arm away and consequently weakens her grip on the cat. She can be seen to be 'forcing' Nasima to let go of the cat. Melisa is reducing Nasima's response options (broadly

speaking compliance and noncompliance) to only one option (compliance) unless Nasima herself upgrades her resistance, which does not happen in this sequence.

In this chapter, we focus on the pursuing party's agentive stance. More specifically, we inspect how an imperative-speaker may manipulate various multimodal resources of her imperative turns to further her CoA despite the recipient's resistance. But first, a review of agency and stance is helpful as the analysis is in part based on how agency and stance are conceptualized in the literature. So in the following, first agency and stance are reviewed (§5.2 and §5.3). Then we inspect an imperative-speaker's agentive stance in three main situations: when she abandons the sequence after resistance (§5.4.1); when she pursues compliance using downgraded Gestalts (§5.4.2); and when she pursues compliance through upgraded Gestalts. Based on the analysis in these sections, a continuum of agency is provided in §5.5. Then §5.6 concludes the chapter with a discussion on the relationship between agency and multimodal gradation in light of Goffman's (1981) footing model (§5.6).

5.2.   Agency

Agency is one of those concepts which has attracted the attention of scholars of remarkably various fields, from disciplines within social science and beyond (see Ahearn, 2000; 2001, for a review). It is, therefore, no surprise that the concept has been defined variously not only across disciplines but also within the same field, depending on the scholars' foci. In most general terms, agency can be understood as "the relatively flexible wielding of means toward ends" (Kockelman, 2007, p. 375). Discussions around agency often bring some elusive issues such as intentionality and free will to the foreground which makes providing a more detailed definition a notoriously complex affair. The limited space here will not do justice to a comprehensive review of agency, from its roots in the philosopher John Locke's ideas on power and free will in the Enlightenment era (see Locke & Woolhouse, 1997, Chapter xxi) to when it gained popularity by sociologists like Giddens (1979) and Bourdieu (1977) and to the very current developments around it across various fields (e.g., Enfield & Kockelman, 2017). Therefore, this section reviews some selective studies from the neighboring disciplines to which the current thesis is based in.

Despite the diversity in the definitions of agency, there is at least one common thread which runs through all of them, and that is *action*. In fact, *The Concise Oxford Dictionary of Sociology* (Marshall, 2009) has taken agency almost as a synonym for action where it says agency "is often no more than a synonym for action" (see also Campbell, 2009 for a review of how other field-specific dictionaries define the term). This conceptualization of agency readily ties it to language since language is considered as a kind of action by a wide variety of scholars. In philosophy, it was Austin (1955) who initially introduced performative utterances – utterances which are part of doing actions. Later, he proposed that all utterances, and not just performatives, are speech acts. Searle (1969) then built Austin's speech acts into his Theory of Speech Act and defined 'conditions' to distinguish different types of speech acts. Even though these early philosophical works were not without problems (Derrida, 1982), they contributed to popularizing the *language as action* school of thought which continues to date (e.g., Levinson, 2013). The implication of such an approach to language is great: speakers of a language can be seen as agents who perform actions to and with one another (see Davidson, 2001 on theory of action).

In theoretical linguistics, however, agency is a grammatical concept and is described in terms of the semantic relations between grammatical categories. Agent is a thetamatic role expressed in grammatical categories (Chomsky, 1993, p. 34). One line of research in this area is ergative or other grammatical features that some languages use to distinguish the 'agenthood' of a grammatical constituent (e.g., Plank, 1979). In simple terms, an ergative construction is one which distinguishes an Agent subject from other kinds of grammatical subjects such as Actor, Experiencer, etc. through grammatical features such as case marking (Coon, Massam, & Travis, 2017).[31]

Anthropologists and especially linguistic anthropologists take language use as one of the best habitats for studying agency (e.g., Ahearn, 2001). Duranti (2004) defines agents as entities who have "control over their own behavior", "whose actions affect other entities", and "whose actions are the object of

---

[31] A capital initial is used to distinguish the grammatical Agent from a social agent.

evaluation" (p. 453). Similarly, Kockelman's (2007) and Enfield's (2013) characterization of agency include flexibility (control over an action and its composition) and accountability for performing the action. Duranti's (1994) anthropological survey shows that some languages have tools to encode agency's constitutive elements. In Samoan, for example, the ergative construction is used by more powerful members of the society to attribute accountability in a blameworthy matter to another member (Duranti, 1994).

In sociology, Goffman's (1961) notes on 'selfhood' is associable with agency. He observes that in even institutions such as asylums or prisons where there is relatively strict surveillance and control over a participant's behaviour, participants use practices to preserve their selfhood despite all the imposing regulations. One such practice, he notes, is using institutionally legitimate actions to pursue self-defining and independent goals. For example, a prisoner may order a book – an authorized action – not necessarily to read it, but to, say, impress the authority or annoy the librarian (see also Shreeya, 2018). Thus, the participant is exercising some degree of agency, in relatively most constraining environments, to pursue their own objectives. An individual's "pursuit of their idiosyncratic objectives" is also how Gibson (2000) broadly conceptualizes agency in conversation (p. 396). His notes on conversational agency illuminate that analyzing agency in interaction yields even a more complicated view than what social theories of agency have offered. This is because in interaction participants face constraints – conversation rules such as one-speaker at a time, relevance, etc. – which they are normatively expected to abide with. He criticizes conversation analysts, the ones who are most directly concerned with interactional constraints, for not tackling the issue in more depth. To address this problem, he distinguishes between two types of agency: a more colloquial type which is "action that successfully advances a person's preconceived objectives", and a more technical type which is employing practices to get round particular conversational constraints (p. 369). For example, to tackle the conversational constraint of relevance, one may use 'topical steering': "putting a particular spin on what was just said, selecting from the available interpretations that which [sic] directs conversation in the more favorable direction" (p. 378). I adopt a similar conceptualization of agency in this chapter: a participant's agency is manifested in their employment of actions or practices which pursues an interactional outcome.

Despite scholars' notes on the significance of studying agency in conversation and language use, research on agency in CA is relatively recent. After some early work on agency in medical encounters (e.g., Peräkylä, 2002), it was mainly Heritage and Raymond's (2005) work on epistemic authority in assessment sequences, which, although did not touch agency directly, planted the first seeds for expanded investigations on agency in interaction. Their study illustrates some practices by which a second assessor, who is agreeing with a first assessor, claims epistemic authority despite being positioned second in the sequence. Later, Enfield (2011) discusses Heritage and Raymond's (2005) work in light of agency. Building upon Kockelman's (2007) account, Enfield breaks down agency into is constitutive components: controlling, composing, and committing to a course of action. He argues that there is an interplay of these three elements in Heritage and Raymond's (2005) assessment sequences, and what they referred to as the 'primary right' for making an assessment can be analyzed in terms of agency components.

Further CA work then started to address agency more directly. In their subsequent work, for example, Heritage and Raymond (2012) investigated agency in request sequences (request for information and request for action). They focus on two kinds of responses: type-conforming (yes/no) and repetitional responses to polar questions. For example, the following two extracts show a yes/no and a repetitional response, respectively.

Extract 5.2. Field: 1:1:14-20, From Heritage & Raymond (2012, p. 182)
```
      01    MUM    'Av your family gone o:ff?
      02           (.)
->    03    LES    Ye:s,
```

Extract 5.3. Breathless, From Heritage & Raymond (2012, p. 182)
```
      01    DOC    Miss:uz Robinson
      02    PAT    Yes.
      03    DOC    Right.>Can I< jus' put you on the machi:ne?
```

```
->    04    PAT    You ca:n.
```

They argue that repetitional responses "exert more agentive leverage" in relation to the terms of the question in both sequence types (p. 188). Repetitionals' more agentive leverage, they discuss, may stem from them being less anaphoric and thus less indexically dependent on the questions. In response to requests for information, repetitions claim epistemic primacy and in response to request for action, they show "a greater degree of commitment to a future course of action" (p. 188).

Clayman (2013) explored prefatory address terms in light of agency. He notes that in an interactional position where a response is relevant, a prefatory address term can function to reduce the responsiveness of the action. These terms of address, he argues, are designed to "recast the action being implemented as autonomously motivated rather than arising from its local sequential context" (p. 296).

Extract 5.4. [10] in Clayman (2013, p. 295)

```
      01    KAR    Will      I       be      invited      next      week?
      02           (0.4)
      03    VIC    tch
      04           (0.2)
->    05    VIC    Karen   you've  been   invited   every   week   b'tchu   dis
      06           never:: (.) get around tih coming do:wn.
```

As Clayman explains, the address term, prefacing the response (line 5), frames the response as arising independently from the request (line 1) and, therefore, indicates an agentic independence of the response.

Furthering the work on assessment sequences, Thompson, Fox, and Couper-Kuhlen (2015) address agency based on the syntactic structure that second assessors use to agree or disagree with a first assessment. They propose a continuum of agency for the following syntactic structures: minimal clausal, expanded clausal, and graded clausal second assessments, exemplified in the following extracts in order.

Extract 5.5. [4.14] in Thompson, et al. (2015, p. 163)

```
      01    SHA    and        they're     usually     very     good.
      02           (0.4)
->    03    NAN    ↑yah they ↑a:re.
```

Extract 5.6. [4.15] in Thompson, et al. (2015, p. 167)

```
      01    ABB    that looks very warm.
->    02    STA    it is very warm.
```

Extract 5.7. [4.21] in Thompson, et al. (2015, p. 181)

```
      03    FEL    (it's) very pretty::[:
->    04    LUC                        [oh it's gorgeous.
```

The extracts reproduced here are taken from their agreement sequences, but they investigate these syntactic types in disagreeing second assessments as well. They conclude that for agreeing second assessments, the display of agency is weakest in Minimal Clausal forms and strongest in Graded Clausal forms, with the Expanded Clausal forms in between. For disagreeing second assessments, they claim, "Expanded Clausal forms embody stronger disagreement than Graded Clausal (downgraded) forms, as well as greater agency" (p. 189).

Assessment sequences have been a locus to investigate agency in other languages as well. In Swedish and German, for example, Auer and Lindström (2021) observe that a second assessment's grammatical dependence on the first assessment turn has implications for agency. The propose a continuum of agency in which syntactically reduced second assessments, which are indexically less dependent on the prior turn, are less agentive compared to full second assessments with the actual evaluative term (e.g.,

the reduced format *ich auch* 'I think so too' compared to full format (*das*) *find ich* (*auch*) + *evaluative term* 'I find that also evaluative term'). They argue the employment of reduced versus full syntactic format has projective potentials in terms of sequential development as well: the reduced formats which do not display disaffiliation project sequence closure whereas after full formats sequence expansion is more likely.

CA investigation on agency continues to grow with more researchers exploring various aspects of interaction in relation to agency. On the *action* level, for example, Koenig's (2011) study on medical interactions shows that an adult patient's non-accepting the doctor's recommendation enhances the patient's agency in the treatment process (see also Sidnell, 2017; Stevanovic & Peräkylä, 2012 on action and agency). On the *turn-taking* level, Vatanen (2018) argues that early-onset recognitional overlaps which agree with an ongoing turn claim agency and independence in the overlapping speaker's assertion. More studies, however, continue to focus on the *verbal composition* level as most of the early CA work on agency did, for example, impersonal deontic imperatives in recruitment sequences in Italian and Polish (Rossi & Zinken, 2016); zero person as opposed to second person references in Finnish (Etelämäki, Voutilainen, Weiste, 2021); and non-anaphoric references in English (Raymond, Clift, & Heritage, 2021). More recently, research on the *bodily composition* level has also started to address agency (Deppermann & Gubina, 2021; Tuncer & Haddington, 2020).

The majority of the studies on agency in CA so far focus on the responding actions in the second position in light of the verbal design features, or more precisely, on how participants resist being the 'second' in the sequence using verbal resources (but see Deppermann & Gubina, 2021; Raymond, et al., 2021; Rossi & Zinken, 2016; Tuncer & Haddington, 2019). What is more, even though these studies show through examples the precise practice or constructional feature that manifests participants' agency, many of them seem to lack a thorough conceptualization of agency: what does it really mean to be agentive in interaction? And how does it work? This chapter is a significant development in that firstly it focuses on the agency in initiating turns; secondly, based on the theories on agency, it makes it explicit how agency is conceptualized here and how it works; and thirdly, it takes a multimodal approach to agency. The following section clarifies how, among all these various approaches to agency, this concept is used in this chapter.

5.3. Agentive Stance

I approach agency as a type of stance, scalar, relational, and dynamic (see Englebretson, 2007, on stancetaking in discourse; on stance in sociolinguistics, see Jaffe, 2009, and for a CA study on stance, see Wu, 2004). Du Bois (2007) defines stance as a social action:

> Stance is a public act by a social actor, achieved dialogically through overt communication means, of simultaneously evaluating objects, positioning subjects (self and others), and aligning with other subjects with respect to any salient dimension of the sociocultural field (p. 163).

In this rather complex definition, my focus is on the "overt communication means" phrase. Based on this definition, in analyzing agency as a stance one can and perhaps should trace it to overt means or practices employed by interactants. This approach to agency may be different from that of some other scholars such as Ahearn (2001), who take agency as the *capacity* to act while the praxis (or practices) are the action itself (p. 118). From these scholars' perspective, "agency is not a quantity that can be measured"; therefore, scalar quantifiers such as 'less' or 'more' are not useful in analyzing agency (Ahearn, 2001; p. 122); instead, different types of agency, for example, "oppositional agency, complicit agency, agency of power, agency of intention, etc." should be focused (p. 130). Approaching agency as a stance, however, has different implications. Stance, as defined by Du Bois, is achieved by overt means which, as we will see, can be scalar in terms of their explicitness. Therefore, it makes sense to talk about a 'stronger' or 'weaker' agentive stance if participants' visible conduct supports it. Consequently, if we view agency as a stance, then we can approach agency as a scalar phenomenon.

The scalar view towards agency is in line with how other types of stance, such as affective and epistemic, are approached in CA and related fields (Sorjonen & Peräkylä, 2012; Heritage, 2012, among others). For example, Ochs (1996) talks about *degrees* of emotional intensity to define affective stance or Goodwin et al. (2012) analyze affective stance in terms of up- or downgraded displays. Similarly,

for epistemics, Heritage (2012) provides a gradient representation of epistemic relationships between participants based on the grammatical realization of their turns at talk.

Furthermore, agency is treated as relational and dynamic here. Using Du Bois's (2007) definition of stance, in the pursuit sequences under investigation, the pursuing party's agentive stance is achieved in positioning themselves as one who sets to get the recipient to do an action and the recipient as the one who resists. In this sense, agency is relational, and it is dynamic since participants can constantly reposition themselves to one another. Using Goffman's (1981) terms, participants may change their agentive relation to one another as the *animator* (the agent who implements an action), the *author* (the agent who decides what should be done and how), and the *principal* (the agent who is accountable for the action) (p. 144).

This chapter aims to show how manipulating the multimodal design of one's action may have implications for their agentive stance. Before presenting evidence for these claims two points are worthy of a mention. By approaching agency as a dynamic stance, I do not treat it as an internal static characteristic of a participant. For the ease of referencing, I sometimes use the term 'agency' rather than agentive stance; however, whenever the term is used in this thesis, it is a shorthand of 'agentive stance as displayed by some observable conduct'. Also, given the relational nature of agency, for most interactional contexts, a complete picture of agency takes into account the agency of both participants and how one influences the other. However, due to limitations of space, this chapter focuses only on the agentive stance displayed by imperative-speakers when pursuing compliance. The complex inter-relationship between the agentive stance of both parties, when they use various methods of pursuit and resistance, should be the aim of future investigations.

5.4. Imperative-Speaker's Agentive Stance in Pursuit Sequences

One way to depict an imperative-speaker's agentive stance when pursuing compliance is to compare it with their stance when they do not pursue compliance. This section compares the agentive stance of an imperative-speaker, whose CoA is impeded by the recipient's resistance, in three sequential environments: (a) non-pursuits, that is sequences in which the imperative-speaker simply *abandons* the sequence after the recipient's noncompliance; (b) downgraded pursuits, and (c) upgraded pursuits. The claim is that the very act of expanding the sequence to pursue compliance displays the imperative-speaker's stronger agentive stance relative to abandoning the sequence. Moreover, a multimodally upgraded pursuit turn embodies a yet stronger agentive stance compared to downgraded pursuits.

5.4.1. Agency and Abandoning

Before delving into pursuit sequences, let us analyze agency in a sequence in which no pursuit occurs; that is, when the imperative-speaker simply abandons the sequence after the recipient's resistance. Consider the relationship between the organization of adjacency pair and agency. The organization of an adjacency pair holds that the relationship between a first and second pair part is that of conditional relevance; that is the relevance of a responsive action is *conditioned* by an initiating action such that the non-occurrence of a response comes off as a noticeable absence (Schegloff, 2007, p. 20). On a normative level, an FPP-speaker limits the infinite number of possible next actions that their co-interactant could have performed to only a smaller subset of type-related actions that relevantly respond to the FPP. Implicitly then an FPP-speaker *influences* or *affects* their co-interactant's next action and how the interactant's next action is interpreted (again at least on a normative level).

Affecting other entities or the world around is a main feature of an agent in many definitions, conceptualizations, or theories of agency (e.g., Duranti, 2004; Kockelman, 2017, pp. 25-38). Thus the very act of producing an FPP is to some extent a display of agency in its own right. This display of agency is more overt when the FPP turn is designed to get the co-interactant to perform a physical action such as closing a window, turning off a tap, handing in a mobile phone and so forth. With such FPPs, not only is the co-interactant normatively expected to produce a response, but the expected response is, again on the normative level, performing a physical action. Stated differently, the FPP-speaker employs their co-interactant to achieve their interactional outcomes. In terms of Kendrick and Drew's (2016) recruitment model, the imperative-recipient is recruited into the imperative-speaker's CoA. Therefore, in light of Kockelman's (2017) conceptualization of agency (i.e., "flexible wielding of means toward

ends", p. 375), the co-interactant is treated as a 'means' towards the interactional end of furthering one's CoA (see Enfield & Kockelman, 2017, on distributed agency).

Then by virtue of the conditional relevance rule, an imperative turn can display its speaker's agentive stance. When the recipient does not comply, they refuse to be recruited by the speaker and hence they resist the imperative-speaker's agency. Facing resistance, the imperative-speaker may abandon the sequence and disengage from the sequence straightforwardly (i.e., not pursuing a preferred response; see Chapter 1, §2.6.1 on abandoning). In this way, they implicitly 'accept' the status quo and succumb to the fact that their CoA is not advanced by the recipient at that moment.[32] In this way, the imperative-speaker displays a weaker agentive stance than if she were to pursue compliance. The following extract illustrates this. The participants (Ebi: father; Haleh: daughter; Amir: Son) are taking turns to look at photos on Haleh's mobile phone (In Figure 5.3, left to right: Haleh, Ebi, and Amir). After Ebi is done looking at a photo, he stretches his hand towards Amir for him to take the phone and look at the photo, but it is Haleh who takes the phone. Amir issues Haleh the imperative 'give me', followed by the account 'I want to see', but Haleh explicitly refuses to comply by saying 'no'. Our focus is on Amir's conduct after Haleh's refusal: He immediately retracts his hand and shifts his gaze down – embodied movements which display that he has abandoned his CoA (lines 4-5).

Extract 5.8. RAY048b, 11.51 [left to right: HAL, EBI, AMI]

```
01      EBI    †biâ#
               IMP-come.2SG
               Here.
               †stretch hand twrd AMI-->
        ami >>*gaze down-->
        fig       #5.3
```

#5.3

```
02             (0.2)+*(0.1)±#(0.6)
        hal        +strtch hand twrd phone-->
        ami             ±strtch hand twrd phone-->
        ami     -->*gaze twrd phone
        fig             #5.4
```

#5.4
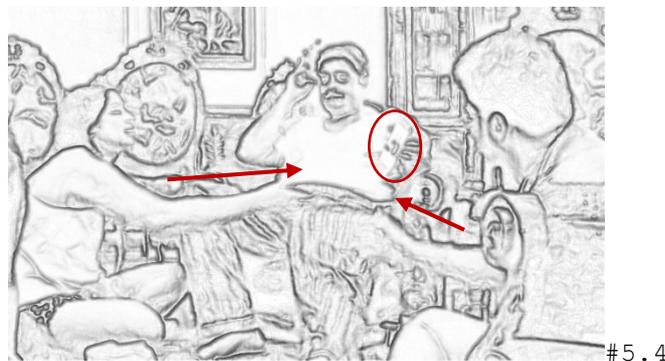
[32] They may try later or may try with a different co-interactant.

```
03    AMI    +be-de      man?# +± mi-kh-âm    be-bin-a[:m#†+±
               IMP-give.2SG 1SG    PRG-want-1SG SBJV-see-1SG
               give me. I want to see.
04    HAL                      # +±               [na#†+±vâstâ
                                                   No   wait
                                                   No   wait
      ebi                                          -->†
      hal  -->+takes phone-------+retracts hand-------------+
->    ami                       -->±move hand twrd HAL-------±retrct-->
      fig                    #5.5                       #5.6
```
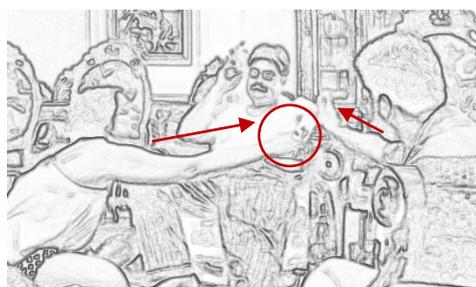
#5.5

#5.6

```
05    HAL    man dâr-am±*# mi-bin-am    felan.
               ISG PRG-1SG   PRG-see-1SG now
               I'm seeing for now.
->    ami            -->±arm on knee-->>
->    ami            -->*gaze down -->>
      fig          #5.7
```
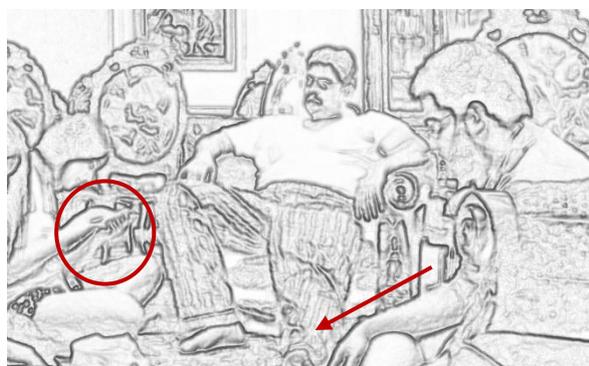
#5.7

```
06           (0.7)

07    HAL    âkh  to  cheghad eshghi ((addressing someone in the photo))
               INTJ 2SG how     love-be.2SG
               Oh How lovely you are
```

Even though Ebi stretches his arm towards Amir and thus selects Amir as the recipient of his offer (line 1; Figure 5.3), both Haleh and Amir stretch their hands towards the phone, competing for it (line 2; Figure 5.4). Haleh wins the competition and takes the phone earlier (Figure 5.5). At this point, Amir could have let Haleh take the phone without entering a competition to get hold of it, but he issues the imperative *bede man* 'give me' (line 3). This imperative, on its own, can display his agentive stance in furthering the CoA that his arm-stretch towards the phone has initiated: the CoA of getting hold of the phone. It makes a response conditionally relevant, and it normatively limits Haleh's next action to only

type-related relevant responsive actions. The imperative is immediately followed by Amir's account 'I want to see' (line 3).[33]

Concurrently with the imperative, Haleh retracts her hand, holding the phone, towards herself, which already projects a dispreferred response (line 4). As Haleh moves the phone towards herself, Amir's stretched arm also moves parallel to the phone, with his gaze also following the movement of the phone. That is, Amir changes the direction towards which his arm is stretched from Ebi to Haleh (Figure 5.6). Haleh then produces a dispreferred response in a preferred format (line 4, 'no' in overlap) (Kendrick & Torreira, 2015), directs Amir to wait (line 4, 'wait'), and offers what appears to be designed as an 'account' for the imperative (line 5, 'I'm seeing for now.'). Haleh's rejection turn is a contingent outcome. That is, even though she is normatively expected to produce a response, she could have just refused to further Amir's CoA on the embodied level (e.g., just keeping the phone to herself) with no explicit verbal rejection. But even in that case, (a) a response is still noticeably absent and (b) Amir's imperative makes Haleh's embodied conduct (holding the phone) to be framed as rejection or resistance rather than a mere bodily action in a series of other bodily actions. So far as the imperative affects how Haleh's next action is understood, the imperative can still display its speaker's agentive stance. Using Duranti's (2004) conceptualization, the imperative *influences* how Haleh's action is understood.

At this point, Amir could have pursued his CoA, but as soon as Haleh produces the rejection token 'no' (even before she brings her turn to an end), he drops his arm on his knee and shifts his gaze down (embodied conduct in lines 4-5; Figure 5.7). These two bodily actions display his straightforward withdrawal from the sequence. After a gap (line 6), Haleh initiates an assessment sequence on a photo she is looking at (line 7, *to* 'you' is addressed at a person in the photo) and continues to keep the phone. Amir has abandoned the sequence. Through abandoning the sequence, he puts a desisting or relinquishing stance on display, a weaker agentive stance than if he were to pursue a preferred response.

The participants' colliding CoAs are visually depicted in Amir's reducing distance between the requested object and his hand by stretching his arm and Haleh's counteracting by moving the phone towards herself and away from Amir's hand. On the verbal level, it is manifested in the imperative and the rejection. Amir could have upgraded his agency in furthering his CoA by stretching his hand further or issuing another imperative, among other means, in which case he would have shown a stronger agentive stance. In the following examples, we consider agency in pursuit sequences.

5.4.2. Agency in Downgraded Pursuits

After a recipient's resistance, the imperative-speaker may choose to pursue compliance. Pursuing compliance involves expanding the base adjacency pair by which the imperative-speaker makes another attempt at recruiting the recipient to further her (imperative-speaker's) CoA. A pursuit turn provides the recipient another slot to comply and advance the CoA. As discussed in the previous chapters, pursuing a response after a recipient's display of resistance treats resistance as inadequate and, therefore, limits the acceptable response choices to include only compliance (the preferred option).

The claim is that pursuing a preferred response after the recipient's noncompliance can display the pursuing party's stronger agentive stance than her stance if she were to abandon the sequence. By pursuing a preferred response post-resistance, not only does the imperative speaker limit the recipient's next action to a range of type-fitting responsive actions, but she also limits the valence of the responsive action to only preferred actions. It is precisely this limitation manifested in *another attempt* and the provision of *another response slot* that can reveal the speaker's stronger agentive stance relative to abandoning. This section presents an example of pursuing with downgraded multimodal features. It

---

[33] In this example, the account does not sound to be pursuit post-rejection. Firstly, the account is latching to the imperative TCU. Secondly, Haleh produces her rejection token after the account, showing her orientation to Amir's imperative and the follow up account as belonging to the same turn.

aims to show that even though the downgraded design projects a potential withdrawal, pursuing still manifests the speaker's stronger agency than straightforward abandoning.

Melisa and her daughters, Minuvash and Nasima (late teenagers), are sitting around the breakfast table with a cut melon on it. There is another uncut melon in the fridge. Melisa has tried several times to get Minuvash to eat something, but she is playing a game on her tablet. Earlier in the recording, Minuvash assesses the melon on the table as not fresh and declares that she wants a melon that is not cut yet. Just prior to the extract, there is a rather long disagreement sequence on whether or not the melon on the table is fresh (not transcribed). Lines 1-3 are the last lines of that sequence. The imperative of interest is *bâz kon* 'cut {it} open' by which Melisa redirects the action of preparing a fresh melon to eat to Minuvash herself (line 5). Minuvash shows no orientations to compliance, and Melisa re-issues the imperative with downgraded multimodal design (line 9).

Extract 5.9. RAY036a (2), 09.47

```
       01   MEL   m: alân un   yeki dige  ham bâz     be-sh-e,
                      now  that one  other too open   SBJV-become-3SG
                   m:: now if the other one gets cut open,

                   ke mese hamin mi-sh-e.
                   PRT like this PRG-become-3SG
                   it also becomes like this.

       02         (0.8)

       03   MIN   na ne-mi-sh-e.
                   no NEG-PRG-become-3SG
                   no it won't.

       04         (0.1)

 ->    05   MEL   *+khob [bo-ro    #bâz*  kon.+*
                      PRT  IMP-go.2SG open do.IMP
                    then go cut open.
                   +head point twrd fridge-----+
                   *gaze twrd frige-----*gz at MIN*
             fig                     #5.8
```


#5.8

```
       06   NAS   [sânieye   avale[sh      khub-e    bade-sh=
                   second-Ez first-Ez-3SG good-be  after-3SG
                   the first second it's good after that

1.5 s  07   MEL                       [bo-ro      az   tu takhchâl darâr,=
                                        IMP-go.2SG from in fridge take.out
                                        go get {it} from the fridge,
```

```
08    NAS  =*[dige khub ni-st          *
               else good NEG-be.3SG
             it's not good anymore.³⁴
```

```
->  09    MEL  =*[barâ khodet    #<<pp>bâz [kon>*
               for  yourself       open do.IMP
             for yourself cut open.
             *gaze down---------------------*
      fig                    #5.9
```


#5.9

```
10    MIN                              [! ah in khofâsh-â ne-mizâr-an.
                                        INJT this bat-PL NEG-allow-3PL
                                        ! damn these bats don't allow.
```

The imperatives in focus are designed to get the recipient to prepare her breakfast for herself (lines 5, 7, 9). After a disagreement sequence about the melon (lines 1-3), Melisa instructs Minuvash to cut the melon in the fridge open to finally get her to have her breakfast (line 5, 'then go cut open'). There is almost half a second, during which Minuvash could have shown at least some embodied preparation for performing the instructed action (e.g., putting the tablet down or getting up); however, she continues to play the game (lines 6-8). Minuvash's continuing her own CoA in a responsive slot is implicative of a dispreferred action. Melisa then pursues compliance by issuing another imperative which clarifies the steps involved in preparing the melon (line 7, 'go get {it} from the fridge') and then reissues her prior imperative with a modified verbal design (line 9, 'for yourself, cut open'). The modified turn design foregrounds Minuvash as the person who benefits from the instructed action and implicitly as one who should perform it (line 9, 'for yourself cut it open'). By this point in the interaction, it becomes clear that as well getting Minuvash to start her breakfast, there is also another issue at stake: between the two participants, who prepares Minuvash's breakfast. The pursuit is another attempt by Melisa to get Minuvash to start her breakfast and at the same time makes it clear that she (Minuvash) is the agent who should prepare her breakfast for herself. Minuvash again shows resistance by displaying involvement in the competing activity of playing game (line 10, 'The bats {in the game} do not allow {it}').

The series of imperatives designed to get Minuvash to prepare her breakfast consists of three: 'then go cut open' (line 5); 'go get from the fridge' (line 7), and 'for yourself cut open' (line 9). Let us focus on the line 5 and its modified repeat, line 9. (line 7 also includes downgraded features but for presentation purposes, we only compare line 5 with 9). Even though the verbal design of the pursuit is modified, in both renditions the imperative verbs occur in the same position within their TCUs (TCU final), which makes the imperatives still prosodically comparable. F_0 and intensity traces for both full TCUs and the imperatives alone are provided separately (Figures 5.10 and 5.11), but here I focus on the prosodic information on the actual imperative verbs (for more on the whole TCUs, refer to Table 5.1). Note that

---

³⁴ Nasima's turn (lines 6, 8) is a 'sarcastic' assessment of any melon; that is, a melon is fresh (good) only in the first second after getting cut. After that, it loses its freshness. The sarcasm targets her sister, Minuvash, who has been arguing with her mother that because the melon on the table has been cut, it is not fresh anymore.

in both versions, the second part of the phrasal verb *bâz kon* is produced whispery, hence no pitch traces corresponding to that part in the spectrograms.

The subsequent imperative is multimodally downgraded compared to the first version (lines 5, 9). The first imperative is produced 3.7 ST below the speaker's average F_0 whereas in the subsequent version this dips to 5.9 ST below the overall average F_0. The intensity level is 70 dB in the first imperative, but only 60.5 dB in the second, which makes it sound very quiet. A similar reduction of measures holds for the span (2.2 ST vs. 0.4 ST) and duration (360 ms vs. 290 ms) of the second version.
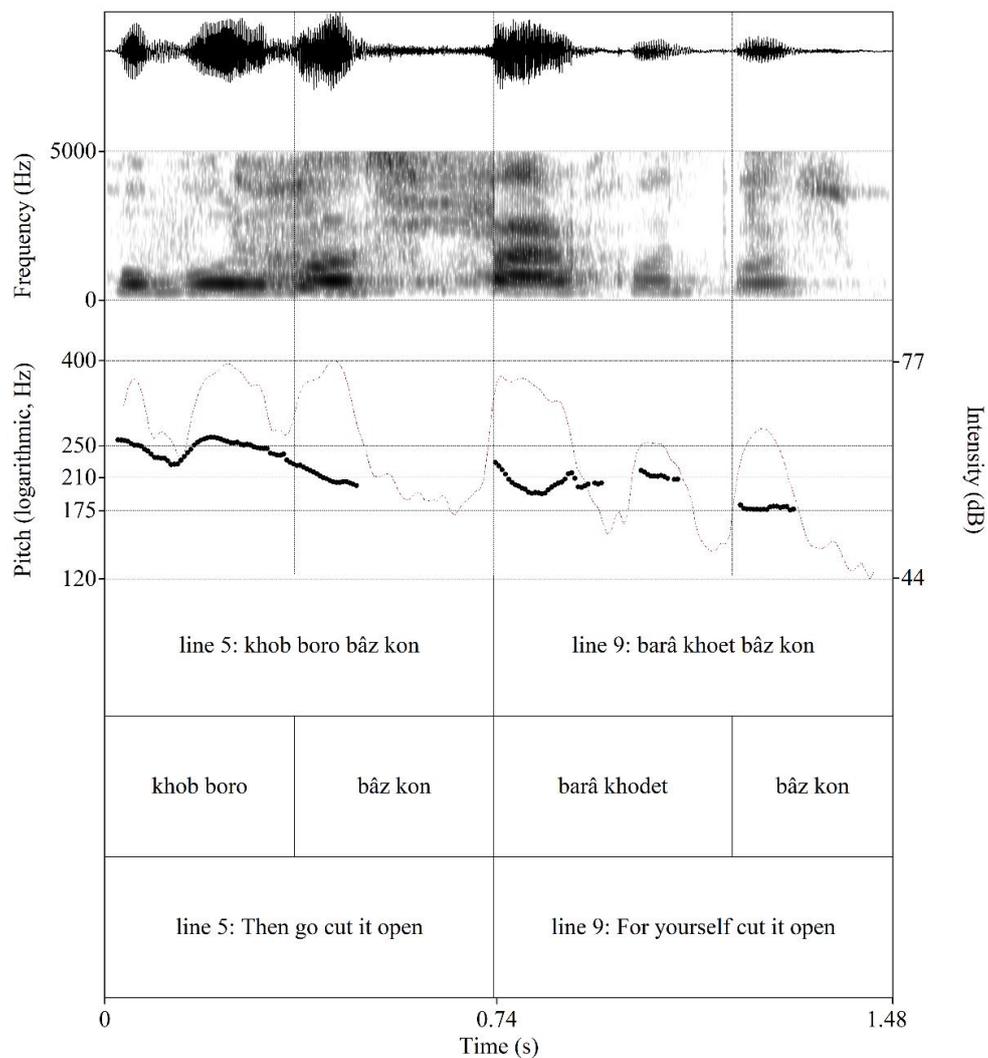
Figure 5.10. Prosodic downgrading considering the whole imperative turns in Extract 5.9.

Table 5.1. Summary of the design features of the imperatives TCUs (dark gray) and verbs (light gray) in Extract 5.9.

Line	TCU's Average F_0	Imperative's average F0	Intensity (dB)	Span (ST)	Duration (ms)	Embodiment
5	239 (- 0.7 ST)	205 (- 3.4)	70	2.2	360	Head point
9	200 (- 3.8 ST)	177 (- 5.9)	60.5	0.4	290	Gaze down

160
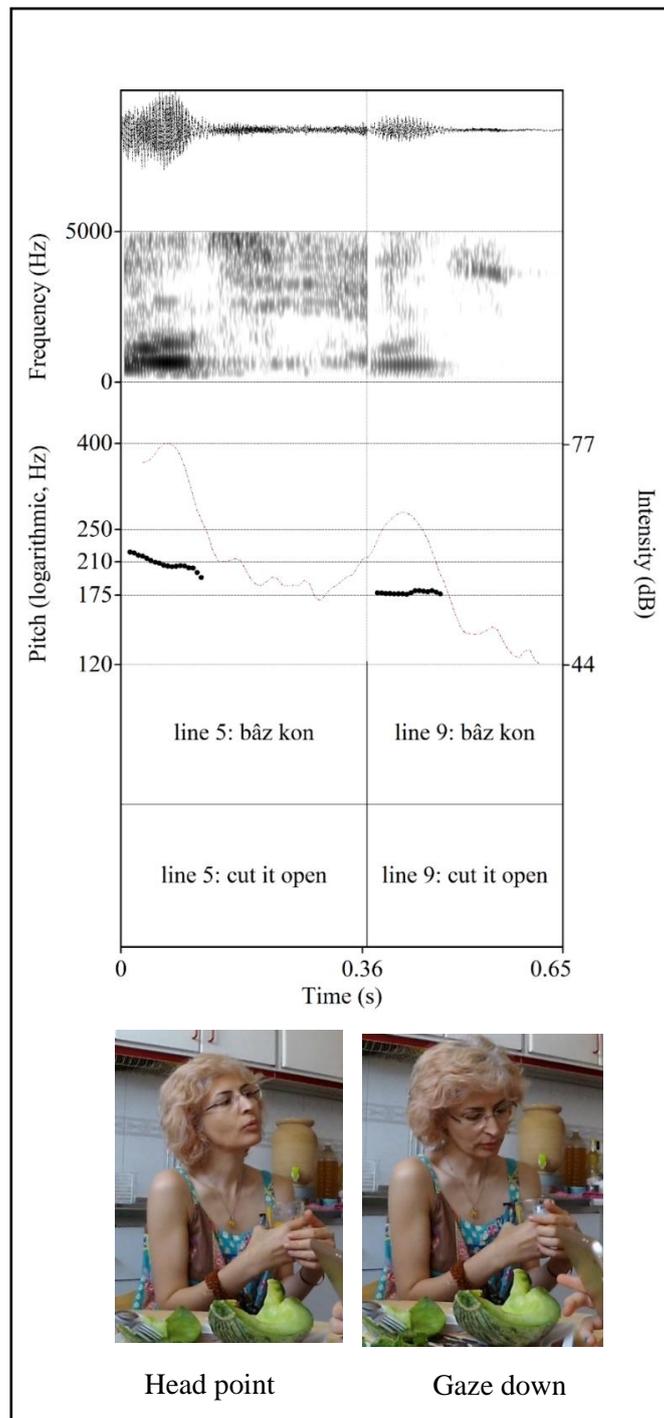
Figure 5.11. Multimodal downgrading in Extract 5.9. With the first imperative, there is a head point towards a relevant object (the fridge) while in the subsequent version, the speaker's gaze is averted.

The imperative-speaker's bodily conduct is also downgraded. The first version is produced as the speaker points towards the fridge with her head (line 5, Figure 5.8) and then shifts her gaze towards the recipient (line 5). These two bodily actions align with the instruction 'go get from fridge' where she alludes that even getting the melon from the fridge should be done by Minuvsh herself. With the modified repeat, however, her gaze is down on her hand which she is scratching (line 9, Figure 5.9). Thus the head point is dispensed with and her gaze is averted. These two bodily actions project the speaker's withdrawal from the sequence (Rossano, 2012). Therefore, disengagement occurs through a process, actualized in the imperative-speaker's reduced involvement in getting the recipient to prepare

her breakfast for herself. Her reduced involvement itself is reflected in the pursuit imperative's downgraded turn design.

Now regarding agency, although Melisa's pursuit has downgraded features, it still appears to embody stronger agency than if she were to simply abandon her sequence (Note that the agentive stance displayed in the subsequent version is downgraded relative to the first). At the sight of Minuvash's resistance, Melisa could have simply withdrawn from the sequence. However, she first clarifies the step involved in getting the melon (line 7, 'go get from the fridge') and then repeats the imperative (line 9, 'for yourself cut open''). The re-doing of an action against resistance when it could have been simply not re-done displays a pursuing party's stronger agency. Recall that the concept of action is at the heart of agency. Consider again some conceptualizations of agency in the related fields: actions towards one's objectives (Gibson, 2000) or "wielding of means toward ends" (Kockelman, 2007, p. 375). In our example here, Melisa re-does her prior action in a more clarified design and hence she performs more actions and resorts to extra means towards the objective or end of getting Minuvash to prepare her breakfast for herself (lines 7 and 9). As mentioned earlier, despite its downgraded design, the subsequent imperative (line 9) still revives the relevance of a response, but since noncompliance is treated as non-acceptable, it limits the acceptable valence to only preferred. This is another aspect that displays the pursuit speaker's stronger agentive stance than if the sequence were to be abandoned.

5.4.3. Agency in Upgraded Pursuits

The discussion about agency in the two previous sections was based on the *organization of sequences*: the conditional relevance rule (§5.4.1) and sequence expansion post-resistance (§5.4.2). In this section we turn our attention to the relationship between agency and a pursuit turn's *design features*. We compare two broad sets of design features: first, upgrades that produce another interactional slot for the recipient to respond; and second, upgrades that diminish the recipient's independent response (e.g., when the pursuing party aims to 'force' compliance). The claims are (a) upgraded pursuits embody a stronger agentive stance compared to downgraded pursuits, and (b) through diminishing the recipient's independent response, the pursuing party displays a stronger agentive stance compared to pursuit by creating another response slot for the recipient.

Upgraded pursuits that create another chance for a response

An imperative-speaker may upgrade the constructional features of her pursuit turn but still recognize the *recipient* as the agent who should advance the CoA. She can create another slot for a response as she *limits* the recipient's responsive actions on a *practical level*. In the example of a downgraded pursuit above, too, the pursuit turn offered another interactional slot for compliance and it limited acceptable responses to only the preferred option. But this limitation was on a *normative level*. In the upgraded pursuits that we are considering in this section, the limitation is on a *practical level*. I elaborate what is intended by a practical limitation with the following extract. Faraneh (leftmost in Figure 5.12) is offering a fig to Amir (rightmost).[35] When Amir rejects the offer, she pursues a preferred response through multimodally upgraded subsequent imperatives (*bokhor* 'eat'): she progressively stretches her hand, holding the fig, closer towards Amir and increases the prosodic measures of each successive imperative. With each upgraded pursuit, she physically limits Amir's personal space more (Hall, 1966, on personal spaces).

---

[35] In Persian culture, offers are sometimes ritualistically repeated. However, the rejections and the subsequent offers in this sequence do not make it recognizable as that ritual. The precise difference of the design and context of those ritualistic offer-rejection sequences and this sequence is an entirely different research project. Here suffice it to only note one difference. In ritualistic offers, when the account for a rejection is "I don't like X", typically the offerer ceases offering which does not happen here.

Figure 5.12. Participants in Extract 5.10. Faraneh (left), Ebi (Middle) and Amir (right) are having fruit.

Extract 5.10. RAY045a, 03.03

```
->      01      #+(0.5)
                far   #+holds hand with a fig stretched twrd AMI-->
                fig   #5.12
```



```
                                                      #5.12
```

```
        02   AMI    n((click))e-mi-khâ-m
                      NEG-         PRG-want-1SG
                          I don't want
```

```
->      03   FAR    bo-[kho.      #
                     IMP-eat.2SG
                     eat.
```

```
        04   AMI       [m      a\#
                        sound sound
                          m a
                     fig           #5.13
```



```
                                    #5.13
```

```
        05   AMI    anji ne-mi-khâ-m
                     fig  NEG-PRG-want-1SG
                     I don't want a fig
```

```
        06        (0.2)
```

```
->      07   FAR     +bo-khor.    +#
                      IMP-eat.2SG
                       eat.
                     -->+strtch more+
                     fig               #5.14
```

163

#5.14

```
08          (0.5)

09   AMI    +((click))+[anji dust na-dâra-m
                   fig  like NEG-have-1SG
                   I don't like figs

->   10   FAR    +         +[bo-khor.    +#
                            IMP-eat.2SG
                            eat.
          far    +holds----+strtch more+
          fig                          #5.15
```


#5.15

```
11   EBI    +±bo-khor.±
                IMP-eat.2SG
                eat.
          far    +holds-->>
          ami    ±........±

12          #*±(0.8)
          ami    #±hand palm up under FAR's hand-->>
          ami     *frowns-->>
          fig    #5.16
```


#5.16

After the first offer, there are three subsequent versions before a third participant joins in the pursuit action. Each time the offerer pursues compliance, she expands the sequence and thus upgrades her agentive stance in furthering her CoA. The very first offer is only embodied: Faraneh stretches her hand holding a fig towards Amir, embodying an offer (line 1, Figure 5.12), which Amir rejects (line 2). At this point, Faraneh could have retracted her hand and treated the sequence as complete; however, she pursues a preferred response by holding her hand stretched and adding the imperative *bokhor* 'eat {it}' to the turn (line 3, Figure 5.13). Amir responds with another rejection (lines 4-5) and Faraneh continues pursuing through a verbatim repeat of the imperative (line 7). Amir rejects this offer too by his click which in Persian is a conventional vocal practice functioning similar to the negation token *na* 'no' (line 9), and then he accounts for his rejection. Overlapping with Amir's account, Faraneh issues yet another imperative (line 10). At this point, Ebi joins Faraneh to pursue a preferred response by repeating the

imperative *bokhor* 'eat {it}' (line 11). Concurrently, he produces an ostensibly aggressive action by holding a fruit knife pointing at Amir. This contributes to sanctioning Amir's repeated rejections (Afshari Saleh, 2020, this thesis Chapter 6, on mock aggression).[36] Finally, Amir's bodily conduct displays compliance: he holds his hand open and palm up near the fig – an embodiment which makes relevant the transfer of the fig (line 12, Figure 5.16) (Figure 5.17 shows the subsequent versions).
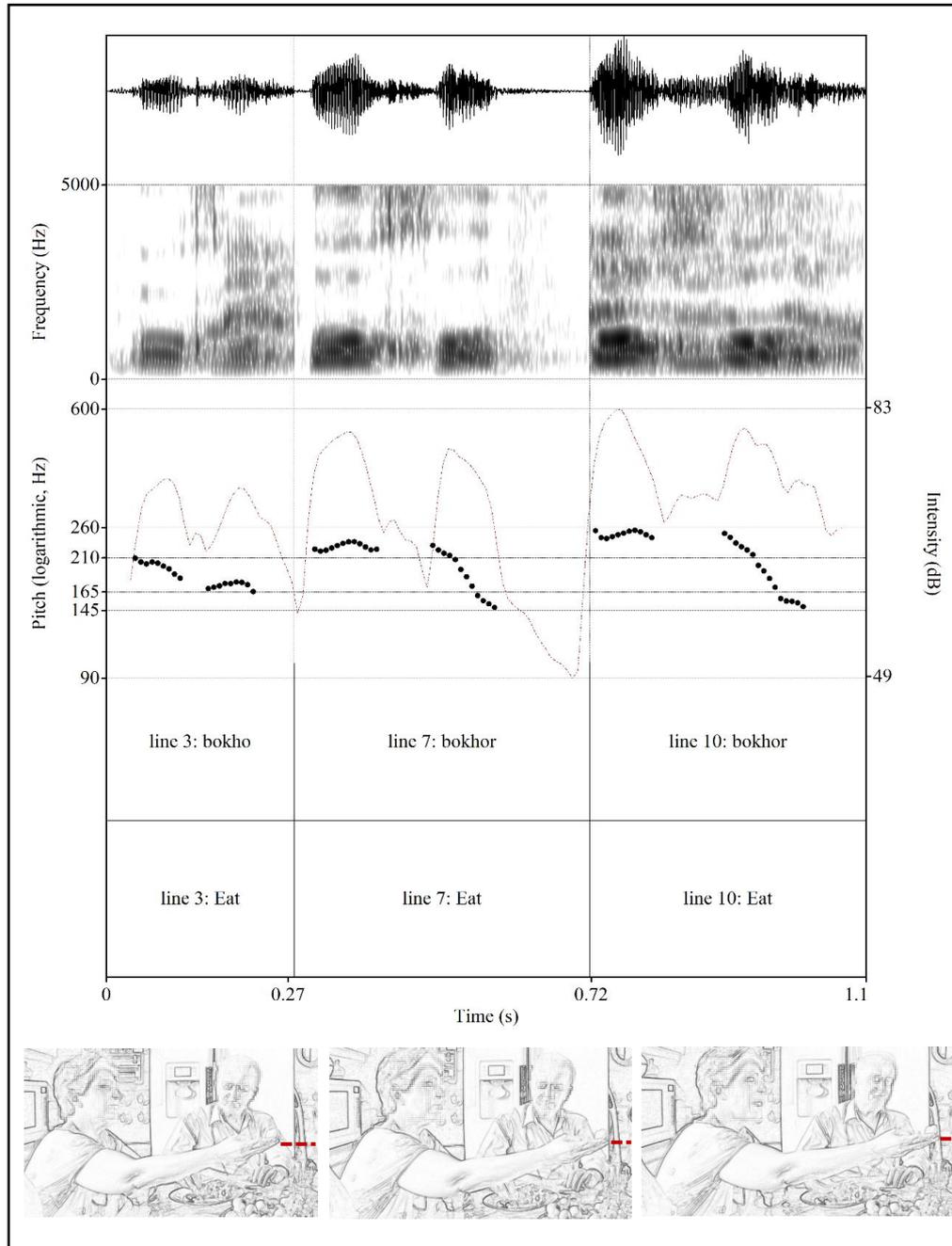
*Figure 5.17. Multimodal upgrading in Extract 5.10.* As the prosodic features (register, intensity, span, and articulation setting) of each successive imperative is upgraded, the speaker's hand is stretched further forward the recipient.

---

[36] An image of this is not presented since Ebi's hand holding the knife towards Amir is positioned right behind Faraneh's hand and is not easily seen in a still image.

The subsequent offers are performed through upgraded multimodal turns. The first subsequent version (line 3) is produced by adding a verbal modality to the initial embodied offer: Faraneh's hand's position does not change compared to her initial embodied offer (Figures 5.12 and 5.13), but an imperative is added to the design. Prosodically, the first subsequent version starts around Faraneh's overall average F_0 and drops low in her register. The TCU is not particularly wide or loud (compared to Faraneh's other stretches of talk in the same recording); it is short and, with the final consonant dropped, it is also underarticulated (Figure 5.17 above, Table 5.2). The second subsequent imperative is produced at a higher register, with a wider pitch span, greater intensity, and tighter articulation setting (line 7). Notice the spectrogram for the final consonant: not only is it produced but it is also relatively elongated (it is produced voiceless – [ɹ̥] – hence no pitch trace). The middle consonant [x] is also relatively longer, compared with its counterpart in the previous subsequent version. All this yields a longer imperative, and these measures (except for duration) are further increased in the third subsequent imperative. The articulation quality sounds as tight as the previous one in the third subsequent version (line 10, Table 5.2). (In the Praat representation, in the third box (line 10), the final voicing striations in the spectrogram, which has no corresponding pitch trace in the pitch panel, belong to Amir's overlapping talk. The pitch trace is corrected.)

Congruent with her verbal production, the embodied aspect is also upgraded. With each successive subsequent version, Faraneh stretches her hand forward, further away from home position. By doing so, she reduces the distance between the object and the recipient of offer – the fig and Amir (Figure 5.17 above. Table 5.2).

Table 5.2. *The prosodic measures of the subsequent imperatives.*

Line	Average F_0 of the TCU (distance from speaker's overall F_0)	Span (ST)	Intensity (dB)	Duration (ms)	Embodiment
3	170 (3.5 *below* the speaker's overall F_0 level)	4	70	270	Arm stretched
7	221 (1.0 *above* the speaker's overall F_0 level)	8	73	450	Arm more stretched
10	233 (2.0 *above* the speaker's overall F_0 level)	9	77	380	Arm most stretched

Faraneh's progressively upgraded agentive stance is observable on sequential organization and turn constructional levels:

- Sequential: Expanding the sequence by adding an imperative after the first embodied offer is rejected (line 3)
- Sequential: Further expanding the sequence after each rejection (line 3, 7, 10)
- Turn Constructional: Progressively stretching hand further away from home position towards the recipient, limiting the recipient's 'personal space' (line 7, 10) (Hall, 1966).

We have already analyzed the first two aspects in previous examples. Our focus here is on the third level. To see how Faraneh upgrades her agency, let us consider the relationship between personal space and agency. In his discussion of proxemics, Hall (1966) defines four levels of interpersonal zones: intimate, personal, social, and public spaces. The distance that is typically observed between members of a family or close friends is called the personal space. Regarding agency, two of the properties of agentive entities, as reviewed in §5.2, is 'control' on their behavior and 'influence' on others (Duranti, 2004). Now, by successively stretching her arm towards Amir, Faraneh is approaching Amir's body. She can be seen to be progressively limiting his personal space. Her successive arm-stretch can project a potential 'breach' of norms on proxemics. By *limiting* Amir's personal space, Faraneh can be seen to be potentially increasing her projected *influence* on Amir's next action: to further resist the offer, he may have to withdraw his body, or to protect his space from a potential 'invasion', he may have to accept the fig (among other possible moves of course). Indeed, in this example, even though Amir has already said that he does not like figs, he eventually holds his hand in a receptive position (i.e., accepting the fig) while frowning (negatively sanctioning Faraneh's upgraded persistence). Even though it is he himself who accepts the fig (e.g., Faraneh does not force it into his mouth), his frown can convey that his acceptance may have been in fact a way to deal with Faraneh's approaching hand and her continued

persistence. In the previous section, we saw that a downgraded pursuit, limits the response options to include only a preferred action on a *normative level*. With upgraded pursuit, in addition to normative limitations, the pursuing party limits the recipient's responsive options on a *practical (embodied) level*. And it is this change that displays stronger agency embodied in an upgraded pursuit compared to a downgrade pursuit. This does not mean that upgraded pursuits always receive acceptance. Rather, the limitations that they create are upgraded from a normative level to a physical level.

In this example, even though the imperative-speaker's agency is upgraded, the practice still creates a slot for the recipient himself to perform or not perform the action named by the imperative. That is, the imperative-speaker still recognizes the recipient as the agent who should further the CoA. In the next section, we will see that the imperative-speaker substitutes himself for the recipient to perform the named action despite the recipient's resistance against it. We will see that by substituting for the recipient, the imperative-speaker upgrades his agency in advancing his CoA even further.

Upgraded pursuits that diminish a chance for an independent response
Facing the recipient's resistance, the imperative-speaker sometimes performs actions that accomplish the pursued outcome of the imperative despite the recipient's resistance. Contrary to the previous group of upgrades, here it is the imperative-speaker himself who either performs the named action or performs other actions that accomplish the pursued outcome. He *substitutes* himself *for* the recipient. An example of this type of upgrade was shown at the beginning of the chapter: the imperative-speaker pulled the recipient's hand away to free the cat. This embodied upgrading happened *as* she repeated the imperative 'let it go'. Despite the verbal pursuit, the imperative-speaker's embodied action diminished a chance for the recipient to produce an independent response. This section focuses on the pursuing party's agency in this kind of upgraded turn.

In the following example, our focus is on the imperative-speaker's agentive stance, but in this extract, the details of the recipient's resistance are also discussed to show that a 'successful' substitution is only contingent and depends on the recipient's resistance. As will be shown, both participants in this extract upgrade the multimodal design of their turns, one to persist in advancing his CoA and the other to resist it.

Haleh (late teenager) has taken a blood test and the inside of her left elbow, where blood was taken from, is covered with a band aid. Earlier she had explained to her family how uncomfortable she felt at the sight of blood taken from her body. Just prior to the extract, Haleh was engaged with her mobile phone in her right hand. The extract begins as she directs her gaze towards Ebi, her father, asking him how she should wash her arm. Ebi offers the advice *bekan* 'Pull {it} off', (it referring to the band aid). After Haleh rejects to follow the advice, Ebi upgrades the multimodal design of the following imperatives (line 5, 8). Haleh too upgrades the multimodal design of her rejections. Amir is Haleh's brother. Ebi is holding a drink with his right hand (Figure 5.18).

Extract 5.11. RAY048b, 18.45

```
01   HAL    •bâbâ man ±vozu mi-khâ-m     ±be-gir-am=
            dad   1SG  vozu PRG-want-1SG SBJV-take-1SG
            Dad I want to wash {my arm}
            •puts arm on knee-->
                     ±gaze at  band aid ±


            =ch[i  kâr    kon-a:[:m]?
            what  action  do-1SG
            What should I do?


02   AMI     [un-o          mi-k[a ]n-i       vozu  mi-gir[-i
             that-OBJ.MAR PRG-pull.off-2SG vozu  PRG-take-2SG
                You pull that off {and} wash
```

167

```
03   EBI                              [be-kan      alân  be-ka:[n
                                      IMP-pull.off now   IMP-pull.off
                                      Pull {it} off now pull {it} off

04   HAL                                             [<<dim>na=
                                                               Na

        =#na>•± h. (.) <<whsp>ne-mtu[-(nam.)>
           no      h. (.)        NEG-can-1SG
         No no I can't.
    hal    -->•
    hal      ±gaze at cell phone-->
    fig    #5.18
```

R-hand holds a glass.
Band aid on the elbow pit
R-hand holds cell phone.
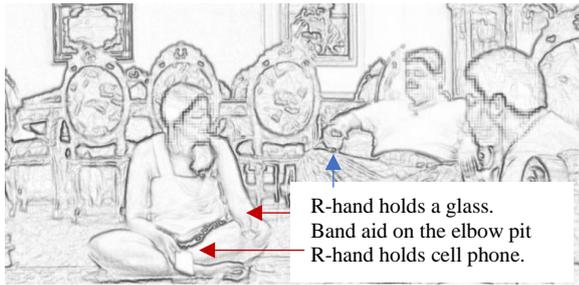
#5.18

```
->   05   EBI                        [+<<fh>vâstâ  bina-m.#(.)bi-â>+±
                                     wait.IMP see-1SG IMP-come
                                     Wait binam. (.) come
         ebi                     +switches hands holding drink+
         hal                                            -->±
         fig                                    #5.19
```

Switches hands to hold the glass.
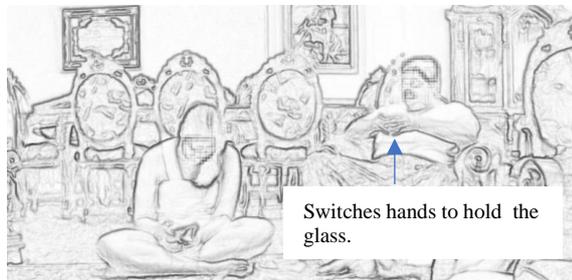
#5.19

```
06        †+(0.1)
    ebi    +stretches hand forward progressively-->
    ebi    †leans forward progressively-->

07   HAL  #*•<<ffhh>na:: •na na# [na>
                     no     no no    no
                   no no no no
    hal   *leans backward progressively-->
    hal   •...........•covers the band aid with hand-->
    fig   #5.20                #5.21
```

#5.20

#5.21

```
->   08   EBI                          [<<ffhh>na #vâst[â:>
                                        no wait.IMP
                                        No wait
          fig                                   #5.22
```


#5.22

```
     09   HAL                                   [<ffhh>na:::na:::h.>
                                                 no   no
                                                 No no


     10           #*•  (1.2)  +†*    (0.7)    † (3.0)#
          ebi                 -->+
          ebi                 -->†leans backward † continues drinking-->>
          hal    -->*scoots away*sits away from EBI-->>
          hal    -->•
          fig       #5.23                                   #5.24
```


#5.23

Arm rests on the chair's

#5.24
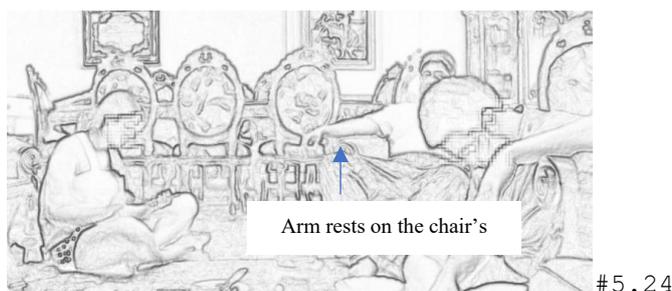
The imperative-speaker first advises a solution for the recipient's trouble (Kendrick, 2021). The solution is sought by Haleh's own request for advice on her trouble – washing her arm with a band aid on it (line 1). Both Ebi and Amir formulate similar solutions (lines 2-3). Ebi's advice is in the form of double imperatives (line 3, 'pull {it} off now pull {it} off'). The double imperatives and time adverb 'now' convey that the action of removing the band aid is already overdue. The advice sets forth an action for Haleh herself to perform. This solution is rejected by Haleh (line 4) by means of two rejection tokens. At this point, Ebi could have treated the sequence as closed and accepted Haleh's rejection by, for example, disengaging from the sequence. However, he expands the sequence and persists in advancing the CoA. He issues the imperative *vâstâ* 'wait' first in line 5 and then a subsequent version in line 8. These two are the imperatives in focus here. The imperatives together with Ebi's co-occurring bodily conduct – stretching his hand towards Haleh's arm – project that Ebi himself is setting off to pull off the band aid. Therefore, following the recipient's resistance against implementing the speaker's solution, the speaker performs actions that project a direct intervention (Kendrick, 2021), and by doing so he can be seen to be *substituting* himself for the recipient.

The prosodic and embodied design of the subsequent imperatives (lines 8, *vâstâ* 'wait') is upgraded relative to the first (line 5, *vâstâ* 'wait'. Figure 5.25; Table 5.3). The first imperative is followed by another lexical components *binam* in its TCU (line 5). *Binam* is originally a subjunctive verb, which literally means 'so that I see'. However, in spoken Persian it has adapted an extra affective function and works similar to a particle. Its literal translation here does not sound to be intended here. Its precise affective function should be investigated, but overall, it conveys a disaffiliative stance towards the imperative-recipient's action. In the subsequent version, this particle is dispensed with; nevertheless, the prosodic and embodied design of the second imperative convey a sense of upgrade in the overall turn design. The first version is prosodically loud (80 dB) and rather high (4.0 ST) compared to the speaker's other stretches of talk in the same recording. Note that this imperative is produced after Haleh has already shown resistance to performing the advised solution, which explains the disaffiliative particle and the prosodic design. These measures are further upgraded in the subsequent version (83 dB loud, and 12 ST high). The duration and span of the two versions are not readily comparable because of the extra lexical component *binam*.

Table 5.3. *Summary of the design features of the imperative 'wait' in Extract 5.11.*

Line	Average F_0 of the imperative *vâstâ* (Hz)	Intensity (dB)	Span (ST)	Duration (ms)	Embodiment
05	227 (4.0 ST above speaker's average F_0)	80	7.0	480	Preparation to move forward
08	370 (12.0 ST above speaker's average F_0)	83	4.5	400	Moving hand forward
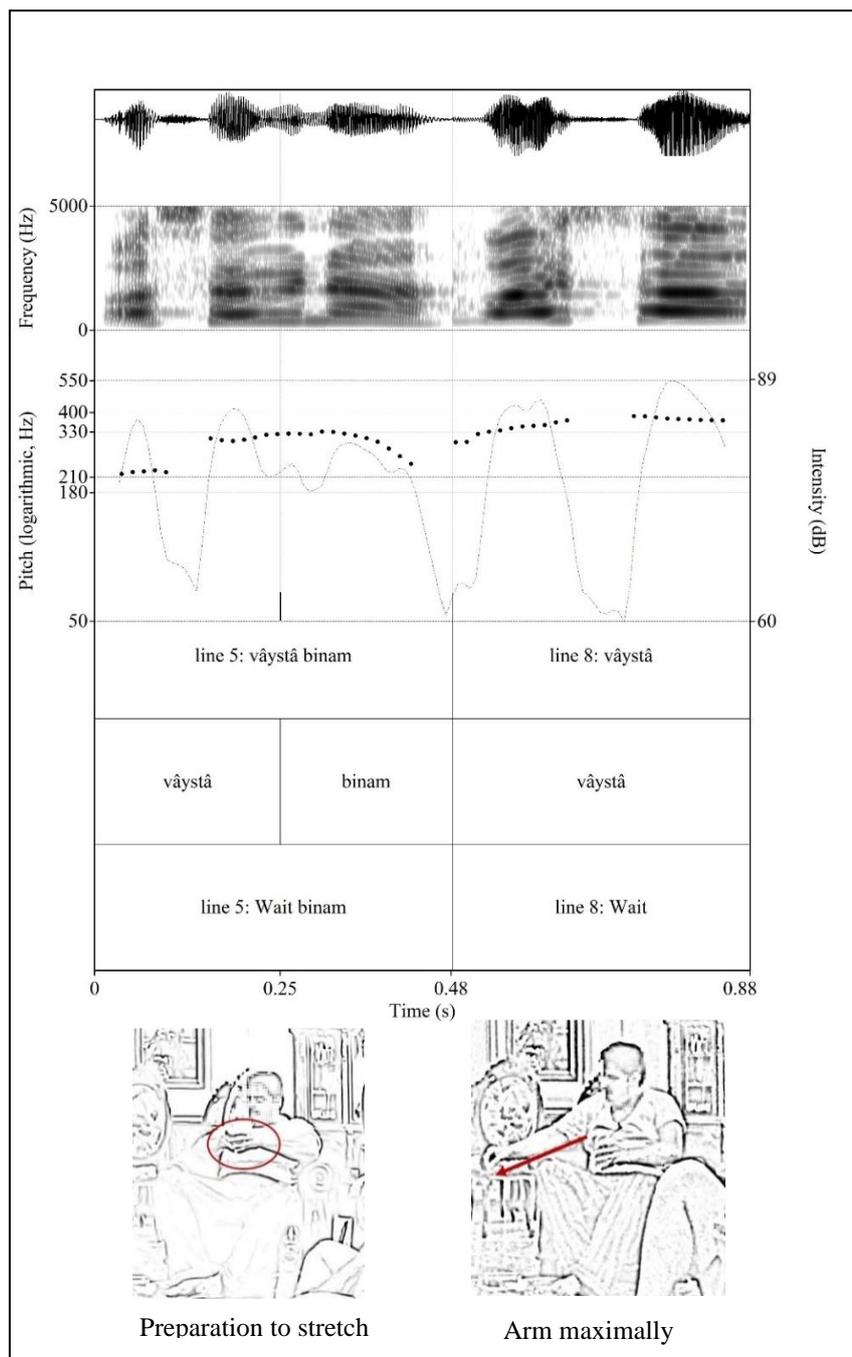
Figure 5.25. Multimodal upgrading in Extract 5.11. The imperative-issuer upgrades the prosody of the subsequent imperative (line 8) relative to the previous one (line 5). His embodied conduct is also upgraded through progressive forward leaning and hand-stretching.

The speaker's bodily conduct is also upgraded. The first version is produced as Ebi switches hands to hold his glass (Figure 5.19). By doing so, he frees his hand that is closer to Haleh's arm. This can be seen as the preparation phase for pulling off the band aid. Following this, Ebi continuously moves closer to Haleh: he stretches his arm and leans his body forwards towards Haleh (Figures 5.20, 5.21) such that by the beginning of the subsequent version, his arm is maximally stretched (Figure 5.22). Despite Haleh's ongoing resistance, he holds his arm maximally stretched for 1.2 seconds (line 10). Thus, with the first version, Ebi starts the forward movement and by the second version he physically reaches the

171

place where he could have implemented his solution, but, as is explained below, Haleh also upgrades her resistance and therefore hinders Ebi's CoA.

The imperative-recipient displays stronger resistance by upgrading the multimodal design of her subsequent rejection tokens (lines 4, 7, 9; Figure 5.26; Table 5.4).[37] In line 4, Haleh's rejection turn (responsive to the initial advice) is composed of only one negation token *na* 'na', followed by an inability account *nemitunam* 'I can't'. The negation token is produced with an intensity of 71 dB and around the speaker's average F_0 (only 1 ST above). Throughout the production of the following account, her turn gets progressively quieter and lower such that the final syllable is barely audible. When Ebi's behavior projects a likely substitution, Haleh's rejection is upgraded (line 7): she produces four rejection particles all very loud and high in her pitch range (84 dB and overall 11ST above Haleh's average F_0). Facing Ebi's continued persistence, her rejection tokens are further upgraded (line 9): they are produced even higher in her pitch range (12 ST above her average F_0). Although the intensity slightly drops, it still sounds very loud compared to her typical conversational intensity level. The number of the tokens are reduced to two; however, each token is produced almost twice longer than each of the previous tokens, making this rejection TCU sound almost as long as her previous turn (see the Praat image in Figure 5.26, below).

Table 5.4. *Summary of the design features of the rejection tokens in Extract 5.11.*

Line	Average F_0 of the rejection TCUs (Hz)	Intensity (dB)	Span (ST)	Duration (ms)	Embodiment
04	291 (1.0 ST above speaker's average F_0)	71	5.4	500	hands in rest position
07	514 (11.0 ST above speaker's average F_0)	84	14.0	920	covers band aid; moves back
09	562 (12.0 ST above speaker's average F_0)	81	6	880	sits further away

Resistance is also upgraded through embodied means. Haleh's initial rejection is produced as she starts to avert her gaze on her mobile phone, a display of returning to her previous activity. This indicates that she is treating the sequence of advice seeking as closed (Rossano, 2012). When Ebi expands the sequence, Haleh's gaze is shifted back at Ebi and she progressively retracts her body away from him and covers the band aid with her opposite hand (lines 6-10; Figures 5.20-5.21). Her gaze displays her re-engagement in the sequence and her body withdrawal is designed to impede the completion of Ebi's CoA. When Ebi persists in approaching Haleh, she lifts her left knee and creates a shield between her arm and Ebi's progressively forward-moving hand (Figure 5.22). And finally, when Ebi reaches the apex of his arm extension and holds his hand ready to implement the solution, Haleh further upgrades her embodied resistance by sliding away from Ebi and sitting on the other side of the room (Figure 5.23-5.24). It is only then that Ebi leans back on his armchair and goes back to sipping his drink. Haleh's gaze is averted back at her phone. At this point the sequence is treated as closed by both participants (Figure 5.24).

---

[37] To compare Haleh's rejection tokens, I have also included her rejection token to Ebi's initial advise in (line 3-4) even though in analyzing Ebi's behavior, only details on his *intervening* turns (lines 5, 8) were presented (not his initial advice, line 3). This is because the nature of rejection and more importantly the means by which Haleh refuses to advance Ebi's CoA is similar in lines 4, 7, and 9. But for Ebi, the lexical design and the nature of his action - advice vs. intervention – are different and not readily comparable.
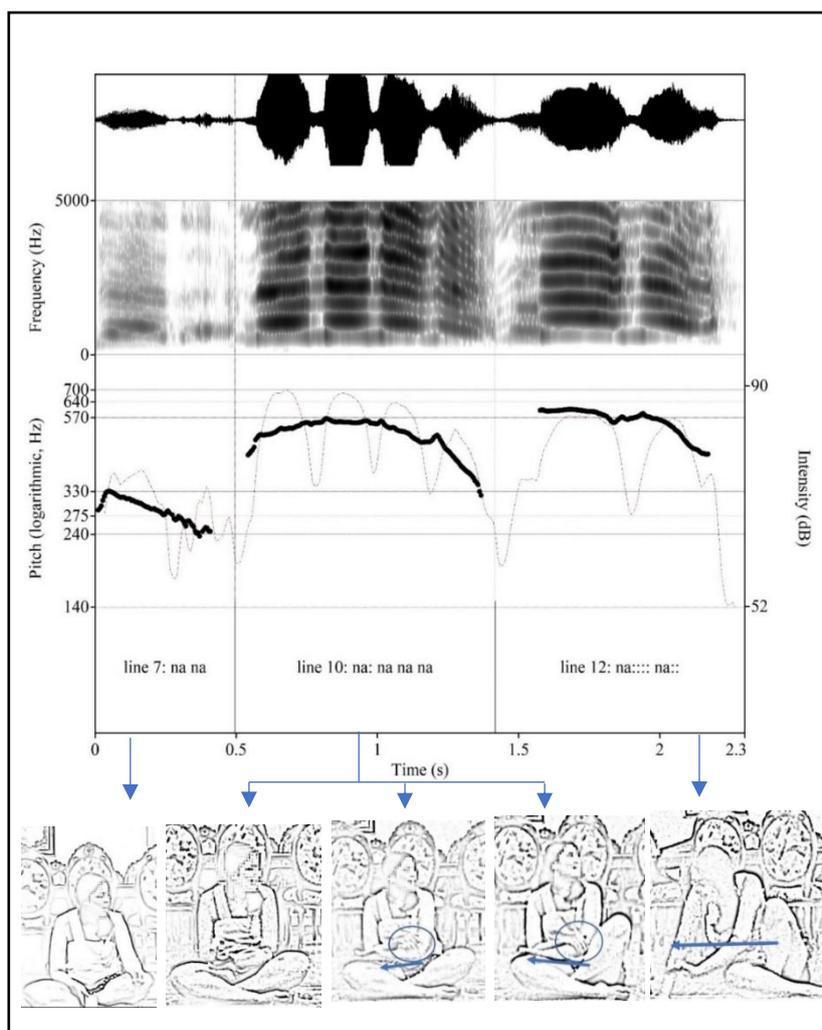
Figure 5.26. Multimodal upgrading of the resisting party's rejection tokens in Extract 5.11.

Now consider the imperative-speaker's agentive stance. The claim is that by substituting for a resisting recipient, the imperative-speaker upgrades his agentive stance compared to the previous group of upgrades where the speaker still provided the recipient a chance for an independent response. Before explaining how, let us have a summary of Ebi's substituting actions.

- Aiming to do the named action itself (line 3, 'pull off'): The imperative pull off is addressed at Haleh, so Haleh is selected as the participant to further the pulling off CoA, but following her rejection, Ebi himself stretches his hand towards the band aid. This projects his substituting himself to perform the named action.

- Doing the reciprocal action to the imperative (line 5, 'go' vs. 'come'): The imperative 'come' is designed to reduce distance between Ebi and Haleh. Again, on the verbal level, Haleh is selected as the participant to make that happen, but on the embodied level, Ebi 'goes' rather than allowing Haleh to choose to 'come'.

Ebi's upgraded agency is manifested in sequential and turn constructional levels.

- Sequential: The imperatives 'pull off', 'wait', and 'come' each limits the recipient's next action to a subset of type-fitting responsive actions (line 3, 5, 8).
- Sequential: Expanding the sequence limits the responsive action's valence to only preferred responses (lines 5, 8).

173

- Turn constructional: Progressively moving forward with hand stretched projects doing the pulling off action and creating a diminished chance of an independent response unless the imperative-recipient upgrades her resistance, which happens in this case (lines 5-9).

In this type of upgrade, the imperatives are re-issued and thus verbally, the recipient is given a response slot; however, on the embodied level, the imperative-issuer appears to be disregarding the recipient's choice. It is the imperative-speaker himself who takes actions to progress his CoA despite the recipient's resistance. The bodily conduct of the imperative-speaker can be seen to be aiming at depriving the recipient of producing a dispreferred response. In the previous group of upgrades, the pursuing party created only practical *limitations* on the speaker's response *choice*. Here, on the other hand, the imperative-speaker aims to *deprive* the recipient of any choice, and in this way, substitution is analyzed as exercising greater agency on the trajectory of the CoA and the recipient's choices.

Of course, the recipient can resist the imperative-speaker's upgraded agency by upgrading her own resisting action, in which case, participants compete for agency to affect the trajectory of the initiated CoA. In this extract, Haleh won this competition despites Ebi's upgraded agency (eventually, the band aid remained on her arm). A full analysis of recipients' agentive stance is left for a separate research project.

5.5. Continuum of Agency and a Summary

The previous sections compared three overall ways in which an imperative sequence may unfold after the recipient's noncompliance: the imperative-speaker may abandon the sequence; they may pursue compliance with downgraded subsequent versions; or they may pursue using multimodally upgraded subsequent versions. Now, a continuum of agency for an imperative-speaker in such sequences can be introduced (Figure 5.27). From left to right, the imperative-speaker's agency is progressively stronger.
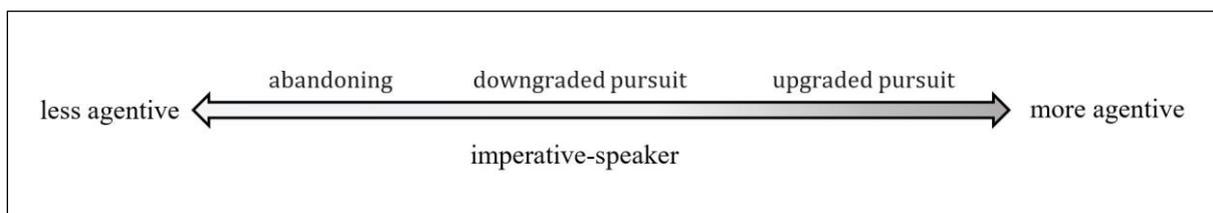
Figure 5.27. The agency continuum.

The evidence for this continuum comes from (a) the norms and normative expectations that are well-established in CA, such as the conditional relevance rule and preference, and (b) some common conceptualizations of agency in the literature: an agentive party is an entity who 'chooses' to perform 'actions' and whose actions 'influence others and the world around', and have 'control' over their actions (see e.g., Kockelman, 2007; Duranit, 2004; Enfield & Kockelman, 2017).

By issuing an imperative that is designed to get the recipient to do an action, the imperative-speaker displays some level of agency. This is because by virtue of the conditional relevance rule, the speaker makes the recipient normatively 'obligated' to produce a type-related response. If the speaker abandons the sequence after a dispreferred response, she effectively ceases to get the recipient to perform the named action, and thus they tacitly accept the recipient's choice between a preferred and dispreferred response types. So the imperative-speaker's agency is only manifested in issuing the imperative and it is relatively weaker than pursuing compliance.

If, however, she chooses to pursue compliance after the recipient's noncompliance, she expands the sequence and limits the valence of the expected response to only preferred. This is because she tacitly treats noncompliance as nonacceptable. So overall, pursuing a preferred response displays a stronger agentive stance than abandoning the same sequence.

174

If the imperative-speaker chooses to pursue compliance, she can do so through an upgraded or a downgraded multimodal Gestalt. An upgraded pursuit can be seen as more agentive compared to a downgraded one. This is evident if we consider two overall patterns created by upgraded pursuits in this collection. In the first one, the pursuing party's embodied Gestalt (that often includes bodily actions that approach the recipient) limits the recipient's personal space. So as well as a normative limitation produced by the pursuit turn, these bodily actions limit the recipient's 'freedom' or choice in performing an independent action. Therefore, due to their constructional details, upgraded pursuits are indicative of stronger agency for the pursuing interactant than downgraded ones.

Through the second group of upgraded Gestalts, the pursuing party's bodily actions project performing the action named by the imperative. In other words, although on the verbal level the imperative-speaker repeats the initial imperative and verbally creates another slot for the recipient's response, on the practical level, he himself sets off to perform actions that complete his CoA despite the recipient's resistance. This type of upgraded Gestalt aims to deprive the recipient of an independent response and from this perspective, they can display even stronger agency than those upgrades that 'merely' limit the recipient's choices.

5.6. Discussion

Implicit in the organization of adjacency pairs is that the sequence initiating participant makes conditionally relevant for the coparticipant to perform a next type-fitting action and in this way, she puts some agency on display (see Sidnell, 2017, on on- vs. off-record methods). But by naming an action for a coparticipant to perform, as the imperatives in this chapter do, a participant displays a stronger agentive stance. A participant's agency is upgraded to even a more explicit level when she expands the sequence and resorts to practices to overcome the coparticipant's resistance against performing the named action. In this chapter, we analyzed four levels of agentive stance for an imperative-speaker: (a) only issuing imperatives, (b) pursuit with downgraded turn design, (c) pursuit with upgraded turn design when the pursuit turn creates another slot for the recipient's independent response (d) an upgraded pursuit when the chances for the recipient's independent response are diminished.

Agency was approached from a multimodal perspective. This study expands the literature on agency which has so far focused more on the verbal turn design (Auer and Lindström, 2021; Clayman 2013; Etelämäki, et al., 2021; Heritage & Raymond, 2012; Raymond, et al., 2021; Rossi & Zinken, 2016; and Thompson, et al, 2015) and more recently on the on the embodied aspect (Deppermann & Gubina, 2021; Tuncer & Haddington, 2019). On the verbal level, Auer and Lindström's (2021) study on reduced versus full evaluative turns shows that reduced evaluations are both lower in agency and projective of a sequence closure whereas full evaluations indicate higher agency and project sequence expansion. On a multimodal level, this study showed that downgraded Gestalts display weaker agency and project sequence closure and upgraded Gestalts can indicate stronger agency and are associated with sequence expansion.

More nuanced than that, a multimodal perspective made it possible to differentiate among various degrees of agency displayed in upgraded turns. Although on the basis of the verbal modality only, in all the provided examples, the sequence is expanded by means of recycling the initial imperative, a complicated picture of agentive relationship between the two participants can be sketched when other modalities are taken into account. Upgraded pursuit turns can be grouped based on whether they afford or diminish chances of an independent response for the recipient. The chapter follows Kendrick and Drew (2016) and Kendrick's (2021) mode of analysis. In the first study, for example, the researchers have focused their attention on different embodied and verbal methods by which a participant (Self) can recruit another participant (Other). They organize the methods of recruitment based on the participant who initiates recruitment. Similarly, this chapter was in part focused on the participant who is projected to complete the imperative CoA, the pursuing party versus the recipient.

The findings have implications in terms of the distribution of the components of action production discussed in Goffman's (1981, p. 144) footing model: *animator* (the agent who implements an action),

author (the agent who decides what should be done and how), and *principal*, the agent who is accountable for the action. Take the example in which the pursuing party offers the coparticipant a fig by extending her hand with the fig in it (Extract 5.10). As discussed, the pursuing participant creates new interactional slots in which the co-participant himself can complete the course of action. It is the co-participant themselves who should carry out the actual acceptance; it also him that can decide how he is going to accept the fig; for example, if he is going to take the fig by hand or if he is going to open his mouth so that the imperative-speaker puts the fig into his mouth, and with this choice comes accountability. Therefore, in light of Goffman's footing model, in this type of upgrading, it is the imperative-recipient who is regarded as the animator of the action and also partially as the author and principal.

The distribution of roles is different in substituting. Consider the example of substitution in which the pursuing participant aims to remove the co-participant's band aid (Extract 5.11). Here the pursuing party upgrades his roles to the animator, author, and principal of the action: he 'decides' that the band aid should be removed and how it should be removed (author); he aims to carry out the removing action himself (animator); and he is potentially accountable for any consequences (principal). Therefore, the sense of upgrade that we perceive is in part an upgrading in the distribution of participants' agentive roles: who does what and how. A similar distribution can be explained using Enfield and Kockelman's (2017) agency model.

The practices illustrated in this chapter were based on some strategies that the pursuing party takes to further their CoA, but the recipient can of course play an agentive role in affecting the trajectory of the sequence. Practices by which an imperative-recipient counters the effect of the speaker's agency should be investigated in future research. Here suffice it to mention that a recipient's quick 'submission' saves the interaction from slipping into a conflictual state even when the pursuing party resorts to rather 'stronger' agentive moves. However, multiple rounds of resistance against the imperative-speaker's moves (even when they are not quite as agentive) risk the interaction to turn into a 'battle ground' where participants compete for agency. Compare Extract 5.10 (the fig example) with the first example of the chapter (the cat example). In the fig example, the recipient's multiple rounds of resistance tilt the interaction towards a trajectory in which a third party intervenes with a mock threatening action (holding a fruit knife towards the resisting party). In the cat example, on the other hand, even though the imperative-speaker forces the recipient's arm away, the recipient submits 'quickly' by straightforwardly allowing her hand to drop on the sofa. This makes the interaction in the cat example look less conflictual compared to the one in the fig episode. Therefore, in analyzing conflictual encounters, the recipient's response types are an indispensable part of the picture.

Mock aggression, as the fig example shows, is one resources that can be employed by participants when a pursuit sequence is on a trajectory of escalation. The following chapter examines this phenomenon and its interactional affordance in and beyond pursuit sequences.

Chapter 6. Mock Aggression: Navigating Affiliation and Disaffiliation in Interaction

6.1.   Introduction

Social interaction is home to a complicated scene of both bond-strengthening and bond- threatening actions between social units. Despite the normativity of avoiding explicit collision (Goffman, 1971/2010, Sacks, 1987), we do not always affiliate with each other. Moments of conflict can arise when interactants insult, accuse, tease, criticize, complain to, or argue with one another. These and other negatively-valenced social actions do occur; however, to sustain a state of equilibrium in terms of social order and social solidarity, members of society are bound to have methods at their disposal to manage these conflictual encounters and navigate affiliation and disaffiliation in their interaction. The question is what those methods are and how they work to solve the social organizational problem of conflict management.

To tackle this question, this chapter takes a wider view on domestic clashes. It goes beyond pursuit sequences and explores a range of different disaffiliative sequences such as insults, accusations, disagreements, and the like. It locates pursuits in the face of resistance among this larger family of disaffiliative interactions. Having seen how what could have been just a momentary clash may get expanded (Chapter 2), and escalated (Chapter 3), we now return to the bigger picture on conflictual interaction, sketched in Chapter 1, and ask how participants may get out of an escalated disaffiliative sequence. We focus on one particular practice which I have called *mock aggression* (Afshari Saleh, 2020): the mimetic enactments or actual deployments of aggressive physical actions (e.g., pinching, slapping, etc.) which are not designed, on the part of the performer, or oriented to, on the part of the recipient, as physically threatening. The examples in this chapter are selected from other sequence types, but at the end of the chapter, an instance of mock aggression in a pursuit sequence is provided. Consider an initial specimen of the phenomenon.

Prior to this extract, Sabâ, the performer of an iconic strangling, had explicitly expressed her worry about some potential bad news and asked Kiân, her husband, to close the news webpages so that they do not get informed about it. Kiân, the recipient of the gesture, has kept the webpages open. When the news is posted, he starts reading it to Sabâ bit by bit with all details. The gesture is performed after the last piece of the news is read.

Extract 6.1. RAY027c, 11.23

```
    01              ((KIÂ finishes reading the bad news))

    02      SAB     <<clenched teeth>#*±cherâ cherâ>        ±
                                       why     why
                                       why why
->                                     ±strangling gesture±
                                       *eyes wide open-->
```


#6.1

```
03      KIÂ    <<smiling>bi-â      man-o      khafe      kon>
                          IMP.come 1SG-OBJ.MAR strangled IMP.do.2SG
                          come strangle me

04      KIÂ    <<laughing>be man che  dige>=
                          to 1SG what PTCL
                          it's got nothing to do with me

05      KIÂ     =*•hehehe•
        sab    -->*
        sab          •smiles•
```

Following the last bit of the unwelcome news, read by Kiân, Sabâ brings her hands forward towards Kiân, keeps them at her head level, with palms facing one another and fingers tense and bent at knuckles (Figure 6.1). This conveys a mimetic enactment of a strangling action. This relatively extreme embodiment (compared to a slapping gesture, for example) can be occasioned by Kiân's persisting in not only keeping the webpages open, but also reading the news (and not just the headline or the gist of it, but in detail) despite his wife's explicit request for not doing so. Concurrent with her gesture, Sabâ stares her gaze at Kiân, with eyes more open than her normal and asks Kiân to provide an account by repeating 'why' twice while keeping her teeth clenched. Given Kiân's blunt non-compliance and Sabâ's enactment, which is addressed at him, the interrogative 'why' can be analyzed as a request for an account for Kiân's news reading rather than an account for the occurrence of the news (line 02). With her bodily behavior (both manual and facial) and her account-soliciting question (see Bolden & Robinson, 2011), Sabâ objects to Kiân's news reading.

In light of Schegloff's (2007, p. 9) distinction between social actions and the vehicles (formats) through which they get performed, Sabâ's action in line 2 has at least two components: the action of objecting, and its format (i.e., her interrogative and embodiment). Kiân first orients to Sabâ's manual embodiment, a component of Sabâ's *objection format*, by saying 'Come strangle me' (line 03). This shows he recognizes Sabâ's embodiment as a depiction of a strangling movement. His directive is ironic: although its imperative format is generally associated with initiating actions such as requests, it is not an initiating action, nor does it make any compliance or non-compliance conditionally relevant. It is a counter-objection, responsive to Sabâ's gesture. Then he orients to Sabâ's *objection action* by denying accountability for the *occurrence* of the news (rather than providing an account for *reading* it; line 04).

In both his turns, Kiân is either smiling or laughing, which shows his orientation to the gesture as physically non-threatening (lines 03-04). Following Kiân's no fault-account, Sabâ stops her tense facial movement and smiles (line 05). This strangling enactment is an example of what I call *mock aggression*: the practice is recognizable to its recipient as representing some aggressive conduct, but it is not oriented to as a serious physical threat. The practice is designed with inhibitory features (see below), for example, here the performer is leaning back while the gesture is being performed (Figure 6.1) and drops her hands as soon as her verbal turn is over (line 02). More on the details of the practice's design features is discussed in §6.5.

This mock strangling gesture is an example of what is titled as mock aggression: the embodiments which, in one way or another, appear aggressive, but are not designed to be, or oriented to as, serious physical threats. This chapter aims to show that mock aggression, when occurring between intimate interactants, has double affordances: it negatively sanctions social transgressions and at the same time facilitates participants' exit from a disaffiliative interaction and thereby contributes to maintaining their social bonds. Before showing evidence on this claim, first a brief review on social transgressions and their connection to disaffiliation in interaction (§6.2) and then a comparison between mock and real aggression (§6.3) is provided.

6.2.   Social Transgressions

Social transgressions, that is, deviations from "relevance rules" or "normative patterns of conduct" (Robinson, 2016, p. 7) can be a trigger point for conflictual interactions. Normative patterns of conduct

comply with participants' common-sense and mutual expectations (Garfinkel, 1967) and are guided by rules of conduct, that is, obligations (in the sense of moral responsibility) and expectations apt for the social encounter (Goffman, 1967). Transgressing normative behavior can take shape through departures from what Goffman describes to be moral principles, that is "norms felt to be desirable intrinsically" (Goffman, 1971/2010, p. 96) such as insulting, a deviation from 'courtesy and respect'. Transgressions can also occur through violations from conventionally expected conduct, such as behaving 'too polite' and avoiding getting personal among close family members (cf. Garfinkel, 1967, p. 47). Performing dispreferred actions (see below), such as telling participants bad news or rejecting a proposal is another method which can generate social transgressions (also see Robinson, 2016, pp. 17-27 on relevance rules). Regardless of how exactly they are actualized, these deviations have in common a socially untoward character, engendered by interlocutors' breach of what Garfinkel (1967) describes to be a 'trusted' expectation in performing proper social actions. Breaches of norms can be seen to be at best a 'momentary lapse' in exercising social responsibility or at worst a 'faulty character' (Goffman, 1971/2010, p. 99) and as such they can occasion some kind of negative sanctioning.

As reviewed in Chapter 1, interlocutors can negatively sanction inapt conduct through a variety of methods such as soliciting accounts (Bolden & Robinson, 2011), issuing accountability-driven imperatives (Kent & Kendrick, 2016), casting sanctioning looks (Kidwell, 2005), among others. Sanctions can be disaffiliative in that they are typically designed to hinder the sanctioned course of action (see Clayman, 2002) and contrary to affiliative actions which endorse interactants' conveyed affective stance (Stivers, 2008), they can be destructive of social solidarity. Stated differently, even though negative sanctions promote social order; they can demote social solidarity between the sanctioning and sanctioned parties as they can be a threat to the sanctioned party's face (cf. Brown & Levinson, 1987). However, if affiliation between the two parties remains in equilibrium even after negative sanctions, there must be some methods to restore the threatened solidarity.

The literature on preference organization in CA has shown a number of practices which mitigate the discordant effect of disaffiliative or transgressive conduct. Heritage (1984) illustrates that dispreferred actions, which are typically disaffiliative in character, are systematically withheld, mitigated, or accompanied with some "remedial" work (Goffman, 1971/2010) like giving accounts, apologies, and excuses; however, preferred actions, which are affiliative in character, are typically performed straightforwardly. These practices can balance out the affiliation-disaffiliation scale in the interaction. Despite a rather extensive literature on the methods by which the transgressive or accountable party moderates disaffiliation in the interaction, there is less research on how the party who is at the receiving end of transgression maintains social accord. So far, the literature mainly shows that responses to breaches of normative behavior can be categorized as utterly disaffiliative with the transgressive conduct at one extreme or affiliative at the other. Although some scholars have shown that disattending the transgression is also a practiced method of dealing with transgressions (e.g., Jefferson, Sacks, & Schegloff, 1987 and Drew, 1987), the literature has mainly postulated a categorical affiliative versus disaffiliative perspective on most social practices.

A similar categorical approach is also adopted about the embodied resources for navigating affiliation and disaffiliation. Affiliative nods in mid-telling of a story (Stivers, 2008), gaze aversion projecting a dispreferred response (Kendrick & Holler, 2017), and pre-beginning frowns foreshadowing disaffiliative turns (Kaukomaa, et al, 2014) are among such embodied resources. However, as this chapter shows, this categorical perspective cannot always explain how social solidarity is managed in interaction.

6.3.   Mock Aggression vs. Aggression

Mock aggression is a kind of embodiment (see Mondada, 2014 on embodiment). A slapping action (hitting an interlocutor's target body part with an open hand), which would be considered as an instrumental or practical action, is one kind of embodiment. So is a slapping gesture, such as quickly raising an open palm, holding it above the interlocutor's target body part and retracting it without making any contact. This kind of gesture, that is, miming a pattern of action referentially associated with some practical action, is called an "enactment" (Kendon, 2004, p. 160), a "mimesis" (Streeck, 2009b, p. 145), an "iconic gesture" (McNeil, 1992, p. 107), or a "kinetograph" (Ekman & Friesen, 1969,

p. 70). Gestures might only include *elements* of the image or action they represent (Kendon, 2004, pp. 307-9). For example, in the slapping gesture described above, only an open hand raise is acted out; nevertheless, the enactment even without the hand's downward movement can, in some contexts, depict a slap. This is because of our understanding of the context (Kendon, 2004, p. 161) and projective potentials (Streeck, 2009a) of gestures.

Mock aggression is in sheer contrast to real aggression (cf. Whitehead, et al., 2018, for a CA approach to violence). Whereas the latter has serious disaffiliative implications (Collins, 2008; Goodwin, 1994), what I refer to as 'mock aggression' is relatable to humor, play, and play-fighting (Morreall, 2016). Goffman (1974/1986) describes playful activities as one of the central kinds of social activities where "a given activity, one already meaningful in terms of some primary framework, is transformed into something patterned on this activity but seen by the participants to be something quite else" (pp. 43-44). Playful embodiments do not occasion the same responses which their serious counterparts would normally do (Sacks, 1992, pp. 671-2). As with the verbal level, where specific turn-design features can work to display the non-seriousness of an utterance (Schegloff, 2001), at the embodied level, too, a Gestalt-like complex of detailed features (Mondada, 2014) including facial expressions, the phonetic, prosodic, and verbal design of the co-occurring talk, and the details of bodily movements contribute to conveying an embodiment as playful. To some anthropologists and ethologists, playful fight-like actions carry a "this is play" (Bateson, 1972, pp. 177-93) or "never mind, it's all fun" (Darwin, 1872/1998, p. 66) message. Through this message, participants get engaged in 'play': an "activity, motor or imaginative, in which the center of interest is process rather than goal" (Miller, 1973, p. 97). For play-fighting, Bateson (1972, p. 179) describes such a process as "an interactive sequence of which the unit actions or signals [are] similar to but not the same as those of combat". In other words, participants get engaged in only a "display of fighting", simulating fighting acts, inhibiting fighting outcomes (Sutton-Smith, 2001, p. 23; see also Fagen, 1981).

Rather than relying on metacommunicative messages or abstract signals, the current chapter uses empirical data to unpack the detailed physical and sequential features of some only ostensibly aggressive embodiments. First, the identified types of the ostensibly aggressive embodiments are outlined (§6.4) and then their common design features are listed (§6.5). Next, we examine the social action that mock aggression can perform when it occurs after social transgressions (§6.6). The affiliative potentials of the phenomenon is illustrated in §6.7. We then analyze an example of mock aggression in pursuit sequences (§6.8), and finally the implications of the findings in terms of our understanding of (dis)affiliation in interaction are discussed (§6.9).

6.4. Identified types of the phenomenon

The following table presents the different types of samples in the collection (Table 6.1).

Table 6.1. Types and frequency of mock aggression samples in this collection.

Embodiments	Number
Slapping	10
Biting	7
Punching	5
Pinching	4
Pulling	3
Kicking	2
Pulling Hair	2
Strangling	2
Throwing an Object	2
Clawing	1
Smacking	1
Wrestling	1
Total	40

To re-state, the collection based on which the analysis in this chapter is conducted includes also other sequence types than pursuits. The collection is made based on only 27 hours of the whole corpus. To limit the scope of the chapter, the collection was narrowed down to samples which were targeted at the co-participants (rather than at oneself or an absent third party).[38] Also, only cases in which the embodiments were responsive to some prior action were included. The samples could be performed with or without contact with the recipients' bodies so long as no orientations to any real physical threat was observed in the participants. 40 such cases were identified

The collection includes both actual deployments of aggressive-looking actions and their mimetic enactments or, to borrow Kendon's (2004, p. 307) terminology, the "real ostension" and "virtual ostension" of aggressive actions, respectively. In both types, the bodily movements are designed such that real harm is avoided and co-participants orient to them as physically non-threatening. The following section shows some of these design features.

6.5. Design Features

The practice is typically recognizable to be playful or not an actual physical threat mainly through its design. Three general sets of features which mock aggression is generally designed with are a) an over- or underplayed quality of bodily movements compared to the typical corresponding aggressive counterpart; b) an inhibitory control by which physically 'dangerous' consequences are inhibited by the performer (cf. Fagen, 1981); and c) the integration of multimodal resources which are typically associated with affiliative displays (e.g., smiles[39]) with those which represent some disaffiliative conduct, for example, a slap (see Mondada, 2014, p. 139, on multimodal Gestalts). Not all samples include these features altogether; therefore, this section inevitably shows only some of these features. More will be unpacked throughout the chapter under different extracts.

Extract 6.2 is an excerpt from a pursuit sequence in which Kiân insisted that Sabâ (his wife) play red hands with him several times (to remain focused on the design features of the phenomenon, the pursuit sequence is not presented).[40] Sabâ rejected Kiân's repetitive requests for different reasons, the last of which was the noise the game could make on the research microphones. Kiân claimed that the game would not produce any noise, so they played once (Figure 6.2), but as the beginning of the extract shows, the game makes a slapping sound.

Figure 6.2. Participants in Extract 6.2. Saba (left) and Kian (Right) are playing red hands.

---

[38] This happens in the context of story- or trouble-telling when, for example, the performer of mock aggression is complaining about an absent third party.

[39] I acknowledge that not all kinds of smile or laughter display affiliation, but for all practical purposes, other kinds of these practices are not considered in this paper.

[40] Player A, holding her hands palm up, tries to hit the back of Player B's hands, which are put palm down on A's hands.

Extract 6.2. RAY027e, 04.10

```
01                 ((slapping sound))

02                 (0.5)

03      SAB     khob #+mi-khor-e    +he±hehehe#       ±
                PRT      PRG-hit-3SG
                khob⁴¹ it'll hit
        kiâ            +strtch arms+
        sab                            ±slps K's hnd±
        fig         #6.3                     #6.4
```

#6.3

#6.4

```
04             #+    (1.2)    +
        kiâ    +slps S's arm+
        fig    #6.5
```

#6.5

```
05             #±    (1.2)                   ±
        sab    ±moves fist twds K's chest±
        fig    #6.6
```

#6.6

```
06             #+    (0.7)                      +
        kiâ    +mimes pinching near S's knee+
        fig    #6.7
```

#6.7

---

⁴¹ Translating this particle in this position needs an investigation of its use in this sequential environment, but it seems to function as a counter-positional token here.
```

```
07          #±     (0.6)                    ±
     sab    ±mimes pinching near K's knee±
     fig    #6.8
```


#6.8

```
08          #+     (2.2)      +
     kiâ    +pulls S's hair+
     fig    #6.9
```


#6.9

```
09          #±     (1.8)      ±
     sab    ±pulls K's nose±
     fig    #6.10
```


#6.10

```
10   KIÂ    *e:[:::::::
            cry.INTJ
            E:::::::::
            *crying face-->
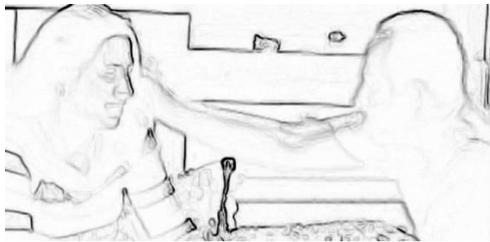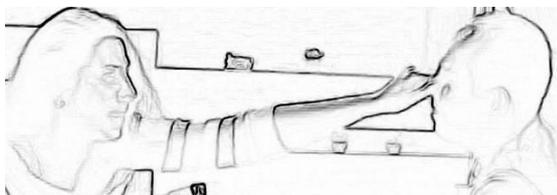
11   SAB       [he±hehe
               ±....-->

12   SAB    khob #±bâsh-e* bash-e ±
            Ok    be-3SG  be-3SG
            OK there there
              -->±taps KIÂ's nose±
     kiâ              -->*
     fig         #6.11
```

#6.11

After a slapping sound is produced, Sabâ repeats her reason for not playing red hands (line 03). Despite this, Kiân stretches his hands towards Sabâ, holding them in the play position, embodying a request from Sabâ to continue playing (line 03). Thus, the sequence starts with an embodiment and this occasions Sabâ's bodily method of doing rejection; she slaps Kiân's hand away (line 03). Following the embodied request-rejection sequence, they get engaged in a whole sequence of mock aggression: Kiân slaps Sabâ's arm (line 04); Sabâ enacts a punch towards Kiân's chest (line 05). He performs a mimetic pinch near Sabâ's leg (line 06); she reciprocates the same gesture near his leg (line 07). He pulls her hair (line 08); she pulls his nose (line 09). Then Kiân 'does being a child' by making a 'crying face' and producing a mock cry sound (line 10). Sabâ laughs at this (line 11) and 'does being an adult' by tapping Kiân's nose twice while comforting him (line 12: 'there there'). Reciprocating mock aggression then creates a 'pathway' from the disaffiliative request-rejection sequence to an affiliative mock cry-sympathy. (Two other common ways by which mock aggression contributes to more affiliative sequences are elaborated in more detail in §6.7)

This extract shows different ways by which the practice can be recognized as non-serious.

- The practice is underplayed by reducing the extent of body contact. Here, the slap is not made with the full palm (Figure 6.4).
- In the case of contact, the intensity or the speed of the bodily motion is not high enough to cause any serious harm. Here, the slap does not even exert enough force to make its recipient's body orientation to change. The recipient's body remains almost still; also, note the recipient's smile as she receives the slap (Figure 6.5).
- The practice is underplayed by not completing the trajectory of all the necessary phases for an aggressive embodiment to be performed; that is, the movement excursion of body parts is cut short. This can be achieved by not completing the stroke phase of the gesture/embodied action (cf. Kita, et al. 1998, on the phases of embodied actions). Here, for example, the arm is stretched in a punching gesture towards the recipient, but it stops before touching the chest (Figure 6.6). Or sometimes only the preparation phase is produced, in which case the preparation phase itself projects the stroke.
- The movement excursion of body parts depicts a complete version of the aggressive counterpart (e.g., for a pinching gesture, the thumb and index finger are drawn close to one another, then a twisting movement is exerted on them); however, the practice is pantomimed off the recipient's body (Figures 6.7 and 6.8).
- Participants inhibit fighting outcomes by choosing an 'inefficient' target body-part. For example, given the bony structure of a knee, inducing serious harm by pinching someone on their flexed knee is harder than say the arm since pinching requires some flesh to be gripped by fingers and twisted or squeezed (Extract 6.3).
- Contradictory design features which are typically associated with affiliative and disaffiliative actions are juxtaposed to form a Gestalt-like whole (e.g., Extracts 6.3, 6.5).
- Sometimes the movements are overplayed, producing a theatrical effect and contributing to the practice being recognized as playful (Extract 6.5).
- The embodiment is designed to be projectable. In this way, the performer gives time and preparation chance to the recipient to take measures to inhibit harm or to block the occurrence of the embodiment (e.g., by withdrawing the body or blocking the projected action; Extract 6.4).

The phenomenon can take shape in both real (e.g., extract just above, lines 04, 08, 09) and virtual (e.g., extract just above, lines 05, 06, 07) ostension of aggressive actions. In both ways, mock aggression achieves the same social action.

6.6. Social Action

I will now show that mock aggression is used as a resource to negatively sanction social transgressions between intimate interactants. I distinguish between the *social action* that is performed by employing the practice and its *interactional outcome*. My main claim is that, despite its ostensibly aggressive form, the practice functions as an exit in the path of a disaffiliative interaction, helping participants to move away from disaffiliation. I claim that this interactional outcome is achieved through the playful design of the practice. This section focuses on the action (negative sanctioning) and the next section focuses on its affiliative outcomes (§6.7).

Mock aggression implements a type of social action: it can negatively sanction a transgression. There is a wide range of actions which performers orient to as sanctionable. In Extract 6.3, for example, the sanctionable is an accountably inappropriate word (line 6: 'ass') used in front of the research camera. While setting up the camera, the researcher had told the participants that they could behave in *any* way (formal, informal, etc.) that they would normally interact with one another. Niyushâ (Mâhnush's daughter) is sitting with her back facing the camera. Ashkân is Mâhnush's partner and is sitting very close to her. 'She' in line 07 refers to the researcher.

Extract 6.3. RAY030a, 08.24

```
    01    MÂH          mâmân cherâ posh neshast-i [(haha)
                       mom   why   back sit-2SG
                       sweetie why are you sitting backward


    02    NIY                                         [(   )


    03    NIY          bekhâterike ghiâf-am ne-mikh-âm [ (0.8)  ] dide\
                       because     face-my NEG-want-1SG            seen
                       Because I don't want my face    (0.8)     seen


    04    ASH                                         [hehehehe]


    05    ASH          â::: mig-e   ghiâf-am ne-mikh-âm\ hehe=
                       INTJ say-3SG face-my  NEG-want-1SG
                       Â::: she says I don't want my face\hehe


->  06    ASH          =kun be tas•*vir ne±shest-e±he*=
                        ass to film     sit-PRF.3SG
                        she's sitting with her ass to the camera
          mâh                        *eyebrow raised--*
          mâh                        •smiles-->>
->        mâh                           ±pinches±


    07    ASH          =gof    esh[kâl na-dâr-e
                        say.3SG problem NEG-have-3SG
                        she said it's not a problem


    08    NIY                     [language please  ((said in English))
```

The sequence starts with Mâhnush's question addressed to Niyushâ (line 1: 'sweetie why are you sitting backward'). Niyushâ gives an account for her backward sitting (line 3: 'I don't want my face (0.8) seen\'). Ashkân marks Niyushâ's account as new information by his change-of-state token *A:::,* followed by quoting her directly (line 05: 'A::: she says I don't want my face\hehe'). The quotation is

185

cut-off by Ashkân's laugh, which is immediately followed by his next TCU, explicating the upshot of Niyushâ's account (line 06: 'she is sitting with her ass to the camera'). Right after he uses the word 'ass', Mâhnush flashes her brows, orienting to Ashkân's lexical choice as inapposite (hence, unexpected), smiles widely, and slightly moves her hand which is already very close to Ashkân's knee towards his knee and produces an iconic pinch on his knee (line 06). With this, she negatively sanctions Ashkân's 'inappropriate' word. Ashkân's orientation to Mâhnush's gesture as a form of negative sanction is clear in his next TCU where he accounts for his language (line 07: 'she [the researcher] said it [our own normal way of behaving] is not a problem'). Quoting from the researcher to account for his turn design shows Ashkân's understanding of Mâhnush's gesture as sanctioning using the word *in front of the research camera* rather than using the word per se. In overlap with his account, Niyushâ, too, verbally sanctions Ashkân's lexical turn design by directing him (through a noun phrase) to watch his language (line 09: 'language please'). This is further evidence for Ashkân's interlocutors orienting to his lexical design as inappropriate.

In terms of its social action, therefore, the phenomenon can be responsive to a transgression and negatively sanction it. The practice is a method through which participants police one another's social behavior, highlight 'red lines', and re-enforce the broken norms. As we saw above, just after the practice, the transgressor produces some remedial work (Goffman, 1971/2010) by giving an account for his accountably inapt behavior in front of the camera.

Design-wise, the pinch is underplayed: the performer's hand only slightly touches the recipient's knee, which as explained before, is not even the most efficient target body-part for a pinching action. Also, the pinch (typically associated with aggression) is concurrent with the performer's broad smile, which can be indicative of positive stance. This integration of *contradictory* design features into a multimodal Gestalt (Mondada, 2014), whose holistic interactional function is different from the separate effect of its individual parts, conveys the embodiment as non-serious. Through this non-serious design, the performer achieves to disaffiliate with the transgression while inhibiting disaffiliative outcomes in her interaction with the transgressor. More on the affiliative outcome of the practice is presented below.

6.7. Orientations and Affiliative Outcome

In the previous sections, the design features and social action of the phenomenon were identified. In this section, an interactional affordance of mock aggression is discussed: the practice can provide participants an opportunity to move away from a disaffiliative interaction. Put differently, mock aggression can offer an 'affilliative exit' off the path of a disaffiliative interaction. Depending on how the recipient orients to mock aggression, this affiliative outcome can be achieved in different ways. Below two common ways observed in the current collection are explained: orienting to the playful aspect of mock aggression (§6.7.1), renewing disaffiliative actions in less disaffiliative design (§6.7.2).

6.7.1. Orienting to the Playful Aspect of Mock Aggression

The affiliative potential of the practice may be realized through its Janus-faced quality. Being associated with both antagonism (through its similarities with fighting patterns) and play (through its physically 'non-threatening' design, see Holt, 2013, on playful/non-serious interaction), mock aggression has double affordances. When facing a sanctionable which is disaffiliative but non-serious (e.g., a playful criticism), the recipient of the transgression can be in an interactional predicament as to which aspect of the sanctionable should be addressed (e.g., the disaffiliative criticism or the affiliative play). The double associations of mock aggression make it a double-faceted resource which can simultaneously *sanction* the transgression through its similarities with aggression and *reciprocate* the transgressor's playfulness through its similarities with play. The recipient of the practice can orient to either of these two aspects but as this collection shows, laughing and more playful conduct are frequent responses to the practice. In this way, mock aggression offers an affiliative exit through which participants can continue their interaction, attending to the non-serious aspect of one another's turns, rather than continuing on the disaffiliative highway, as it were. Prior to Extract 6.4, Farnâz started a playful criticism on Amir's excessive suspicion about everyday sounds (e.g., car alarms, doors closing, etc.). The extract starts with Farnâz's last exaggerated bit of criticism where she explains that a popular song about suspicion (by a Persian singer, Shâdmehr) is sung especially for Amir.

Extract 6.4. RAY018c, 17.30

```
01    FAR    <<giggles>un   sher  az shâdmehr hast   ke
                       that song  of name     be.3SG that
                       there is this song from Shâdmehr which says

02           mashkuk-am       [mashkuk-am>
             suspicious-be.1SG  suspicious-be.1SG
             I'm suspicious I'm suspicious

03    AMI                             [((giggles))

04    FAR    •un-o          barâ  to  khund-an
             that-OBJ.MAR   for   2SG sing-3PL
             they've sung it for you
             •smiles-->

05           *    (4.0)   •*    (0.4)     *
      far    *gaze at AMI * gaze averted*
      far              -->•

06    AMI    az   gâz mi-sh-e      deitâ jam     kard?
             from gâz PRG-be-3SG   data  collect do.3SG
             is it OK if a gâz⁴² gets recorded?

07           (0.5)

08    FAR    gâz?
             gâz?

09    AMI    âre.
             yeah
             yeah.

10    FAR    gâz chi-e
             gaaz what-be.3SG
             what is a gâz

11           #± (0.2)#† (0.2) ±
->    ami      ± biting gesture±
      far            †holds hand in defense-->
      fig    #6.12   #6.13
```


#6.12   #6.13

```
12    FAR    ±.h::<<ff>hehehehe±hehehehehehehehehe†hehehe.h> ((3.3s))
      ami    ±,,,,,,,,,,,,,,,,±
      far                                          -->†
```

---

⁴² The noun 'gâz' in Persian refers to different concepts: bite, stove, fume, etc.

This extract shows how the simultaneous affiliative and disaffiliative potentials of the practice make it a useful resource to deal with the sanctionables which make incompatible responses simultaneously relevant. Here, the sanctionable is Farnâz's criticism of Amir, that is, her implicit description of him as a 'paranoid' (lines 01-02, 04). She designs her criticism as playful by an exaggerated assessment and giggling (lines 01-02). When Farnâz's turn is still incomplete, Amir accompanies Farnâz with giggling (line 03). As it becomes clear that the object of Farnâz's play is Amir himself, he stops giggling; however, Farnâz's smile and maintained gaze on Amir act as an invitation for joining in the now-completed joke (lines 04-05). Accepting such an invitation is challenging as the object of the joke is Amir himself. Therefore, the playful design of Farnâz's turn calls for an affiliative smile whereas her criticism can make a range of other actions (e.g., defense, counter-criticism, etc.) relevant. For four noticeable seconds of silence, Farnâz, the joke-teller, maintains her smile and gaze at her co-participant who should now take the turn (line 05, see Hoey, 2015 on lapses). Their interaction is on the verge of termination when Farnâz stops smiling and averts her gaze away from Amir (line 05) until Amir revives it with a 'pre' (Schegloff, 2007) to his mock aggression. He asks if a *gâz* 'bite' can be recorded in the research camera (line 06). This question could potentially make an upcoming bite highly projectable. In effect, he is requesting permission for a bite from the prospective recipient of it[43]. Designing the practice in this maximally cooperative and projectable way reveals its playful nature. The term *gâz,* having multiple meanings (e.g., bite, stove, fume), is ambiguous in this context. Farnâz's first round of repair initiation and Amir's first solution (lines 08-09) do not clarify which meaning is intended. Then, Farnâz initiates another repair (line 10: 'what is a *gâz*'). The repair solution is a biting gesture. Amir's very fast forward-movement towards Farnâz and his concurrent wide mouth-opening (line 11, described as biting gesture, Figure 6.12), which makes Farnâz quickly withdraw her body and hold her hand in a defensive position (lines 6.11 and 6.12, Figure 6.13), shows that the gesture is not simply a repair solution; rather, it targets Farnâz as a recipient, and hence sanctions the criticism. Also, note that Amir does *not* delay the production of the gesture to first receive a grant from Farnâz for a bite. This turns his pre (line 06) into a pro-forma pre.

In this extract too, contradictory design features are juxtaposed to form a playful holistic shape: the practice is designed fast (0.4 second, in this context, disaffililative: giving the recipient less time to prevent receiving the bite), but the mouth movement is designed incomplete for a bite: the performer does not close his mouth (until he initiates the retraction phase) or make any contact with the recipient's body. Thus, he inhibits himself from causing physical harm to the recipient.

The playful design of the sanctionable does not necessarily exempt it from disapproval, but at the same time warrants some acknowledgement. Through a non-aggressive design of a potentially aggressive embodiment, the performer of the practice displays disapproval of the sanctionable but he also acknowledges and reciprocates the playfulness of the transgression. Farnâz starts laughing loudly (line 12, for 3.3 seconds), showing her orientation to the affiliative face of the practice, which leads to more play (not transcribed for brevity). Contrary to mere *sanctioning* or *disattending* the transgression (shown by the long silence in line 05), *reciprocating* the transgressor's playfulness saves the sequence/interaction from pre-maturely ending in disaffiliation.

6.7.2. Renewing Disaffiliative Actions in Less Disaffiliative Designs

Mock aggression can implement an interruption in the forward progress of a disaffiliative interaction at a point where some serious conflict is projected if interactants pursue their uncooperative turns further. A very common orientation to mock aggression is participants' renewing their dispreferred actions but in a more cooperative design that facilities an exit from the disaffiliative sequence. For example, they provide an account for their transgressive conduct or downgrade the design of their disaffiliative turns. Participants' less disaffiliative turns can gradually 'steer the wheel' of the interaction towards a more 'affiliative path', hence calling for the co-participants' more affiliative responses. In this way, they can gradually step away from disaffiliation. Extract 6.5 shows how the recipient of mock aggression downgrades the linguistic design of her disaffiliative question after the practice and how more cooperatively the performer responds to the question, putting an end to their disaffiliative interaction.

---

[43] Amir has a K+ status regarding the requirements of recording as I had explained everything to him (rather than Farnâz). Therefore, his question is heard as requesting permission rather than information.

Katâyun is using her mother's (Farâneh's) needles to do her knitting. Prior to the extract, Katâyun repeatedly alluded to her search for some knitting needles and her inability to find any other than Farâneh's needles. Katâyun's wrist is hurt. The extract starts with Farâneh challenging Katâyun's search for needles.

Extract 6.5. RAY008a, 07.20

```
        01      FAR    >>†asan†  gasht-i•
                       at.all search-2SG.PST
                       did you search at all
                       †....†point to KAT-->
                kat    >>frowns---------•

->  02                 (0.2)± (0.4 )  ±#      (0.1)      ±†•     (0.3)       †
                kat         ±.........±# biting gesture±,,,,,,,,,,,,,,-->
                kat                                     •smiles-->
                far                                -->†withdraws finger†
                fig                    #6.14
```


#6.14

```
        03      KAT    *âre±•
                       yeah
                       Yeah
                       ,,,,±
                        -->•
                far    *smiles-->

        04             (0.8)

        05      FAR    #†gasht-i
                       search-2SG.PST
                       did you search
                       †Point to KAT-->
                fig    #6.15
```


#6.15

```
        06      KAT    âre†
                       yeah
                       Yeah
                far    -->†
```

189

```
07      KAT   #†+bâ   in   dast-e  cholma†        cholâgh-am+ gasht-am*
                    with this hand-of incmplte.word crippled-1SG search-1SG
                I searched with this crippled hand of mine
              +moves wrist up and down four times--------+
       far    †,,,,,,,,,,,,,,,,,,,,,,,,†
       far                                                         -->*
       fig   #6.16
```

#6.16

```
08            (3.0) ((gazes at one another))

09            (4.0) ((both turn gaze at someone passing behind camera))

10      KAT   pâsho          bo-ro
              get.up.IMP.2SG IMP-go.2SG
              get up and go
```

This extract shows how a disaffiliative sequence of actions (challenge-defense) gets gradually less disaffiliative after mock aggression until a new sequence starts. It starts with Farâneh, addressing Katâyun with an interrogative while pointing at her (line 01: 'did you search at all'). Farâneh's turn can be analysed as a strong challenge against Katâyun for three reasons: a) Katâyun had previously alluded to her search for needles several times (not shown in the transcript); therefore, questioning the very occurrence of any searching action has confrontational implications; b) the adverb *asan* 'at all', an instance of what Pomerantz (1986) refers to as an extreme case formulation, calls for a disconfirmation (i.e., 'No')[44], which is in contrast with what Katâyun has repeatedly alluded to, and c) the (potentially) accusatory pointing gesture can heighten conflictual consequences.

Given Katâyun's multiple defense (not transcribed), an upgraded verbal defense would risk the interaction to get 'out of control'. (Note Katâyun's frown in line 01). Of the two verbal and gestural aspects of Farâneh's challenge, the gesture continues even after her verbal utterance is over (lines 01-2). Katâyun first deals with this aspect of Farâneh's challenge: she opens her mouth widely while quickly moving her head towards Farâneh. Then she closes her mouth very close to Farâneh's pointing finger without actually biting it (Figure 6.14). With this, she completes her iconic bite. In the retraction phase of her gesture (i.e., when she is moving back), Katâyun smiles (line 02) as Farâneh withdraws her hand. In terms of design, given the relatively small size of the target body-part (i.e., a fingertip), Katâyun's rather wide mouth-opening besides her fast forward movement gives an overplayed quality to her iconic bite. This and her smile help the gesture to be recognized as playful. The biting gesture can be occasioned by the ready accessibility of the recipient's pointing finger right in front of the performer's mouth. Also note that she is holding a rather sharp needle with one of her hands and the other is hurt, so using an embodiment which involves her hands to playfully sanction the transgression is less likely. Action-wise, Katâyun manages to regulate, at least temporarily, the visible component of Farâneh's challenge. Thus, she negatively sanctions Farâneh's pointing and/or her persistent challenging by physically 'forcing' her a step back.

Katâyun, then, deals with Farâneh's verbal action by answering her interrogative with a minimal token (line 03: 'yeah') as she is still smiling. Simultaneous with Katâyun's minimal token (responsive to her biting gesture), Farâneh too mildly smiles. Thus participants exchange smiles showing their orientation

---

[44] To the best of my knowledge, the preference organization of polar questions in Persian has not been studied, but intuitively the question in line 01 prefers a negative response.

to the gesture as 'non-serious'. Despite Katâyun's clear (e.g., not in overlap) response, Farâneh partially repeats her question while recycling her pointing gesture (line 05: 'did you search?', Figure 6.15), still with a mild smile. However, this time, she downgrades the linguistic design of her interrogative by dropping the adverb 'at all' which makes the question sound less negatively charged. Farâneh's renewed question, which had already received a fitted response, shows her orientation to Katâyun's turn in line 03 as not sufficient. In response, Katâyun does not repeat the biting gesture. After recycling her affirmative response token (line 06: 'yeah'), she adds some clarification about her search as she moves her hurt wrist up and down several times (lines 07: 'I searched with this crippled hand of mine'). Katâyun's clarification shows her orientation to Farâneh's renewed question as a method of inquiring about the quality of her search rather than challenging the whole searching action. Through the dramatized assessment of her hand as *cholâgh* 'crippled', she alludes to the difficulty of the search for her and with her repetitive wrist movement (Figure 6.16), she figuratively shows the physical movement she had to make while searching. Therefore, her verbal and embodied actions convey that her search was thorough enough. By adding some clarification and not just responding minimally as she had done before, Katâyun, too, is being more cooperative. After this, Farâneh stops the disaffiliative sequence of questioning (lines 08-09) and a new sequence of suggesting starts (line 10). They do not go back to the dispute afterwards. Thus, participants' reissuing their actions in less disaffiliative designs shows their orientation to the phenomenon as a point which signals that some co-operative work is needed to gradually de-escalate the argument and to prevent it from reaching a confrontationally tense state.

6.8. Mock Aggression in Pursuit Sequences

In the pursuit collection, both parties are observed to use mock aggression. It can be used by the imperative-speaker after the recipient's noncompliance, in which case the practice negatively sanctions the recipient's noncompliance. Sometimes it is used by the imperative-recipient to sanction the imperative-speaker's persistence. Since this thesis has been focused on the pursuing party's turns, the following is an example of mock aggression employed by the pursuit-speaker. Roshana and Azita have just learnt that the results of their university application are announced. They are trying to find the 'results' page online, using their mobile phones (Figure 6.17). Azita is searching for the 'results' link on a webpage. Roshana is looking for a link that someone called Ekteshafi had earlier sent them. As they are involved in their searching actions, Azita keeps asking Roshana questions. Roshana issues ten imperatives, directing Azita to give her time or be quiet so that she (Roshana) can tell her (Azita) where to find the link. But Azita keeps talking about various things and effectively does not give Roshana time. This is treated as the sanctionable by Roshana. After several times being cut-off, Roshana negatively sanctions Azita's intervening turns with a mock punching gesture (line 41, highlighted gray, Figure 6.19). The imperatives are indicated by a single arrow (**->**). Double arrows (**=>**) show Roshana's turns that are left syntactically incomplete due to Azita's intervening talk.

Extract 6.6. RAY041a, 07.30

```
01   AZI   #yâ ali yâ ali yâ ali
           INTJ   INTJ   INTJ
           oh God oh God oh God
     fig   #6.17
```


#6.17

```
        02              (0.7)

->      03   ROS    bezâ[:r
                    IMP-let.2SG
                    wait

        04   AZI        [khob in    alân [kodum guri berim ((gaze on her phone)))
                         PRT  this  now   which grave SBJV-go-1PL
                         khob where the hell should we click

->      05   ROS                            [bâbâ ah    khafe       sho        dige
                                             PRT  INTJ  suffocated  become.2SG PRT
                                             bâbâ ah shut up dige

        06   ROS    hamash      zer      mi-zan-i.     hamash      ze:r
                    constantly  nonsense PRG-hit-2SG   constantly  nonsense
                    you keep talking nonsense a:ll nonsense

        07   ROS    por  az z[er-i
                    full of nonsense-be.2SG
                    you're full of nonsense

        08   AZI           [bâbâ ku
                            PRT  where
                            bâbâ where

        09   ROS    [por  az zer-i
                     full of nonsense-be.2SG
                     you're full of nonsense

        10   AZI    [khob na-goft-e    kodum
                     PRT  NEG-say-3SG  which
                     khob it {the webpage} hasn't said where

                    guri  be-r-in.
                    grave SBJV-go-2PL
                    the hell you should click.

        11              (0.2)

        12   ROS    alân be-het mi-g-am       tu [sâ\ b\
                    now  to-2SG PRG-say-1SG    in  incomplete.words
                    now I'll tell you          in  sa\ b\

        13   AZI                                 [chie
                                                  what
                                                  what

        14              (0.1)

->      15   ROS    be-bin     age âmad-e   bâshe khob?
                    IMP-look   if  come-3SG AUX    OK
                    look if it is announced OK?

        16              (0.1)

        17   AZI    kho[:
                    OK
                    OK
```

```
->    18   ROS      [alân be-het mi-g-am        #+be-zâr          +
                    now  to-2SG PRG-say-1SG     IMP-wait
                    now I'll tell you           wait
                                                 +hold open hand+
          fig                                   #
```



```
                                                                  6.18

      19        (0.2)

      20   ROS   alân behet mi-g-[am
                 now  to-2SG PRG-say-1SG
                 now I'll tell you

      21   AZI                   [age umad-e   ke [neveshte umad-e
                                  if  come-3SG PRT written  come-3SG
                 what do you mean 'if'. it {webpage} says it's been announced

->    22   ROS                              [sab      kon
                                            patience IMP.do.2SG
                                            wait

      23        (0.2)

=>    24   ROS   uni  ke   ekteshâfi neveshte bud     zohr tuye
                 that that Name      written  PRT.3SG noon in
                 the one {the link} that Ekteshafi had sent at noon in

      25        (1.0)

      26   AZI   to dâr-i     shomâre dâvtalabit-o
                 2SG have-2SG number application-OBJ.MAR
                 do you have your application number?

      27        (0.1)

->    28   ROS   âre  be-b-in      eh   eh
                 yeah IMP-look-2SG INTJ INTJ
                 yeah look eh eh

=>    29   ROS   uni  ke   ekteshâfi ferestâde bud     zohr
                 that that Name      sent      PRT.3SG noon
                 the one {the link} that Ekteshafi had sent at noon

      30   AZI   vali Roshanâ man mi-kh-âm     ratbamo      be-dun-am
                 but  Name    1SG PRG-want-1SG mark-OBJ.mar SBJV-know-1SG
                 Roshanâ but I want to know my mark

      31        (0.7)

      32   ROS   bale bekhâtere inke miduni       khob shod-i
                 yes  because   this PRG-know-2SG good become-2SG
                 yes because you know that {your mark} is good
```

```
     33          (0.2)

     34   AZI    na mi-dun-am    kheili ef*[teza shod-am
                 no PRG-know-1SG very  terrible become-1SG
                 no I know it's very terrible

->   35   ROS                              *[khafe      sho              *
                                             suffocated become.2SG
                                             shut up
          ros                              *gaze from corner of eyeball*


     36          (0.4)

->   37   ROS    be-bin (.) be-bin (.)  ye lahze   gush kon
                 IMP-look   IMP-look  one second ear  IMP.do
                 look look listen one second

=>   38   ROS    uni  ke   ekteshafi zohr ferestâde bud
                 that that Name      noon sent      PRT.3SG
                 the one {the link} that Ekteshafi had sent at noon

     39   ROS    [barâye rotbehâye bartar
                  for     mark-PL   better
                  {the link} for better marks

     40   AZI    [man ke  +[ne-midun-am\
                  1SG that NEG-know-1SG
                  I don't know\

≡>   41   ROS                +[<<h,ff>negâ# vasate ±harfam hamash mi-par-i>=+±
                               PRT middle talk   constantly PRG-jump-2SG
                               Hey you keep cutting me off
          ros              +mock punching--------------------------------+
          azi                                ±smiles------------------±
          fig                     #6.18
```



```
                                                            #6.19


     42   ROS    [((laughs))

     43   AZI    [((laughs))=

     44   ROS    = [<<laugh, p, l>uni ke    zohr ekteshâfi>hehehehe
                               that that noon Name      laugh.PRT
                       the one {the link} that Ekteshafi at noon hehehehe
     45   AZI     =[((laughs))

     46          (2.0)
```

194

```
47   ROS   [<<pp>ekteshâfi>
              Name
              Ekteshafi


48   AZI   [bebin    yani    soojeye ingilisâ mi-sh-im
            IMP-look meaning project English PRG-become-1PL
            look we'll be a project {to be analyzed} in England
```

There are several disaffiliative or dispreferred features that shows this pursuit sequence is on an escalation trajectory: repetition of the imperatives by the pursuing party (Roshana); continued resistance by the recipient (Azita); using culturally 'offensive' lexical components by both, particles and interjections that show negative stance by both; and overlapping turns.

Roshana issues ten imperatives (lines 3, 5, 15, 18, 22, 28, 35, 37). Some of them are repetitions of a prior version, and some are new imperatives. Although the sequential positions of these imperatives and the precise action that is directed by them are different, all the imperatives seem to be designed to achieve a similar interactional outcome: Roshana securing some time and recipiency while searching for the link (see Table 6.2 for a summary). The verbs *bezâr* and *sab kon* both can be translated to 'wait' in English. After issuing the imperative *bezâr* twice (lines 3, 18), Roshana replaces it with the imperative *sab kon* (line 22) (see Chapter 2 on 'replacing' as one verbal method of pursuit). In Persian, the verb *bebin* 'look' and *gush kon* 'listen' can be sometimes used to secure the recipient's recipiency. In this way, *bebin* 'look' and *gush kon* 'listen' have similar interactional outcomes. Again, after three times issuing the verb *bebin* (15, 28, 37)*, Roshana replaces it with verb *gush kon* (line 37). Rather than examining these imperatives separately, suffice it to say that the whole extract depicts a complicated pursuit sequence in which Roshana uses various verbal methods of pursuit (repeating and replacing) to request time and recipiency.

Table 6.2. Summary of the imperatives and their potential interactional outcome in the sequence.

Imperatives	Line numbers	Preferred responsive turn	Potential interactional outcome of the imperative
bezâr 'wait'	Twice (3, 18)	Give Roshana time	Securing time
sab kon 'wait'	Once (22)	Give Roshana time	Securing time
bebin 'look'	four times (15, 28, 37, 37)	Do recipiency	Securing recipiency
gush kon 'listen'	Once (37)	Do recipiency	Securing recipiency
khofe sho 'shut up'	Twice (5, 35)	Be quiet	In the slot in which Azita 'keeps quiet', Roshana can search for the link or produce her own verbal turn, so the potential outcome of this imperative is also securing time and/or recipiency.

Azita displays resistance against Roshana's imperatives. Her verbal resistance is manifested in those turns in which she keeps asking questions rather than giving Roshana time (lines 4, 26) or challenges Roshana's positioning regarding the status of the results in the middle of her search (line 21) or declares what she (Azita) intends to do after they find the link (line 30). All these occur as Roshana is requesting time and recipiency, and in this context, they embody noncompliance. Her resistance is especially clear in those turns that are produced in the middle of Roshana's TCUs. Roshana produces the phrase *uni ke ekteshâfi neveshte bud zohr* 'the one {the link} that Ekteshafi had sent at noon' three times (lines 24, 29, 38), and each time the TCU is left syntactically incomplete by Azita's intervening turns (line 26, 30, 40). It is after the third time that Roshana produces a mock punching gesture, as she explicates the sanctionable: 'hey you keep cutting me off' (line 41). We will analyze the gesture shortly.

After the first 'unmarked' imperative (line 2, *bezâr* 'wait'), the sequence depicts a rather escalated trajectory. There are disaffiliative particles, interjections, and 'offensive' lexical components in both participants' turns. Let us start with Azita's turns. Firstly, she designs her question with the intensifier *guri* 'the hell' (line 4, 'khob where the hell should we click' and line 10: 'khob it {the webpage} hasn't said where the hell you should click'). This intensifier is typically considered 'offensive'. Secondly, she uses the particle *khob* in both of her questions. This particle has implications for 'opposition'. There is also the particle *bâbâ* in her subsequent question (line 8, 'bâbâ where'). The particle *bâbâ* can potentially display annoyance (overall, a negative stance).

The disaffiliative practices are observable in Roshana's turns too. Firstly, responsive to Azita's questions on the whereabouts of the 'results' link, she produces the imperative *shut up* (line 5, 'bâbâ ah *shut up* dige). This imperative is culturally considered an 'offensive' method to direct someone to keep quiet. Secondly, the imperative is accompanied with disaffiliative particles and interjection: *bâbâ, ah,* and *dige*. It is prefaced with the disaffiliative particle *bâbâ* and the interjection *ah*. The interjection *ah* is a response cry responsive to unpleasant stimuli, and it is followed by the particle *dige* which is sometimes used as a sanctioning practice to show that performing the directed action is overdue. Thirdly, Roshana assesses Azita's question on the whereabout of the 'results' link as 'nonsense' four times: 'you keep talking nonsense' (line 6); 'all nonsense' (line 6); 'you're full of nonsense' (line 7) and 'you're full of nonsense (line 9). Note the practice of extreme case formulation in the lexical design of the assessment which further contributes to escalation (Pomerantz, 1986).

The overlapping turns also contribute to the escalated status of the interaction (lines 4-5; 9-10; 20-21-22; 40-41). Through overlapping turns participants are seen to compete for turns. The overlapping participant in effect deprives the other from advancing her CoA, which is one of the features observed in conflictual talk (see Chapter 3, Extract 3.7 for a similar phenomenon).

Some of the imperative turns have upgraded multimodal features. There are overall two reasons why producing multimodal figures of the kind presented in the previous chapters may not be helpful for this complicated sequence. Firstly, imperatives are lexically different, and they are placed in different positions within the imperative turns (which makes them prosodically not readily comparable). Another aspect of complication is that the comparison can be made in different ways: comparing each imperative with the immediately preceding imperative or comparing imperatives that are lexically identical. Since the focus of the sequence is the mock punching gesture, here I just mention some of the upgraded multimodal features in only a couple of the imperatives.

The first imperative (line 3, *bezâr* 'wait') is issued with no 'intensified' prosody, for example, Roshana speaks around her average F_0 level and not particularly loud. Her next imperative *khafe sho* 'shut up' (line 5) is designed higher and louder. As mentioned before, this imperative is accompanied with disaffiliative particles which make the turn lexically upgraded. Now compare the *khafe sho* turn in line 5 with the one in line 35. On the face of it, the second *khafe sho* may look downgraded (line 35), but accompanying with this second rendition, Roshana casts a sanctioning look at Azita. So even though particles are dispensed with, a disaffiliative embodied feature is added. Therefore, the upgrade can be perceived in Roshana's embodiment (Since the footage must be anonymized, the sanctioning look cannot be easily seen in a filtered still image, so a figure on this is not provided).

In respect to the two renditions of the imperative *bezâr* 'wait' in line 3 and 18, the second rendition is designed with an upgraded bodily action. Roshana adds what looks like to be a 'stop' gesture in the production of the second *bezâr* (line 18). She opens her hand and keep it in an oblique orientation with her palm facing out towards Azita. This canonical stop gesture achieves an embodied upgrading. The prosody of the two versions is not straightforwardly comparable since contrary to the first version, the second one is produced turn final.

Having seen Roshana's persistence, Azita's resistance, and the disaffiliative practices that escalate their incompatible CoAs, we can now analyze the mock punching gesture. As mentioned above, the gesture is produced after Azita cuts Roshana's turns three times (lines 24-26; 29-30; 39-40). The gesture is accompanied by a verbal turn in which Roshana explicitly mentions Azita's transgression: 'hey you keep cutting me off' (line 41) (see Rossi, 2018 on factual declaratives doing criticism). This turn is produced high in the speaker's pitch range, and louder and faster compared to the speaker's preceding turns. As for the gesture itself, Roshana fists her left hand, stretches her arm towards Azita, and moves her closed fist up and down three times as in a punching movement. The combination of the verbal turn, the prosodic design of the turn, and the co-occurring punching gesture produces a Gestalt by which Roshana holds Azita accountable for not allowing her to finish her turns (or in other words for not complying with her earlier imperatives).

The gesture is not designed as physically threatening nor is it oriented to as such by the recipient. Roshana remains sitting in her chair throughout the production of the gesture. Even though the up and down hand movement is fast (which matches the fast rate of the articulation of the verbal turn), there is distance between the two participants and no physical contact is projected in Roshana's bodily conduct. Just in the middle of Roshana's gesture, Azita smiles widely which can potentially show her treatment of the punching gesture as only 'mock'.

The gesture has affiliative potentials. Right after the gesture, both participants start laughing. Hence both can be seen to be orienting to the 'playful' aspect of the conduct. Roshana recycles the TCU that was repeatedly cut off by Azita (line 44). However, the recycled turn is laughter infused; it is designed quiet and low; and it is abandoned by Roshana herself as she continues laughing loudly (line 44). Then, she recycles only a part of the previously cut off turn in very quiet design (line 47). The quiet design of it makes it to be heard more as 'self-talk'. So after the mock punching, Roshana can be seen to be gradually desisting from the sequence which she was trying to pursue with Azita. Both participants continue laughing, and in this way an escalated collision in the two participants' CoAs is de-escalated. Mock aggression between these two intimate interactants, therefore, provides an exit from their sequence in which both were trapped in a loop of disaffiliative conduct. It provides a transit to the next sequence in which Azita assesses what just happened in their interaction as 'worthy of analysis' by the researcher in England.[45]

6.9. Discussion

Mock aggression refers to the embodiments whose holistic forms are associated with aggressive actions (e.g., strangling), but the design of which make them different from aggression and as such they do not produce the same interactional outcome as aggression would; that is, they do not constitute any serious physical threats. This study showed that mock aggression can negatively sanction a transgression, but, due to its design features, it often does so by offering an exit out of disaffiliation into a more affiliative sequence. When it occurs in an escalated pursuit, it functions as a pivot point where a persisting participant starts to gradually desist from the sequence or at least pursues compliance with less disaffiliative constructional features.

The literature in CA distinguishes between alignment, "the structural level of cooperation" and affiliation, "the affective level of cooperation" with the former targeting "the formal design preference of the turn" and the latter, "the level of action and affective stance" (Stivers, Mondada, & Steensig, 2011, pp. 20-21; Stivers, 2008). The findings here propose that when it comes to affiliation, an additional distinction between *social actions* and their *interactional outcomes* should be considered. For example, disaffiliating with the co-participant's just performed action may not necessarily have

---

[45] The recordings were made only in a couple of months after I started my PhD project, and even *I* did not know that mock aggression would become the focus of my research. Recordings were made with no a priori focus on any social phenomenon. Azita's anticipation is accidental.

disaffiliating outcomes as the case of mock aggression showed. (Dis)Affiliative outcomes depend on a combination of factors including the *method* of doing (dis)affiliation.

In CA, methods of promoting/demoting social solidarity are usually discussed in the light of the possibility of different *alternatives* of actions in the same sequential position (Heritage, 1984). Even when Jefferson et al. (1987) suggest a continuum of affiliation-disaffiliation, they still analyze participants' responses to impropriety talk as falling on the *either* affiliative *or* disaffiliative side of the continuum. CA has thus taken this categorical alternative-based perspective on the (dis)affiliative status of social actions. This study raises questions about this rather strict dichotomy from five perspectives: the format of some practices like mock aggression, their sequential organization, participants' orientations to them, the social actions that they perform, and their outcomes in the interaction.

Regarding the format, the physical design of mock aggression is a mixed composition of features which could typically be seen as promoting social solidarity *and* demoting it. For example, in Extract 6.5 (lines 1-3), we saw that the performer of mock aggression, who had been frowning, stops frowning as she initiates her biting gesture and smiles in partial overlap with the gesture. Stopping a frown and initiating a smile, which is typically associated with affiliative actions co-occur with a biting gesture which could *potentially* demote social solidarity (because it could potentially harm the recipient of the bite). Even without other co-occurring gestures, mock aggression in itself has contradictory design features. For example, a biting gesture can be seen as disaffiliative because of the harm it represents and not as such because of its, for example, overplayed theatrical qualities. To borrow Mondada's (2014, p. 137) terminology, the performer uses a web of resources packaged in a "complex multimodal Gestalt", but in the case of mock aggression these resources are contradictory in terms of their typical affiliative connotations. Rather than a paradoxical *message*, which some play-researchers (e.g., Bateson, 1972) have theorized to be the necessary ingredient of play, the current empirical data suggest that it is the juxtaposition of these typically affiliative and disaffiliative *design features* (and co-participants' orientations to them, see below) which convey mock aggression as playful.

The contradictory design features are also observable at the sequence organizational level. In Extract 6.4 (line 6), for example, the performer of mock aggression produces a pre, verbally requesting a permission for a *gâz* 'bite'. This sequential design makes his *gâz* highly projectable and thus less disaffiliative (because of the chance that it gives the co-participant to block the projected bite and by implication, the potential physical harm). The participant is requesting permission – a cooperative action – for a bite – a potentially uncooperative action. However, as we saw earlier, without the co-participant's granting any permission, the mock bite is performed. This turns the pre to one which is performed just pro-forma. This study, therefore, opens a new window to our understanding of social actions' design features (both at turn and sequence levels) by proposing that a dichotomy of preferred-dispreferred design cannot explain the mixed composition with which some practices like mock aggression are systematically shaped.

The way co-participants orient to mock aggression also reveals its double-faceted status. For example, in Extract 6.4, the recipient of the iconic bite, holds her hand in a defensive position, a practice used to avoid the gesture, but at the same time laughs loudly, a practice which can be analyzed as displaying a positive stance toward the gesture. This shows the contradictory orientations that the recipients themselves take towards mock aggression. Overall, at a single moment of interaction or in rapid succession, one can see participants both affiliating and disaffiliating with each other (i.e., figuratively and sometimes quite literally, both approaching and distancing themselves from one another). Therefore, from the emic perspective of participants themselves, the practice can simultaneously occasion both an affiliative *and* dissaffiliative treatment.

Action-wise, in this corpus, mock aggression was analyzed to *sanction* transgressions (a method for disaffiliating with transgressions), but because of its mixed design features and orientations, it can be conveyed that the performer of mock aggression is also *playing* with the recipient concurrently. That is, participants perform one action the overall design of which resembles another action; however, its detailed design and co-participants' orientations make it different from what it resembles. These two actions are fused to make a whole, the interactional implications of which is different from those of each action separately (See also Goffman, 1974/1986, p. 82, for the analogy adding a *layer* or

lamination to an activity"). Whereas sanctioning can have disaffiliative consequences in the interaction, playful sanctioning creates opportunities for affiliation. From this perspective, Walker's (2014) discussion on the independence of form and function is supported in the realm of embodiment; that is, the same bodily movements which can otherwise provoke/escalate a conflict function to dissolve/manage the conflict when more granular details of form are taken into account.

The outcome of the practice can be further explained in light of the preference for progressivity in interaction (cf. Stivers & Robinson, 2006). Clayman (2002) explains that one way in which a preferred response (e.g., an acceptance to a request as opposed to a rejection) is cooperative is that it "advances the course of action that has been initiated and thus collaborates with the agenda being pursued by the prior speaker" (p. 231). In this respect, mock aggression has double potentials: it is designed to hinder a disaffiliative course of action, but it often does so by opening an exit into a more affiliative course of action. For example, it can work as a pivot point after which the same disaffiliative sequence continues (e.g., Extract 6.5) or resumes (e.g., Extract 6.4) but with a more cooperative design. All this makes it almost impossible to decide on the affiliative-disaffiliative status of mock aggression with an either-or approach.

The findings also show that although the particular kind of embodiment used by the performer of the practice is not precisely projectable, there are a number of factors which affect the choice of embodiment: (non)proximity of the participants (Extracts 6.2-6.3), accessibility of the recipient's target body-part (Extract 6.5), availability of the performer's body parts (Extract 6.5), and the kind of transgression (Extract 6.1). Therefore, as Mondada (2014) notes, participants locally choose the exact format of the embodied practice based on the local contingencies and contextual constraints.

Mock aggression is a pathway around an interactional tension (compare with Collin's, 2009, concept of violence as a pathway around an emotional confrontational tension). Participants' shared knowledge about the norms on when, how, and with whom to use it furnishes a safe setting for its use. This study showed some of these norms; for example, the practice is used between intimate participants and constitutes no serious physical threats. When the normative conventions of its use are observed, the practice can help participants to disaffiliate with transgressions while maintaining affiliation with the transgressors. These double affordances make mock aggression a valuable resource to consolidate both social bonds and social norms in conflictual encounters. Future research can explore cases where what seems to be designed as mock aggression is not oriented to as playful and investigate how that influences the trajectory of interaction; what other resources are used to dissolve conflicts; and how they are different from/similar to mock aggression.

Chapter 7. Conclusions

This thesis focused its analytic lens on pursuit in social interaction: a ubiquitous social phenomenon that has been known to conversation analysts for almost 40 years (Jefferson, 1981; Pomerantz, 1984). Despite this rather long history, there is still much to explore about this phenomenon. By taking a multimodal approach to data, this study makes some novel and significant contributions to our understanding of pursuits in terms of sequence structure, turn design features, agentive stance, and the management of escalated episodes of interaction.

Chapter 1 delimited the broad scope of the thesis to conflictual and disaffiliative interaction. We started with Goffman's (2010, pp. 5-6) metaphor of social members as vehicular units whose pursuits may collide in their daily encounters. We narrowed our attention to imperative sequences, and among all different types of imperatives, we considered only those that are designed to get a recipient to perform an action. We further sharpened our focus on those imperative sequences in which the imperative-recipient displays resistance against complying. And then the enquiry began to find answers to the following questions:

1. What verbal practices do participants use to pursue compliance in the face of resistance?
2. Do the details in the composition of the pursuing action have consequences for the trajectory of the interaction?
3. What implications do the multimodal design features of a pursuit turn have for the structure of the sequence?
4. What implications does the action of pursuing have for the participants' agency?
5. Do different practices of pursuit vary in terms of the agentive stance that the pursuing party adopts?
6. How may participants exit an escalated pursuit?

The section below summarizes the findings in respect to the above questions.

7.1. Summary of the Findings

Chapter 2 inspected the structure of the pursuit sequences under scrutiny. Firstly, the organization of an imperative sequence was dissected. Several types of responses to an imperative turn were exemplified: preferred, dispreferred, and less than preferred responses, and then resistance was operationalized as dispreferred and less than preferred responses. Then a number of verbal practices of pursuit were presented: incrementing, addressing, accounting, replacing, reformatting, and repeating. Of all these verbal methods, we aimed our attention on pursuit by repeating a prior turn after a co-interactant's resistance. The chapter makes a significant contribution to the literature by identifying two types of pursuit based on the stances that a pursuing party displays. These stances are manifested in the design features of a pursuit turn and/or the co-occurring practices or actions. In the first type, the pursuing party *persists* in furthering her CoA while in the second type she *gradually desists* from the sequence. While in both types, a pursuit turn expands the sequence, persisting projects more sequence expansions whereas in gradual desisting a contingent sequence closure is more likely. So the two pursuit types differ in terms of their projective potentials.

In Chapters 3 and 4, we took a more detailed look at the multimodal design of pursuit turns and examined a novel phenomenon: *multimodal gradation*: a 'stronger' or 'weaker' manifestation of a multimodal Gestalt relative to a previous point of reference (Mondada, 2014 on multimodal Gestalts). Multimodal gradation, as conceptualized in this thesis, involves a coordinated integration of a turn's

design features from different modalities which yields a unified whole. In *multimodal upgrading,* the focus of Chapter 3, this Gestalt is a more 'intensified' or 'stronger' version compared to a prior point of reference, and in *multimodal downgrading*, introduced in Chapter 4, the Gestalt is a 'subdued' or 'weaker' version relative to a prior counterpart.

In these two chapters, we analyzed the implications of multimodal up- or downgrading of a pursuit turn. The claim was that multimodal upgrading is more associated with sequence expansion while multimodal downgrading is more connected with disengagement and sequence closure. Two sets of observation were used as the basis for this claim: observations on the co-occurring practices that are used with or within multimodal Gestalts and observation on the sequential position of multimodal Gestalts.

The practices or actions that are observed to co-occur with/within multimodal Gestalts in a pursuit turn consist of sanctioning (disaffiliative), pleading (not necessarily disaffiliative), positive assessment of the recipient (affiliative), and the like. With the thesis revolving around disaffiliative conduct, we only focused on the first group of practices; that is, disaffiliative practices by which a pursuing party holds the recipient at fault for resistance. The occurrence of these practices in/with a pursuit turn puts the pursuing party's persisting stance on display. Stated differently, the occurrence of these practices in/with a pursuit turn shows that the speaker is likely to produce more turns to make her CoA progress if she receives a noncompliant response.

Now, in the majority of sequences under scrutiny, *multimodal upgrading* is accompanied with practices listed above. That is between upgraded and downgraded Gestalts, the overall distribution of these practices is higher in *upgraded* pursuits rather than downgraded ones. On the basis of this observation, multimodal upgrading is argued to be more associated with sequence expansion and downgraded pursuits with sequence closure and withdrawal. The logic behind this reasoning is that when the pursuing party does not treat resistance as sanctionable, they are more likely to disengage or back down and vice versa.

The second observation is related to different sequential positions where these two groups of Gestalts occur. Both up- and downgraded Gestalts occur after a recipient's resistance, but after compliance, in this collection, only downgraded Gestalts are observed to occur. (Sequence closing imperatives with virtually 'same' multimodal design features as a prior were not observed in this collection). After compliance, pursuing is not relevant anymore. So when an imperative is repeated after the recipient's compliance, it must be in the service of other social actions than pursuing. In this sequential position, a sequence closure or potential disengagement is more likely and thanks to the co-occurring features of these repetitions (such as confirmatory nods, etc.), these types of repetitions were argued to be sequence closing imperatives.

Now, in this collection, sequence closing imperatives are never designed with upgraded features; rather they have downgraded features. Again, the occurrence of only downgraded features where a sequence closure is more likely makes downgrading rather than upgrading more associable with sequence endings and withdrawal.

To recap, upgraded pursuits are projective of more sequence expansions and when they occur with sanctioning practices, they project escalation. The opposite is true for downgraded pursuits. They project a contingent sequence closure and when they occur in an escalated situation, they contribute to de-escalating the clash (see Figure 7.1. for a summary).

1. A: Imperative designed to get the recipient to do an action
2. B: Resistance
3. A: ─────────

Abandoning or closing the sequence   Pursuing compliance

Downgraded pursuits (desisting/relinquishing stance)

Upgraded pursuits co-occurring with affiliative practices (persisting stance)

Upgraded pursuits co-occurring with disaffiliative practices (persisting stance)

Contingent sequence closure

Contingent sequence expansion

Contingent sequence expansion

If the pursuit is already escalated, de-escalation is likely.

Future research

Escalation

Figure 7.1. The projective potential of multimodal upgrading and downgrading in pursuit sequences.
The gray lines show a trajectory, observed in the data, but not focused on in this manuscript.

Chapter 5 analyzed multimodal gradation in relation to the agentive stance of a pursuing party. It was shown that by pursuing compliance, an interactant displays a stronger agentive stance relative to abandoning the same sequence. Also, a pursuit with an upgraded multimodal design manifests stronger agency for the pursuing party relative to downgraded pursuit turns. We further distinguished two types of upgraded pursuit turns. In the first type, the pursuing party bodily movements (progressively) limit the recipient's personal space (Hall, 1966); nevertheless, they pursuit turn still creates another opportunity for the *recipient* to comply. In the second type, even though a verbal pursuit is produced, bodily actions of the pursuing participant aim to deprive the recipient to perform an independent response. For example, the speaker forces the recipient to perform the directed action while repeating the imperative. The claim was that in forcing the pursued outcome to happen, the pursing interactant displays stronger agency that in merely approaching the recipient.

Finally, Chapter 6 introduced another novel phenomenon: mock aggression (Afshari Saleh, 2020). Mock aggression is one way that participants are observed to manage the collision of their pursuits. It is described as the embodiments which, in one way or another, appear aggressive (punching, pinching, slapping, etc.) but are not designed to be, or oriented to as, serious physical threats. When it occurs between intimate interactants, it negatively sanctions transgressions and at the same time provides systematic opportunities for participants to engage in more affiliative interaction. The findings show that despite its aggressive appearance, mock aggression facilitates participants' exit from a disaffiliative interaction, owing to its detailed design features and thereby contributes to maintaining their social bonds.

Below the implications of the findings and the contributions they make to the relevant literature are discussed.

7.2. Contributions and Implications of the Findings

The findings in this project contribute to the literature on pursuits (Bolden, et al., 2012; Jefferson, 1981; Pomerantz, 1984; Stiveres & Rassano, 2010), multimodality (Mondada, 2016, 2018, 2019b); turn design (Drew, 2012), sequence structure in adjacency pairs (Schegloff, 2007), preference organization (Clayman, 200; Heritage, 1984; Kendrick & Torreira, 2015; Pillet-Shore, 2017; Pomerantz & Heritage), affiliation and disaffiliation (Afshari Saleh, 2020; Lindström, & Sorjonen, 2013), conflictual interaction (Clift, 2020; Clift & Pino, 2020; Coutler, 1979/1990; Dersely & Wootton, 2001; 2010; Glenn, 2019; Goodwin, 1980, 1982, 1983; Hoey et al, 2020; Holt, 2012; Lloyd, 2015; 2017; Lloyd & Mlynář, 2021; Maynard, 1985; Whitehead, 2015; Whitehead, et al, 2018), and action formation (Levinson, 2013). Detailed implications of the findings in respect with each of these themes are presented at the discussion section of each chapter. Here, I address the bigger picture and discuss the implications in five sections: Turn design and sequence structure (§7.2.1), stance, preference, and affiliation (a contradiction conundrum, §7.2.2), conflictual encounters (§7.2.3), prosody beyond CA (§7.2.4), and multimodality (§7.2.5).

7.2.1. Turn Design and Sequence Structure

In CA, the gradation of a turn's design features is already discussed for separate modalities: lexicon (Heritage & Raymond, 2005; Hoey et al., 2020; Lerner, et al, 2012; Pomerantz, 1984), prosody (Couper-Kuhlen, 2012; Curl, 2006; Local, Auer, & Drew, 2010; Ogden, 2006; 2010; Sikveland, 2019; and Zellers & Ogden, 2013), and body (Hauser, 2019; Keel, 2015). In these studies, the gradation of a turn's constructional features is reported to afford various social functions depending on the interactional context. This thesis shows that several design features involving different modalities can simultaneously be up- or downgraded. Using Mondada's (2014) terms, a whole multimodal Gestalt can undergo gradation from one turn to the next. The findings here are in line with those studies in which multimodal gradation is discussed in light of sequence structure.

Focused on the prosodic design of complaint sequences, for example, Ogden (2010) shows that upgraded complaints seek affiliative responses while downgraded ones project sequence closure. Sequence expansion or closure is also discussed in terms of clusters of prosodic features, and not particularly graded features. For example, Local and Walker (2005) distinguish two types of 'so' in English data. *Holding* 'so' projects more on-topic talk by the 'so'-speaker and is designed as "relatively high in pitch, relatively loud, with final glottal closure", while *trail-off* 'so' projects sequence ending and is designed as quieter, lower in pitch, and never with a final glottal closure (pp. 6-7). Reviewing the findings in the literature (e.g., Curl, Walker & Local, 2006; Ogden, 2012). Ogden (2017) distinguishes two clusters of prosodic features: low-quiet (LQ) and high-loud (HL) clusters which have distinct interactional affordances. The former has affordances for turn or sequence expansion and the latter turn or sequence closure.

Sequence structure is also discussed in relation to participants' bodily conduct (Goodwin, 1981; Raymond & Lerner, 2014; Rossano, 2006; 2012; Schegloff, 1998, among others). Gaze aversion, for example, is shown to signal sequence ending and disengagement (Goodwin, 1981; Rossano, 2012). The same also applies for retracting body to home position (Schegloff, 1998). Lexical resources, such as expletive insertion, have also been shown to have consequences for sequence structure (Hoey et al, 2020).

This research contributes to these studies from two perspectives. Firstly, it shows that many of these features that are separately reported to project expansion or closure of a sequence may in fact belong to a larger multimodal Gestalt. Thus, it invites further multimodal investigations to consider the totality of the resources that may be operative in structuring a sequence.

Then, it also contributes to our understanding of post-expansions in CA. As Figure 7.1 shows, participants' understanding of where a sequence is heading towards depends on a complicated interplay between the details of a turn's design features and other co-occurring practices. Depending on participants' displayed stances, post-expansions have terminative or expansive potentials. Schegloff (2007) has already distinguished minimal from non-minimal post-expansions with the former closing the sequence and the latter expanding it. This thesis further shows that non-minimal post-expansions, themselves, can be of two types: those that project yet more non-minimal post-expansions if they do not receive a preferred response, and those that are likely to be (among) the last even if they do not receive a preferred response. In other words, post-expansions can be of persisting or desisting type.

Relatably, one of the striking observations is that downgraded multimodal pursuits project sequence closure *as* the sequence is still expanding. In other words, the very turn that projects sequence closure expands the sequence as well. In this way, multimodal downgrading in pursuit turns can embody a process of attrition in which a sequence comes to closure in a stepwise fashion. This observation can be discussed as a contradiction conundrum in CA.

7.2.2. Stance, Preference, and Affiliation: A Contradiction Conundrum

In sequences in which gradual desisting occurs, the sequence expands since more pursuit turns are being produced but they have design features that are more associable with sequence endings. Therefore, the speaker can be seen as both pursuing and at the same time letting go. Pursuing is achieved by the verbal repetition of the imperative. Letting go is signaled by the design features of the imperative turn such as downgraded prosody and the like. This observation can point our attention to a potential conundrum underlying some CA analytic concepts, and that is seemingly contradictory analytic concepts being simultaneously applicable to one turn, one social action, or one social phenomenon.

This thesis showed that such contradiction may occur for some other CA concepts such as affiliation and preference. For example, mock aggression was shown to include contradictory design features: those that are conventionally associable with affiliation (e.g., smiling) and those that may be indicative of disaffiliation (e.g., an ostensibly aggressive movement). The phenomenon was also shown to afford seemingly contradictory functions: disaffiliating with social transgressions and at the same opening opportunities for affiliation. The same goes for preference. In this thesis, we considered preferred responsive actions as those that allow the progression of an initiated CoA and vice versa (Clayman, 2002; Pillet-Shore, 2017; Schegloff, 2007). But we saw that the design features of a single responsive action can be seen as allowing for the progression of the initiating action but also at the same time not quite so (e.g., A recipient passes the bottle of water in response to the imperative 'pour water' and hence making water accessible to the requester. This makes it possible for the requester herself to further the CoA of drinking water. Even though the recipient has contributed to the progression of the CoA, he has not performed the directed action per se). CA has traditionally considered anything less than a preferred response as dispreferred, but is this a right approach? As we saw, participants themselves may show orientations to both preferred and dispreferred aspects of such actions (see Chapter 2, §2.5 or Chapter 4, §4.8.2).

Even choosing titles for such phenomena may not be straightforward since the title should encompass contradictory features: it is not entirely desisting; it is rather a *gradual* desisting. It is not aggression per se; it is more of *mock* aggression, or it is not totally dispreferred but rather *less than preferred.* To handle such a conundrum, this research proposes that some social phenomena should be understood as a process or continuum rather than discrete units or categories. Rather than only two categories for the preference structure, a continuum of possibilities can be observed where preferred and dispreferred design features are located on the two extremes, but there are also other possibilities for a mixture of both preferred and dispreferred design features in formatting one social action. The same goes for affiliation, stance, and the projected sequence structure. For example, rather than treating sequence closure as an abrupt point in interaction, it can be seen as a process shaped through multiple turns.

7.2.3. Conflictual Encounters

The projective potential of multimodal gradation in pursuit turns has implications for conflict management. Multimodal upgrading and downgrading can be two hallmarks of escalation and de-

escalation processes, respectively. The following points out the contributions of the research for conflict studies.

Escalation and De-escalation

Recent research has discussed the relationship between upgrades as a turn constructional feature with escalation as an action, achieved by upgrades (Hoey et al., 2020). This research associates upgrading with sequence expansions rather than directly with escalation as action. Here, escalation is identified as one type of sequence trajectory: sequence expansions actualized through more and more overt disaffiliative actions. A direct relationship between upgrading and escalation could probably work for Hoey et al.'s study since they are focused on a particular type of lexical resource, an expletive insertion. By virtue of the semantics of the lexical resource, the upgrade has disaffiliative affordances and can potentially escalate the discord. But in the sequences in this collection a variety of resources are used in an upgraded Gestalt. Even though we focused on upgrades that accompany disaffiliative practices, as mentioned earlier, upgrading can also project sequence expansions that do not necessarily contain disaffiliative actions. Associating upgrading with sequence expansion rather than directly with escalation leaves the door open for further investigations on sequences in which upgrading a pursuit turn projects more sequence expansions, but those expansions do not particularly escalate the situation. This is reflected in Figure 7.1, presented in §7.1 (see §7.3 on direction for further research). The figure illustrates that while upgrading the design features of a pursuit turn can potentially display the speaker's persisting stance (i.e., contingent sequence expansions), it does not automatically escalate the clash.

Such a view to escalation implies that rather than being the result of a particular design feature or even a particular type of a complex Gestalt, escalation is more likely to be associated with the affiliative valence of the design feature or the Gestalt. It is *disaffiliative* upgraded Gestalts that escalate the situation rather than an upgraded Gestalt.

This research shed light on two different practices that may de-escalate a conflictual interaction: multimodal downgrading and mock aggression. Multimodal downgrading embodies a gradual process through which the conflictual sequence undergoes attrition. Mock aggression, on the other hand, is likely to be used when disaffiliation is escalated. When it occurs between intimate interactants, it offers an exit from the escalated situation potentially at its peak.

Professional vision

In his paper on 'Professional Vision', Goodwin (1994) shows the importance of precise coding and graphic representations in analyzing an episode of interaction. On the case of Rodney King, for example, he shows that coding the interaction between the police force and Rodney King transforms a massive beating scenario into separate beating episodes "each with its own sequence stages" oscillating between escalation and de-escalation periods (p. 617). In this thesis, too, participants' re-doing their previous actions, when coded, segmented, and aligned with their other vocal and bodily behavior, demonstrates various types of pursuit (with different implications for escalation and de-escalation) which otherwise could have been seen as just an episode of pursuit.

Goodwin continues to emphasize the power of demonstrating rather than stating: "As talk and image mutually enhance each other, a demonstration that is greater than the sum of its parts emerges" (p. 620-621). Also, focused on the micro-sociology of violence, Collins (2009) states:

> It is not literally true that a picture is worth a thousand words. Most people will not see what is in a picture, or will see it through the most readily available visual cliche´s. It takes training and an analytical vocabulary to talk about what is in a picture, and to know what to look for. A picture is worth a thousand words only for those who already have internalized an adequate vocabulary. This is particularly so when we have to train ourselves to see micro-details (p. 5).

The graphic presentations throughout this thesis can potentially advance the literature's coding and graphic visualizations that are designed to show the relationship between various features from a multitude of modalities. Even though graphic presentations have always been used in CA ever since pioneers such as Goodwin (1981) started analyzing video data, this thesis makes use of multifaceted figures in which multiple aspects of production from various modalities are shown and aligned. These presentations take the clarity of multimodal arguments to the next level and thus contribute to future research on multimodality in general and in particular to enquiries into conflicts, aggression, and violence. Following Collins' (2009) point, the argument in this thesis offers technical vocabulary that one may use in analyzing an episode of conflict.

7.2.4. Prosody beyond CA

The literature on prosody beyond CA has considered phonetic manipulations to an utterance for communicative purposes. To explain the relationship between prosodic forms and their communicative interpretations, Gussenhoven (2004) proposes some 'codes' based on which speakers of a language may interpret affective attributes of a speaker (pp. 71-96). Two of these codes are 'effort code' and 'production code'. Gussenhoven offers detailed explanations on the phonetic exponents of these codes, for example a higher level of effort in the production of an utterance includes a 'wider excursion of the pitch movement', which is one of the phonetic exponents observed in upgraded Gestalts in this thesis (p. 85). And the production code is related to the subglottal air pressure (p. 89). A higher level of subglottal air level results in louder and higher utterances, which in this thesis are two phonetic features of upgraded Gestalts. Gussenhoven proposes some affective interpretations of these codes. For example, a higher effort code signals 'authoritative', 'insistent', and 'enthusiastic' while lower effort codes are indicative of 'lacking commitment' and 'uninterested'. For the production code, he states, 'high endings signal continuation, low endings finality and end of turn' (p. 89).

There are some similarities between the interactional affordance of the up- and downgraded Gestalts reported in this thesis and the 'paralinguistic interpretations' of prosody proposed by Gussenhoven (2004, pp. 71-96). For example, he mentions 'insistence' and 'lack of commitment' for utterances with higher and lower effort codes, respectively. This may roughly go with my claim on upgraded and downgraded Gestalts being associated with 'persisting' and 'gradual desisting', respectively. However, this thesis provides a more sophisticated picture. Firstly, the findings here imply that the distinction between the two categories of code – 'effort' versus 'production' – may be an artificial one. In naturalistic interaction, phonetic exponents attributed to these categories of code go hand in hand, create an integrated Gestalt, and together perform interactional functions. Secondly, in face-to-face interaction, upgraded prosodic exponents are claimed to display persistence when there is no 'mismatch' between prosodic and other relevant design features. Future research should explore the interactional potentials of prosodic features when they 'mismatch' those of other modalities. For example, can a higher and louder utterance still be indicative of persistence when the co-occurring bodily actions are downgraded? Without considering a speaker's co-occurring conduct, a claim such as a 'higher effort code signals insistence' can be questioned for naturalistic face-to-face data. And finally, rather than conveying interpretations associated with an utterance, the phonetic features investigated in this research are a part of a Gestalt that performs interactional functions (e.g., expanding the sequence or moving towards closure). Without considering the details of how the sequence unfolds (other than prosodic details), a firm claim on the interactional affordance of the Gestalt seems to be simplistic. For example, an upgraded pursuit turn can display persistence in part because the pursuit turn *re-does* an action that has already been resisted, but the function of the same Gestalt may differ in a different sequential context.

7.2.5. Multimodality within and beyond CA

As discussed in Chapters 3-4, categorizing a bodily action as up- or downgraded is not so straightforward; nevertheless, this study is perhaps among the first to document some recurrent bodily movements that are observed to co-occur with prosodic and lexical up- and downgrades. Traditionally,

resources by which social actions are performed are associated with separate modalities: verbal, prosodic, and embodied. And from a more granular perspective, these modalities each consist of various smaller-scaled cues such as syntactic, morphological, lexical and the like for the verbal level; the same goes for prosodic and bodily resources. The observation that multiple resources across a range of modalities are modified in a coordinated fashion re-establishes the fact that there is a tight connection between them. (e.g., Kendon, 1980; McNeill, 1985; 1992).

We may take a further step on the relationship between these resources and consider the possibility that although traditionally attributed to different modalities, these resources are in the first instance produced as one. It is not just that speech and the co-speech bodily actions are pragmatically co-expressive (e.g., McNeill, 1992) or that the modalities are temporally coordinated (e.g., Kendon, 2004, pp. 108-126). Rather, in re-doing a social action, a multitude of employed resources are congruently modified all at once. This implies that the selection of the resources is so systematic and the connection between them is so tight that when a participant 'needs' to re-do the action, not only is the whole Gestalt re-employed but it is also modified in a coordinated fashion.

Much evidence for the integration of various modalities comes from studies on perception or psycholinguistic processing (e.g., Hagoort & Berkum, 2007; Holler & Levinson, 2019. See Goldin-Meadow & Brentari, 2017 and Perniss, 2018 for a review). When it comes to production, researchers have been mainly focused on sign languages and/or on the connection between modalities within utterances. This research traced the speakers' verbal, prosodic, and bodily behavior throughout a whole sequence of action. The findings open a new window for our inquiry into the psychological integration of multimodal resources that convey meaning or perform social actions. It poses the question on how our brain may handle such coordination.

Concerned with the *processing* aspect of multimodal signals and Gestalts, Holler and Levinson (2019) have recently asked how our brains segregate the relevant from irrelevant movements while processing the "complex orchestration of multiple articulators" (p. 639), and how our brains process a multitude of movements that may not be perfectly aligned as one coherent message. Multimodal gradation, of the type that has been identified in this thesis, puts forward similar sets of questions on the *production* side of the story. In re-formatting a social action, the modification effect – up or downgrading – is applied to the multitude of cues that are involved. That is, the effect is applied to a number of resources. The question is how, for example, the 'intensification command' is sent to various modalities. Furthermore, in the process of up- or downgrading, there are resources that do not change even though they are involved in the production of the social action. For instance, not all prosodic measures may change. Or, as another example, both hands and brows may be involved in the production of a social action, but only the movements of brows are upgraded, and not the hands. In these instances, the question is how the 'intensification command' applies only to some resources involved in the production of the social action? Does the multimodal Gestalt consist of those resources that are up- or downgraded unanimously or does it consist of all the resources that contribute to the production of the social action even if they are not up- or downgraded with the rest of resources? Finding answers to these questions may take us closer to answers on the processing aspect of multimodal actions.

7.3. Further Directions

This project opened various windows into future investigations on multimodality and conflictual talk. Firstly, a next step following this project is a quantification of the findings. The claim that multimodal upgrading is associated with sequence expansions is made based on the frequent co-occurrence of, for example, sanctioning practices with/within such upgrades and their rare happening with/within downgraded Gestalts. The overall claim is in line with what the literature has shown, so from this perspective, the findings are supported. However, for a firmer ground for such a claim, quantifying the data can help. Secondly, as mentioned in §7.1, affiliative practices can also occur with upgraded

Gestalts. Future research can unpack the details of such affiliative practices and how they may shape the trajectory of the sequence.

Further research can also examine different sequence types in which persisting may occur. People may persist in performing other social actions such as complaining, insulting, disagreeing, and so forth. Constructional features of persistent turns that perform other social actions than directing can contribute to our understanding of conflicts in general and how they may escalate or de-escalate. Another aspect that helps the picture of conflict studies is the interlocutors' behavior. In this study, we were focused on the pursuing party, but a conflict takes at least two parties. The practices by which interlocutors may resist against a pursuing party's CoA should be investigated in the future. These practices may especially have implications for how agency is regulated between the two participants in a conflictual interaction.

Multimodal transcript conventions (short version)
(Mondada, 2019)

* *	Descriptions of embodied actions are delimited between
+ +	two identical symbols (one symbol per participant and per type of action)
Δ Δ	that are synchronized with correspondent stretches of talk.
*--->	The action described continues across subsequent lines
---->*	until the same symbol is reached.
>>	The action described begins before the excerpt's beginning.
--->>	The action described continues after the excerpt's end.
.....	Action's preparation.
----	Action's apex is reached and maintained.
,,,,,	Action's retraction.
ric	Participant doing the embodied action is identified when (s)he is not the speaker.
fig	The exact moment at which a screen shot has been taken
#	is indicated with a specific sign (#) showing its position within turn at talk.

Adapted summary of GAT 2 transcription conventions
(Selting, et al. 2011)

Sequential structure

`[ ]` `[ ]`	overlap and simultaneous talk

Pauses

`(.)`	micro pause, estimated, up to 0.2 sec duration appr.
`(0.5)(2.0)`	Measured pause of appr. 0.5/2.0 sec. duration (to tenth of a second)

Laughter

`haha`
`hehe`
`hihi`
	syllabic laughter

Other conventions

`((coughs))`	non-verbal vocal actions and events non-verbal vocal actions and events with indication of scope
`<<coughing>`	
→	refers to a line of transcript relevant in the argument

Sequential Structure

=	fast, immediate continuation with a new turn or segment (latching)

Other segmental conventions

`:`	lengthening, by about 0.2-0.5 sec.
`::`	lengthening, by about 0.5-0.8 sec.
`:::`	lengthening, by about 0.8-1.0 sec.

Final pitch movements of intonation phrases

?	rising to high
,	rising to mid
–	level
;	falling to mid
.	falling to low

Other conventions

`<<surprised> >` interpretive comment with indication of scope

Pitch jumps

↑	smaller pitch upstep
↓	smaller pitch downstep
↑↑	larger pitch upstep

↓↓	larger pitch downstep

Changes in pitch register

`<<l> >`	lower pitch register
`<<h> >`	higher pitch register

Loudness and tempo changes, with scope

`<<f>`	`>`	forte, loud
`<<ff>`	`>`	fortissimo, very loud
`<<p>`	`>`	piano, soft
`<<pp>`	`>`	pianissimo, very soft
`<<aal>`	`>`	allegro, fast
`<<dim>`	`>`	diminuendo, increasingly softer

Changes in voice quality and articulation with scope

`<<creaky> >`	glottalized
`<<whispery> >`	change in voice quality as stated

List of glossing abbreviations

adapted from Leipzig Glossing Rules (2015)

1	first person
2	second person
3	third person
PRT	particle
OBJ.MAR	object marker
PRG	progressive
PRF	perfect
PL	plural
SG	singular
PST	past

References

Afshari Saleh, R. (2020). Mock Aggression: Navigating Affiliation and Disaffiliation in Interaction. *Research on Language and Social Interaction*, *53*(4), 481–499.

Afshari Saleh, R. (2022). Orientations to accountability through address forms: the case of Persian V-pronoun addressing an intimate co-interactant, *the International Society for Conversation Analysis* (8). Retrieved from URL https://www.conversationanalysis.org/orientations-to-accountability-through-address-forms-the-case-of-persian-v-pronoun-addressing-an-intimate-co-interactant/

Afshari Saleh, R., Teleghani-Nikazm, C., Monfaredi, E., & Rezaee, P. (2022). *Managing coherence in Interaction: The case of Persian 'khob'* [Manuscript submitted for publication]. Department of Language and Linguistic Science, University of York.

Ahearn, L. M. (2001). Language and Agency. *Annual Review of Anthropology*, *30*, 109–137.

Aikhenvald, A. Y. (2010). *Imperatives and Commands*. Oxford University Press.

Antaki, C., & Kent, A. (2012). Telling people what to do (and, sometimes, why): Contingency, entitlement and explanation in staff requests to adults with intellectual impairments. *Journal of Pragmatics*, *44*(6), 876–889.

Ambrazaitis, G., & House, D. (2017). Multimodal prominences: Exploring the patterning and usage of focal pitch accents, head beats and eyebrow beats in Swedish television news readings. *Speech Communication*, *95*, 100–113.

Antaki, C. (1994). *Explaining and Arguing*. London: Sage.

Anvari, H. & Ahmadi Givi, H. (2011). *Dastoore zabane farsi 2* [Grammar of Persian 2]. Tehran: Fatemi.

Aronsson, K., & Cekaite, A. (2011). Activity contracts and directives in everyday family politics. *Discourse & Society*, *22*(2), 137–154.

Atkinson, J. M., & Drew, P. (1979). *Order in court: The organization of verbal interaction in judicial settings.* London: Macmillan.

Auer, P. (1996). On the prosody and syntax of turn-taking In E. Couper-Kuhlen, M. Selting (Eds). *Prosody and Conversation* (pp. 57–100). Cambridge, U.K.: Cambridge University Press.

Auer, & Lindström. (2021). On agency and affiliation in second assessments: German and Swedish opinion verbs in talk-in-interaction. In J. Lindström, R. Laury, A. Peräkylä, & M. Sorjonen, (Eds.) *Intersubjectivity in Action*: *Studies in language and social interaction,* (pp. 81-107). Amsterdam, John Benjamins.

Austin, J. L. (1975). *How to do things with words*. Oxford university press.

Bateson, G. (1972). *Steps to an ecology of mind*. Chicago, IL: University of Chicago Press.

Betz, E. M., Taleghani-Nikazm, C., Drake, V., & Golato, A (2013), Third position repeats in German: The case of repair- and request-for-information sequences, *Gesprächsforschung: Online-Zeitschrift zur verbalen Interaktion*, 14, 133–166.

Bilmes, J. (1988). The concept of preference in conversation analysis. *Language in Society,* 17, 161–181.

Bilmes, J. (1992), Dividing the rice: a microanalysis of the mediators role in a Northern Thai negotiation, *Language in Society*, vol. 21, no. 4, pp. 569–602.

Blythe, J., Gardner, R., Mushin, I., & Stirling, L. (2018). Tools of Engagement: Selecting a Next Speaker in Australian Aboriginal Multiparty Conversations. *Research on Language and Social Interaction*, *51*(2), 145–170.

Bolden, G. (2017). Requests for here-and-now actions in Russian conversation. In M. Sorjonen, L. Raevaara, & E. Couper-Kuhlen (Eds). *Imperative Turns at Talk The design of directives in action* (pp. 175–211). John Benjamins Publishing Company.

Bolden, G. B., Mandelbaum, J., & Wilkinson, S. (2012). Pursuing a Response by Repairing an Indexical Reference. *Research on Language and Social Interaction*, *45*(2), 137–155.

Bolden, G. B., & Robinson, J. D. (2011). Soliciting accounts with why-interrogatives in Conversation. *The Journal of Communication*, *61*(1), 94–119.

Bourdieu P. (1977). Outline of a Theory of Practice. Cambridge, UK: Cambridge Univ. Press.

Brown, R. & Gilman, A. (1968) "The pronouns of power and solidarity." In: J. A. Fishman (ed), Readings in the Sociology of Language *Style in Language*. The Netherlands: Mouton Publishers, 252–276.

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge, England: Cambridge University Press.

Brown, L. & Prieto, P. (2021). "Gesture and prosody in multimodal communication." In: M. Haugh, D. Kadar, & M. Terkourafi (Eds.), *The Cambridge Handbook of Sociopragmatics*, 430-453. Cambridge: Cambridge University Press.

Butler, C. W., Danby, S., & Emmison, M. (2011). Address Terms in Turn Beginnings: Managing Disalignment and Disaffiliation in Telephone Counseling. *Research on Language and Social Interaction*, *44*(4), 338–358.

Butt, M. (2010). The light verb jungle: Still hacking away. In M. Amberber, B. Baker, & M. Harvey (Eds.), *Complex Predicates: Cross-linguistic Perspectives on Event Structure* (pp. 48-78). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511712234.004

Button, G., & Casey, N. (1985). Topic Nomination and Topic Pursuit. *Human Studies*, *8*(1), 3–55.

Campbell, C. (2009). Distinguishing the power of agency from agentic power: A note on Weber and the "black box" of personal agency. *Sociological Theory*, *27*(4), 407–418.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81–105.

Cantarutti, M. N. (2021). Co-animation and the Multimodal Management of Contextualisation Problems when Jointly "Doing Being" Others. *Social Interaction. Video-Based Studies of Human Sociality*, *4*(4). https://doi.org/10.7146/si.v4i4.128166

Chomsky, N. (1993). *Lectures on Government and Binding: The Pisa Lectures*. Walter de Gruyter.

Clift, R. (2020). Stability and visibility in embodiment: The "Palm Up" in interaction. *Journal of Pragmatics*, *169*, 190–205.

Clift, R., & Pino, M. (2020). Turning the Tables: Objecting to Conduct in Conflict Talk. *Research on Language and Social Interaction*, *53*(4), 463–480.

Clayman, S. E. (2002). Sequence and solidarity. In S. R. Thye & E. J. Lawler (Eds.), *Group cohesion, trust and solidarity* (Vol. 19, pp. 229–253). Oxford, England: Elsevier Science Ltd.

Clayman, S. E. (2010). Address terms in the service of other actions: The case of news interview talk, *Discourse & Communication*, 4(2), 161–183.

Clayman, S. E. (2013). Agency in response: The role of prefatory address terms. *Journal of Pragmatics*, *57*, 290–302.

Clayman, S., & Heritage, J. (2014). *Benefactors and beneficiaries: benefactive status and stance in the management of offers and requests*. 55–86.

Clyne, M., Kretzenbacher, H., Norrby, C., & Schüpbach, D. (2006). Perceptions of variation and change in German and Swedish address. *Journal of Sociolinguistics*, *10*(3), 287–319.

Cohen, A., & Hart, J. T. (1968). On the anatomy of intonation. *Lingua. International Review of General Linguistics. Revue Internationale de Linguistique Generale*, *19*(1-2), 177–192.

Collins, R. (2009). *Violence: a micro-sociological theory*. Princeton, N.J.; Oxford: Princeton University Press.

Coon, J., Massam, D., & Travis, L. D. (2017). *The Oxford Handbook of Ergativity*. Oxford University Press.

Coulter, J. (1990). Elementary properties of argument sequences. In G. Psathas (Ed.), *Interaction competence* (pp. 181–204). Lanham, MD: University Press of America.

Couper-Kuhlen, E. (1992), Contextualizing discourse: the prosody of interactive repair, In P. Auer & A. di Luzio (Eds.) *The Contextualization of Language* (pp. 337–364). Amsterdam, John Benjamins.

Couper-Kuhlen, E. (2004). Analyzing language in interaction: the practice of never mind. *English Language and Linguistics*, *8*(2), 207–237.

Couper-Kuhlen, E. (2010). Commentary on Stivers and Rossano: "Mobilizing Response." *Research on Language and Social Interaction*, *43*(1), 32–37.

Couper-Kuhlen, E. (2012). Exploring affiliation in the reception of conversational complaint stories. In A. Peräkylä & M.-L. Sorjonen (Eds.), *Emotion in Interaction.* (pp. 113–146). Oxford, New York: Oxford University Press.

Couper-Kuhlen, E., & Ono, T. (2007). "Incrementing" in conversation. A comparison of practices in English, German and Japanese. *Pragmatics*, *17*(4), 513–552.

Craven, A., & Potter, J. (2010). Directives: Entitlement and contingency in action. *Discourse Studies*, *12*(4), 419–442.

Cruttenden, A. (1997). Intonation. Cambridge: Cambridge University Press.

Crystal, D. (2008). A Dictionary of Linguistics and Phonetics, 6th ed. Malden, MA: Blackwell Publishing.

Curl, T. S. (2005). Practices in Other-Initiated Repair Resolution: The Phonetic Differentiation of "Repetitions." *Discourse Processes*, *39*(1), 1–43.

Curl, T. S., & Drew, P. (2008). Contingency and Action: A Comparison of Two Forms of Requesting. *Research on Language and Social Interaction*, *41*(2), 129–153.

Curl, T. S., Local, J., & Walker, G. (2006). Repetition and the prosody–pragmatics interface. *Journal of Pragmatics*, *38*(10), 1721–1751.

Darwin, C. (1998). *The expression of the emotions in man and animals* (3rd ed.). London: Harper Collins.

Davidson, D. (2001). *Essays on Actions and Events: Philosophical Essays Volume 1*. Clarendon Press.

Davidson, J. A. (1984). Subsequent versions of invitations, offers, requests, and proposals dealing with potential or actual rejection. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 102–128). Cambridge, England: Cambridge University.

Deppermann, A. (2013). Multimodal interaction from a conversation analytic perspective. *Journal of Pragmatics*, *46*(1), 1–7.

Deppermann, A. (2018). Instruction practices in German driving lessons: Differential uses of declaratives and imperatives. *Vigo International Journal of Applied Linguistics*, *28*(2), 265–282.

Deppermann, A., & Gubina, A. (2021). When the Body Belies the Words: Embodied Agency With darf/kann ich? ("May/Can I?") in German. *Frontiers in Communication*, *6*, 74.

Dersley, I. & Wootton, A. J. (2000). Complaint sequences within antagonistic argument, *Research on Language and Social Interaction*, 33, 375-406.

Dersley, I., & Wootton, A. J. (2001). In the heat of the sequence: Interactional features preceding walkouts from argumentative talk. *Language In Society*, *30*(4), 611–638.

Derrida. (1982). Signature Event Context." Glyph 1: 172-97. 1977b. *Limited Inc.*

Drake, V. (2015). Indexing Uncertainty: The Case of Turn-Final Or. *Research on Language and Social Interaction*, *48*(3), 301–318.

Drew, P. (1987). Po-faced receipts of teases. *Linguistics*, 25 (1), 219–253.

Drew, P. (2009), "Quit talking while I'm interrupting": A comparison between positions of overlap onset in conversation, In M. Haakana, M. Laakso, & J. Lindström, (Eds.), *Talk in Interaction: Comparative Dimensions,* (pp. 70–93). Helsinki, Finnish Literature Society (SKS).

Drew, P. (2012). Turn design. In J. Sidnell & T. Stivers (Eds). *The Handbook of Conversation Analysis* (pp. 131–149). Oxford: Wiley-Blackwell.

Drew, P., & Walker, T. (2009). Going too far: Complaining, escalating and disaffiliation. *Journal of Pragmatics*, *41*(12), 2400–2414.

Du Bois, John. 2007. The stance triangle. In *Stancetaking* In R. Englebretson (Ed) *discourse: Subjectivity, evaluation, interaction*, (pp.139–182). Amsterdam: John Benjamins.

Duranti, A. (1994). *From Grammar to Politics: Linguistic Anthropology in a Western Samoan Village*. University of California Press.

Duranti, A. (2004). *A Companion to Linguistic Anthropology*. John Wiley & Sons.

Edwards, D. (2000). Extreme Case Formulations: Softeners, Investment, and Doing Nonliteral. *Research on Language and Social Interaction*, *33*(4), 347–373.

Ekman, P. & Friesen, W. V. (1969). The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding. *Semiotica, 1(1),* 49-98.

Ekman P, Friesen W. V. (1982), Felt, false, and miserable smiles *Journal of Nonverbal Behavior*. 6: 238-252. DOI: 10.1007/BF00987191

Enfield, N. J. (2011). Sources of asymmetry in human interaction: Enchrony, status, knowledge and agency. In T. Stivers, L. Mondada, & J. Steensig (Eds.), *The morality of knowledge in conversation* (pp. 285-312). Cambridge: Cambridge University Press.

Enfield, N. J. (2013). *Relationship Thinking: Agency, Enchrony, and Human Sociality*. Oxford University Press.

Enfield, N. J., & Kockelman, P. (Eds.). (2017). *Distributed Agency*. Oxford University Press.

Englebretson, R. (Ed). 2007. *Stancetaking in discourse*. Amsterdam: John Benjamins.

Etelämäki, M. & Couper-Kuhlen, E. (2017). In the face of resistance. In M. Sorjonen, L. Raevaara, & E. Couper-Kuhlen (Eds). *Imperative Turns at Talk The design of directives in action* (pp. 215-240). John Benjamins Publishing Company.

Etelämäki, M., Voutilainen, L., & Weiste, E. (2021). Distributing Agency and Experience in Therapeutic Interaction: Person References in Therapists' Responses to Complaints. *Frontiers in Psychology*, *12*, 585321.

Fagen, R. (1981). *Animal Play Behavior*. New York; Oxford: Oxford University Press.

Farshidvard, K. (2009). *Dastoore mokhtasare emrooz bar pâyeye zabânshenâsie jadid* [A Brief Grammar of Today based on New Linguistics]. Tehran: Sokhan.

Floyd, S., Rossi, G., & Enfield, N. J. (2020). *Getting others to do things: A pragmatic typology of recruitments*. https://doi.org/10.5281/zenodo.4017493\

Ford, C. E., Fox, B. A., & Thompson, S. A.(1996), Practices in the construction of turns: The "TCU" revisited, *Pragmatics (6)*3, 427–454.

Ford, C. E., & Thompson, S. A. (1996). Interactional units in conversation: syntactic, intonational, and pragmatic resources for the management of turns, In E. Ochs, E. A. Schegloff, & S. A. Thompson (Eds). *Interaction and Grammar* (pp. 134–184). Cambridge: Cambridge University Press.

French, P., & Local, J. (1983). Turn-competitive incomings. *Journal of Pragmatics*, *7*(1), 17–38.

Frick, M., & Palola, E. (2022). Deontic Autonomy in Family Interaction. *Social Interaction. Video-Based Studies of Human Sociality*, *5*(1). https://doi.org/10.7146/si.v5i2.130870

Gago, P. C. (2017). Narrative accounts and conflict escalation in legal family mediation. *Journal of Applied Linguistics and Professional Practice,* 10(3), 361–387.

Garcia, A. (1991). Dispute resolution without disputing: How the interactional organization of mediation hearings minimizes argument. *American Sociological Review*, 56 (6), 818-835.

Garfinkel, H. (1967). *Studies in ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.

Ghaffari, M. (2020). Persian as an interlanguage. In P. Shabani-Jadidi (Ed). *The Routledge Handbook of Second Language Acquisition and Pedagogy of Persian* (pp. 546-566). Abingdon: Routledge.

Gibbon, D. (2009). Gesture Theory is Linguistics: On Modelling Multimodality as Prosody. *PACLIC*.

Gibson, D. R. (2000). Seizing the Moment: The Problem of Conversational Agency. *Sociological Theory*, *18*(3), 368–382.

Giddens, A. (1979). *Central problems in social theory: Action, structure, and contradiction in social analysis* (Vol. 241). Univ of California Press.

Goffman, E. (1961). *Asylums: Essays on the Social Situation of Mental Patients and Other Inmates*. AldineTransaction.

Goffman, E. (1963). *Behavior in public places.* New York: The Free Press.

Goffman, E. (1967). *Interaction ritual: essays on face-to-face behavior.* New York: Pantheon Books.

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

Goffman, E. (1986). *Frame analysis: an essay on the organization of experience*. Boston: Northeastern University Press.

Goffman, E. (2010). *Relations in public: microstudies of the public order.* New Brunswick; London: Transaction Publishers.

Goldin-Meadow, S., & Brentari, D. (2017). Gesture, sign, and language: The coming of age of sign language and gesture studies. *The Behavioral and Brain Sciences*, *40*, e46.

Gonzalez Temer, V., & Ogden, R. (2021). Non-convergent boundaries and action ascription in multimodal interaction. *Open Linguistics*, *7*(1), 685–706.

Goodwin, C. (1981), *Conversational Organization: Interaction between Speakers and Hearers*, London, Academic Press.

Goodwin, C. (1994). Professional Vision. *American Anthropologist*, 96(3), 606–633.

Goodwin, C. (2018). Why Multimodality? Why Co-Operative Action? (transcribed by J. Philipsen). *Social Interaction Video-Based Studies of Human Sociality*, *1*(2). https://doi.org/10.7146/si.v1i2.110039

Goodwin, M.H. (1980), He said/she said: formal cultural procedures for the construction of a gossip dispute activity, *American Ethnologist*, 7(4), 674–694.

Goodwin, M. H. (1982). Processes of dispute management among urban black children. *American Ethnologist*, *9*(1), 76–96.

Goodwin, M. H. (1983), Aggravated correction and disagreement in children's conversations", *Journal of Pragmatics*, 7(6), 657–677.

Goodwin, M. H., Cekaite, A. & Goodwin, C. Tulbert, (2012), Emotion as stance, In A. Peräkylä & M. Sorjonen, (Eds.). *Emotion in interaction* (pp. 16-41). Oxford: Oxford University Press.

Goodwin, C. & Goodwin, M. H. (1990). Interstitial argument, In A. Grimshaw (Ed) *Conflict Talk*, Cambridge, Cambridge University Press, pp. 85–117.

Goodwin, C., & Goodwin, M. H. (2005). Participation. In A. Duranit (Ed). *A Companion to Linguistic Anthropology* (pp. 222–244). Oxford: Blackwell.

Glenn, P., & Kuttner, R. (2013). Dialogue, dispute resolution, and talk-in-interaction: On empirical studies of ephemeral phenomena. *Negotiation and Conflict Management Research*, *6*(1), 13–31.

Glenn, P. (2019). Conflict Interaction: Insights from conversation analysis. In M. Evans, L. Jeffries, & J. O'Driscoll (Eds.), *The Routledge Handbook of Language in Conflict* (pp. 215–245). Routledge. https://doi.org/10.4324/9780429058011

Griffiths, P., Merrison, A.J., & Bloomer, A. (Eds.). (2010). *Language in Use: A Reader* (1st ed.). Routledge. https://doi.org/10.4324/9781003060994

Gruber, H. (1998). Disagreeing: Sequential placement and internal structure of disagreements in conflict episodes. *Text, 18,* 467–503.

Guellaï, B., Langus, A., & Nespor, M. (2014). Prosody in the hands of the speaker. *Frontiers in Psychology*, *5*, 700.

Gussenhoven, C. (2004). *The phonology of tone and intonation.* Cambridge University Press.

Hagoort, P., & van Berkum, J. (2007). Beyond the sentence given. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *362*(1481), 801–811.

Hall, E. T. (1966). *The Hidden Dimension*. Anchor Books.

Haugh, M., & Sinkeviciute, V. (2019). Offence and conflict talk. In M. Evans, L. Jeffries, & J. O'Driscoll (Eds). *The Routledge Handbook of Language in Conflict* (pp.196–214). Routledge Handbooks Online. https://doi.org/10.4324/9780429058011-12

Hauser, E. (2019). Upgraded self-repeated gestures in Japanese interaction. *Journal of Pragmatics*, *150*, 180–196.

Heath, C. (1984), Talk and recipiency: sequential organization in speech and body movement, In J. M. Atkinson, & J. Heritage (Eds.). *Structures of Social Action: Studies in Conversation Analysis* (pp. 247–265). Cambridge, U.K., Cambridge University Press.

Heath, C., & Luff, P. (2013). Embodied action and organisational interaction: Establishing contract on the strike of a hammer. *Journal of Pragmatics*, *46*(1), 24–38.

Heinemann, T. (2008). Questions of accountability: yes—no interrogatives that are unanswerable. *Discourse Studies*, *10*(1), 55–71.

Heinemann, T. (2009). *Managing unavoidable conflicts in caretaking of the elderly: humor as a mitigating resource*. *2009*(200), 103–127.

Heinemann, T., & Steensig, J. (2017). Three imperative action formats in Danish talk-in-interaction: The case of imperative + modal particles bare and lige. In M. Sorjonen, L. Raevaara, & E. Couper-Kuhlen (Eds). *Imperative Turns at Talk The design of directives in action* (pp. 139-173). John Benjamins Publishing Company.

Hepburn, A., & Potter, J. (2011). Threats: Power, family mealtimes, and social influence. British Journal of Social Psychology, 50(1), 99–120. doi:10.1348/014466610X500791

Heritage, J. (1984). *Garfinkel and Ethnomethodology*. Cambridge: Polity Press.

Heritage, J. (2012). Epistemics in Action: Action Formation and Territories of Knowledge. *Research on Language and Social Interaction*, *45*(1), 1–29.

Heritage, J. & Atkinson, M. (1984), "Introduction", In J. Heritage, J. M. Atkinson (Eds.), *Structures of Social Action: Studies in Conversation Analysis,* (pp. 1–15). Cambridge: Cambridge University Press.

Heritage J & Clayman. (2010). *Talk in Action: Interactions, Identities and Institutions*. Oxford: Wiley-Blackwell.

Heritage, J., & Raymond, G. (2005). The terms of agreement: Indexing epistemic authority and subordination in talk-in-interaction. *Social Psychology Quarterly*, *68*(1), 15–38.

Heritage, J., & Raymond, G. (2012). Navigating epistemic landscapes: Acquiescence, agency and resistance in responses to polar questions. *Questions: Formal, Functional and Interactional Perspectives*, 179–192.

Hoey, E. M. (2014). Sighing in Interaction: Somatic, Semiotic, and Social. *Research on Language and Social Interaction*, *47*(2), 175–200.

Hoey, E. M. (2015). Lapses: How people arrive at, and deal with, discontinuities in talk. *Research on Language and Social Interaction*, *48*(4), 430-453.

Hoey, E. M. & Kendrick, K. H. (2017), Conversation analysis. In A. M. B. De Groot, & P. Hagoort (Eds). *Research Methods in Psycholinguistics and the Neurobiology of Language: A Practical Guide*. (pp. 151-173). Wiley Blackwell.

Hoey, E. M., Hömke, P., Löfgren, E., Neumann, T., Schuerman, W. L., & Kendrick, K. H. (2021). Using expletive insertion to pursue and sanction in interaction. *Journal of Sociolinguistics*, *25*(1), 3–25.

Holler, J., Kendrick, K. H., Casillas, M., & Levinson, S. C. (2015). Editorial: Turn-Taking in Human Communicative Interaction. *Frontiers in Psychology*, *6*, 1919.

Holler, J., & Levinson, S. C. (2019). Multimodal language processing in human communication. *Trends in Cognitive Sciences*, *23*(8), 639-652.

Holt, L. (2012). Using Laugh Responses to Defuse Complaints. *Research on Language and Social Interaction*, *45*(4), 430–448.

Holt, E. (2013). 'There's many a true word said in jest': seriousness and non-seriousness in interaction. In Glenn, P., & E. Holt (Eds.) *Studies of Laughter in Interaction*. (pp. 69-90). Bloomsbury, London.

Hutchby, I. (1992), Confrontation talk: Aspects of interruption in argument sequences on talk radio, *Text* (12)3, 343–371.

Hutchby, I. (1996). *Confrontation talk: Arguments, asymmetries and power on talk radio*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Hutchby, I. & Wooffitt, R. (1998). *Conversation Analysis: Principles, Practices and Applications*. Cambridge, Polity Press.

Jaffe, A. (2009). *Stance: Sociolinguistic Perspectives*. Oxford University Press.

Jasbi, M. (2020). Some unique semantic properties of Persian. In P. Shabani-Jadidi (Ed). *The Routledge Handbook of Second Language Acquisition and Pedagogy of Persian* (pp. 134-146). Abingdon: Routledge.

Jefferson, G. (1981). The abominable 'ne?': An exploration of post-response-initiation response-solicitation. In P. Schröder, & H. Steger (Eds), Dialogforschung (pp. 53–88). Düsseldorf.

Jefferson, G., Sacks, H. & Schegloff, E. A. (1987). Notes on Laughter in the Pursuit of Intimacy. In G. Button & J.R.E. Lee (Eds.) *Talk and social Organisation*. (pp.152-205). Clevedon: Multilingual Matters.

Kamunen, A. (2019). How to Disengage: Suspension, Body Torque, and Repair. *Research on Language and Social Interaction*, *52*(4), 406–426.

Kaneyasu, M. (2021). In Pursuit of Alignment and Affiliation: The Practice of Anchoring Shared Knowledge in Japanese Conversation. *Discourse Processes*, *58*(2), 93–112.

Karimi, S. (2018). Generative Approaches to Syntax. In A. S. A. Shabani-Jadidi (Ed.), *The Oxford Handbook of Persian Linguistics*. Oxford University Press.

Kaukomaa, T., Peräkylä, A., & Ruusuvuori, J. E. (2014). Foreshadowing a problem: turn-opening frowns in conversation. *Journal of Pragmatics*, *71*, 132–147.

Keel, S. (2015). Young children's embodied pursuits of a response to their initial assessments. *Journal of Pragmatics*, *75*, 1–24.

Keevallik, L. (2014), Turn organization and bodily-vocal demonstrations, *Journal of Pragmatics*, 65, pp. 103–120.

Keevallik, L. (2018), What does embodied interaction tell us about grammar? , *Research on Language and Social Interaction, 51*(1), 1–21.

Kendon, A. (1972). Some Relationships Between Body Motion and Speech: An Analysis of an Example. In A. W. Siegman & B. Pope (Eds.), *Studies in Dyadic Communication* (Vol. 7, pp. 177–210). Pergamon.

Kendon, A. (1975). Gesticulation, speech, and the gesture theory of language origins. *Sign Language Studies*, *9*, 349–373.

Kendon, Adam. (1980). Gesticulation and speech: Two aspects of the process of utterance in M. Key (Ed), *The Relationship of Verbal and Nonverbal Communication* (pp. 207-227). Mouton, The Hague.

Kendon, A. (1990). *Conducting Interaction: Patterns of Behavior in Focused Encounters.* Cambridge: Cambridge University Press.

Kendon, A. (2004). *Gesture: visible action as utterance*. Cambridge: Cambridge University Press.

Kendon, A. (2014). Semiotic diversity in utterance production and the concept of "language." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1651), 20130293.

Kendrick, K. H. (2015). Other-initiated repair in English. *Open Linguistics*, *1*(1). https://doi.org/10.2478/opli-2014-0009

Kendrick, K. H. (2020). Recruitment in English: A quantitative study, In S. Floyd, G. Rossi, N. J. Enfield (Eds). *Getting others to do things: a pragmatic typology of recruitments* (pp. 93-146). Language Science Press.

Kendrick, K. H. (2021). The "Other" side of recruitment: Methods of assistance in social interaction. *Journal of Pragmatics*, *178*, 68–82.

Kendrick, K. H., Brown, P., Dingemanse, M., Floyd, S., Gipper, S., Hayano, K., Hoey, E., Hoymann, G., Manrique, E., Rossi, G., & Levinson, S. C. (2020). Sequence organization: A universal infrastructure for social action. *Journal of Pragmatics*, *168*, 119–138.

Kendrick, K. H. & Drew, P. (2014). The putative preference for offers over requests. In P. Drew & E. Couper-Kuhlen (Eds.), *Requesting in Social Interaction.* Amsterdam/Philadelphia, John Benjamins, pp. 87–114.

Kendrick, K. H., & Drew, P. (2016). Recruitment: Offers, Requests, and the Organization of Assistance in Interaction. *Research on Language and Social Interaction*, *49*(1), 1–19.

Kendrick, K. H., & Holler, J. (2017). Gaze direction signals response preference in conversation. *Research on Language and Social Interaction*, *50*(1), 12–32.

Kendrick, K. H., & Torreira, F. (2015). The Timing and Construction of Preference: A Quantitative Study. *Discourse Processes*, *52*(4), 255–289.

Kent, A., & Kendrick, K. H. (2016). Imperative directives: orientations to accountability. *Research on Language and Social Interaction*, *49*(3), 272–288.

Kidwell, M. (2005). Gaze as social control: how very young children differentiate "the look" from a "mere look" by their adult caregivers. *Research on Language and Social Interaction*, *38*(4), 417–449.

Kim. K. (2007). Sequential organization of post-predicate elements in Korean conversation. *Pragmatics* 17(4), 573–603.

Kim, S. H. & Kim, M. S. (2020), Requesting here-and-now actions with two imperative formats in Korean interaction, *In C. Taleghani-Nikazm, E. Betz, & P. Golato (Eds). Mobilizing Others: Grammar and lexis within larger activities (*pp. 19-46). Amsterdam / Philadelphia: John Benjamins Publishing Company.

Kita, S., van Gijn, I., & van der Hulst, H. (1998). Movement phases in signs and co-speech gestures, and their transcription by human coders. In I. Wachsmuth & M. Fröhlich (Eds). *Gesture and Sign Language in Human-Computer Interaction*, (pp.23–35). Springer Berlin Heidelberg.

Kockelman, P. (2007). The Relation between Meaning, Power, and Knowledge. *Current Anthropology*, *48*(3), 375–401.

Koenig, C. J. (2011). Patient resistance as agency in treatment decisions. *Social Science & Medicine*, *72*(7), 1105–1114.

König, E. & Siemund, P. (2007). Speech act distinctions in grammar. In T. Shopen, (Ed.) *Language Typology and Syntactic Description* (2nd ed.). Cambridge: Cambridge University Press. pp. 276 - 324 doi:10.1017/CBO9780511619434

Kotthoff, H. (1993). Disagreement and concession in disputes: On the context sensitivity of preference structures. *Language in Society, 22,* 193–216.

Krivokapić J. (2014). Gestural coordination at prosodic boundaries and its role for prosodic structure and speech planning processes. *Philosophical transactions of the Royal Society of London*. Series B, Biological sciences, 369(1658), 20130397. https://doi.org/10.1098/rstb.2013.0397

Laver, J. (1994). Principles of Phonetics. Cambridge: Cambridge University Press.

Lee, J. (2019). Scaling as an argumentative resource in television talk shows. *Journal of Pragmatics*, *150*, 133–149.

Lehiste, I. (1970). Suprasegmentals. Cambridge, MA: MIT Press.

Leipzig Glossing Rules. (2015). Max Planck Institute for Evolutionary Anthropology. https://www.eva.mpg.de/lingua/resources/glossing-rules.php.

Lerner, G. H. (1996). Finding "Face" in the Preference Structures of talk-in-interaction. *Social Psychology Quarterly*, *59*(4), 303–321.

Lerner, G. H. (2004). On the Place of Linguistic Resources in the Organization of Talk-in-Interaction: Grammar as Action in Prompting a Speaker to Elaborate. *Research on Language and Social Interaction*, *37*(2), 151–184.

Lerner, G. H., Bolden, G. B., Hepburn, A., & Mandelbaum, J. (2012). Reference Recalibration Repairs: Adjusting the Precision of Formulations for the Task at Hand. *Research on Language and Social Interaction*, *45*(2), 191–212.

Levinson, S. C. (2013). Action formation and ascription. In J. Sidnell & T. Stivers (Eds). *The Handbook of Conversation Analysis* (pp. 103–130). Oxford, U.K., Wiley-Blackwell.

Lieberman, P. (1967). Intonation, perception, and language. *M.I.T. Research Monograph*, *38*, 210.

Lindström, A. (2005). Language as social action: a study of how senior citizens request assistance with practical tasks in the Swedish home help service, In A. Hakulinen, M. Selting, (Eds.) *Syntax and Lexis in Conversation: Studies on the Use of Linguistic Resources in Talk-in-Interaction* (pp. 209–230). Amsterdam, John Benjamins.

Lindström, A. and Sorjonen, M.-L. (2013). Affiliation in Conversation. In J. Jack Sidnell & T. Stivers (Eds). *The Handbook of Conversation Analysis,* Oxford, U.K., Wiley-Blackwell, pp. 350- 369. https://doi.org/10.1002/9781118325001.ch17

Lloyd, M. (2017a). On the Way to Cycle Rage: Disputed Mobile Formations. *Mobilities*, *12*(3), 384–404.

Lloyd, M. (2017b). When Rules Go Awry: A Single Case Analysis of Cycle Rage. *Human Studies*, *40*(4), 681–706.

Lloyd, M., & Mlynář, J. (2021). Hand-ling "Road Rage": Embodiment in Conflict on the Move. Social Interaction. *Video-Based Studies of Human Sociality*, 4(4). https://doi.org/10.7146/si.v4i4.125177

Local, J. (2003). Variable domains and variable relevance: interpreting phonetic exponents. *Journal of Phonetics*, *31*(3), 321–339.

Local, J., Auer, P., & Drew, P. (2010), Retrieving, redoing and resuscitating turns in conversation. In D. Barth-Weingarten E. Reber, & M. Selting, (Eds). *Prosody in Interaction* (pp. 131-160). Amsterdam, John Benjamins

Local, J., & Kelly, J. (1986). Projection and "Silences": Notes on Phonetic and Conversational Structure. *Human Studies*, *9*(2/3), 185–204.

Local, J. K. Kelly, J., & W. H. G. Wells. (1986). Towards a Phonology of Conversation: Turn-Taking in Tyneside English. *Journal of Linguistics*, *22*(2), 411–437.

Local, J., & Walker, G. (2005). Methodological imperatives for investigating the phonetic organization and phonological structures of spontaneous speech. *Phonetica*, *62*(2-4), 120–30.

Locke, J., & Woolhouse, R. S. (1997). *An essay concerning human understanding*. London, New York.

Loehr, D. P. (2012). Temporal, structural, and pragmatic synchrony between intonation and gesture. *Laboratory Phonology*, *3*(1), 71–89.

Mahootian, S. (1997). *Persian: Descriptive Grammars*. Routledge.

Marshall, G. (2009). The concise Oxford Dictionary of Sociology.

Maynard, D. W. (1985). How children start arguments. *Language in Society*, 14, 1–29.

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

Megerdoomian, K. (2020). Linguistic competence of Persian heritage versus second language speakers. In P. Shabani-Jadidi (Ed). *The Routledge Handbook of Second Language Acquisition and Pedagogy of Persian* (pp. 53-99). Abingdon: Routledge.

Miller, S. (1973). Ends, Means, and Galumphing: Some Leitmotifs of Play. *American Anthropologist*, *75*(1), 87–98.

Mondada, L. (2007). Multimodal resources for turn-taking: pointing and the emergence of possible next speakers. *Discourse Studies*, *9*(2), 194–225.

Mondada, L. (2014). The local constitution of multimodal resources for social interaction. *Journal of Pragmatics*. 65, 137-156.

Mondada, L. (2016). Challenges of multimodality: Language and the body in social interaction, *Journal of Sociolinguistics*, 20(3), 336–366.

Mondada, L. (2017). Precision timing and timed embeddedness of imperatives in embodied courses of action. In M. Sorjonen, L. Raevaara, & E. Couper-Kuhlen (Eds). *Imperative Turns at Talk The design of directives in action* (pp. 65-101). John Benjamins Publishing Company.

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

Mondada, L. (2019a). Conventions for transcribing multimodality. Retrieved November 11, 2019 from https://www.lorenzamondada.net/multimodal-transcription.

Mondada, L. (2019b). Contemporary issues in conversation analysis: Embodiment and materiality, multimodality and multisensoriality in social interaction. *Journal of Pragmatics*, *145*, 47–62.

Morreall, J. (2016). Philosophy of Humor. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Retrieved 10 November from https://plato.stanford.edu/archives/win2016/entries/humor/

Muntigl, P., & Turnbull, W. (1998). Conversational structure and facework in arguing. *Journal of Pragmatics, 29,* 225–256.

Nanbakhsh, G. (2012). Moving beyond T/V pronouns of power and solidarity in interaction : Persian agreement mismatch construction. In *Linguistica* (Vol. 52, Issue 1, pp. 253–266). https://doi.org/10.4312/linguistica.52.1.253-266

Nespor, M. & Vogel, I. (1986). *Prosodic Phonology*. Dordrecht: Foris Publications.

Nevile, M. (2015). The Embodied Turn in Research on Language and Social Interaction. *Research on Language and Social Interaction*, *48*(2), 121–151.

Nishizaka, A. & Hayano, K. (2015), Conversational preference, In K. Tracy, C. Ilie, & T. Sandel, (Eds.), *The International Encyclopedia of Language and Social Interaction.* London, John Wiley & Sons, vol. 1, pp. 229–236.

Nolan, F. (2003). Intonational equivalence: an experimental evaluation of pitch scales, 771–774.

Ochs, E. (1996). Linguistic resources for socializing humanity. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 407–437). Cambridge University Press.

Ogden, R. (2006). Phonetics and social action in agreements and disagreements. *Journal of Pragmatics*, *38*(10), 1752–1775.

Ogden, R. (2010). Prosodic constructions in making complaints. In D. Barth-Weingarten, E. Reber, & M. Selting (Eds.), *Prosody in interaction* (pp. 81–104). Amsterdam: John Benjamins.

Ogden (2017). Iconicity in spoken conversation [Unpublished manuscript].

Ogden, R. (2021). The Phonetics of Talk in Interaction. In R. Knight & J. Setter (Eds.), *The Cambridge Handbook of Phonetics* (Cambridge Handbooks in Language and Linguistics, pp. 657-681). Cambridge: Cambridge University Press. doi:10.1017/9781108644198.027

Okada, M. (2018). Imperative Actions in Boxing Sparring Sessions. *Research on Language and Social Interaction*, *51*(1), 67–84.

Ong, B., Barnes, S., & Buus, N. (2021). Downgrading Deontic Authority in Open Dialogue Reflection Proposals: A Conversation Analysis. *Family Process*, *60*(4), 1217–1232.

Peräkylä, A. (2002). Agency and Authority: Extended Responses to Diagnostic Statements in Primary Care Encounters. *Research on Language and Social Interaction*, *35*(2), 219–247.

Peräkylä, A., & Ruusuvuori, J. (2012). *Facial Expression and Interactional Regulation of Emotion*. In A. Peräkylä & M. Sorjonen (Eds) *Emotion in interaction* (pp. 64–91). Oxford: Oxford University Press.

Perniss, P. (2018). Why We Should Study Multimodal Language. *Frontiers in Psychology*, *9*, 1109.

Persson, R. (2013). Intonation and sequential organization: formulations in French talk-in-interaction. *Journal of Pragmatics, 57*, 19-38.

Plank, F. (1979). 'Ergativity, syntactic typology and universal grammar: Some past and present viewpoints,' in F. Plank (Ed.), *Ergativity. Towards a theory of grammatical relations*. (pp. 3–36). London: Academic Press.

Pillet-Shore, D. (2010). Making Way and Making Sense: Including Newcomers in Interaction. *Social Psychology Quarterly*, *73*(2), 152–175.

Pillet-Shore, D. (2012). Greeting: Displaying Stance Through Prosodic Recipient Design. *Research on Language and Social Interaction*, *45*(4), 375–398.

Pillet-Shore, D. (2017). Preference organization, In H. Giles, J. & Harwood (Eds), *The Oxford Research Encyclopedia of Communication.* Oxford, Oxford University Press.

Plug, L. (2014). On (or not on) the 'upgrading–downgrading continuum': The case of 'prosodic marking' in self-repair. In D. Barth-Weingarten, & B. Szczepek Reed (Eds), *Prosodie und Phonetik in der Interaktion / Prosody and phonetics in interaction* (pp. 70-86). Göttingen, Verlag für Gesprächsforschung.

Pomerantz, A. (1984a). Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 57–101). Cambridge, England: Cambridge University Press.

Pomerantz, A. (1984b). Pursuing a response. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 152–164). Cambridge, England: Cambridge University Press.

Pomerantz, A. (1986). Extreme case formulations: a way of legitimizing claims. *Human Studies*, *9*(2), 219–229.

Pomerantz A. & Heritage, J. (2013), Preference. In J. Jack Sidnell & T. Stivers (Eds). *The Handbook of Conversation Analysis,* Oxford, U.K., Wiley-Blackwell, pp. 210–228.

Raevaara, L. (2017). Adjusting the design of directives to the activity environment: Imperatives in Finnish cooking club interaction. In M. Sorjonen, L. Raevaara, & E. Couper-Kuhlen (Eds). *Imperative Turns at Talk The design of directives in action* (pp. 381–410). John Benjamins Publishing Company.

Raymond, C. W., Clift, R., & Heritage, J. (2021). Reference without anaphora: on agency through grammar. *Linguistics and Philosophy*, *59*(3), 715–755.

Raymond, G. T., & Lerner, G. H. (2014). A body and its involvements: Adjusting Action for Dual Involvements. In *Multiactivity in Social Interaction* (pp. 227–246).

Reber, E. (2012). *Affectivity in Interaction: Sound objects in English*. John Benjamins Publishing Company.

Rendle-Short, J. (2010). "Mate" as a term of address in ordinary interaction. *Journal of Pragmatics*, *42*(5), 1201–1218.

Rickford, R. (2019) *Turn-taking, turn-organisation and social action: an examination of increments in post-other-talk position.* Unpublished. University of York MA Social Research Dissertation

Robinson, J. D. (1998). Getting Down to Business: Talk, Gaze, and Body Orientation During Openings of Doctor-Patient Consultations. *Human Communication Research*, *25*(1), 97–123.

Robinson, J. D. (2013). Epistemics, action formation, and other-initiation of repair: The case of partial questioning repeats. In M. Hayashi, G. Raymond, & J. Sidnell (Eds.), *Conversational Repair and Human Understanding* (pp. 261–292). Cambridge University Press.

Robinson, J. D. (2016). *Accountability in social interaction*. New York, NY: Oxford University Press.

Robinson, J. D. (2020). Revisiting Preference Organization in Context: A Qualitative and Quantitative Examination of Responses to Information Seeking. *Research on Language and Social Interaction*, *53*(2), 197–222.

Robinson, J. D., & Bolden, G. B. (2010). Preference organization of sequence-initiating actions: The case of explicit account solicitations. *Discourse Studies*, 12(4), 501–533. https://doi.org/10.1177/1461445610371051

Robinson, J., & Kevoe-Feldman, H. (2010). Using Full Repeats to Initiate Repair on Others' Questions. *Research on Language & Social Interaction*, *43*(3), 232–259.

Romaniuk, T. (2013). Pursuing Answers to Questions in Broadcast Journalism. *Research on Language and Social Interaction*, *46*(2), 144–164.

Rauniomaa, M. (2008). *Recovery through repetition: returning to prior talk and taking a stance in American-English and Finnish conversations*. University of Oulu.

Ross, E. A. (1910). *Social control: A survey of the foundations of order*. New York: The Macmillan Company.

Rossano, F. (2006). *When the eyes meet: Using gaze to mobilize response*. Paper presented at the International Conference on Conversation Analysis, Helsinki, Finland.

Rossano, F. (2012). *Gaze behavior in face - to – face interaction*. Unpublished PhD Dissertation, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.

Rossano, F. (2013). Gaze in conversation. In J. Sidnell & T. Stivers (Eds). *The Handbook of Conversation Analysis* (pp. 308–329). Oxford: Wiley-Blackwell.

Rossano, F., Brown, P., & Levinson, S. C. (2009). Gaze, questioning, and culture. In *Conversation Analysis: Comparative Perspectives* (pp. 187–249). Cambridge University Press.

Rossi, G. (2012). Bilateral and Unilateral Requests: The Use of Imperatives and Mi X? Interrogatives in Italian. *Discourse Processes*, *49*(5), 426–458.

Rossi, G. (2017). Secondary and deviant uses of the imperative for requesting in Italian. In M. Sorjonen, L. Raevaara, & E. Couper-Kuhlen (Eds). *Imperative Turns at Talk The design of directives in action* (pp. 103–137). John Benjamins Publishing Company.

Rossi, G. (2018). Composite Social Actions: The Case of Factual Declaratives in Everyday Interaction. *Research on Language and Social Interaction*, *51*(4), 379–397.

Rossi, G., & Zinken, J. (2016). Grammar and social agency: The pragmatics of impersonal deontic statements. *Language*, *92*(4), e296–e325.

Ruusuvuori, J. (2001). Looking means listening: coordinating displays of engagement in doctor±patient interaction. *Social Science & Medicine*, *52*, 1093–1108.

Sacks, H. (1987). On the preferences for agreement and contiguity in sequences in conversation. In G. Button, J. R. E. Lee (Eds). *Talk and Social Organisation*. (pp. 54-69). Clevedon, Multilingual Matters.

Sacks, H. (1992). *Lectures on conversation* (Vols. 1 & 2). (G. Jefferson, Ed.). Cambridge, England: Blackwell.

Sacks, H. & Schegloff, E. A. (1979), Two preferences in the organization of reference to persons in conversation and their interaction. In G. Psathas (Ed). *Everyday Language: Studies in Ethnomethodology*, New York, Irvington, pp. 15–21.

Sacks, H., & Schegloff, E. A. (2002). Home position. *Gesture*, *2*(2), 133–146.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, *50*(4), 696–735.

Sadock, J. M., & Zwicky, A. M. (1985). Speech act distinctions in syntax. In T. Shopen (Ed.), Language typology and syntactic description (pp. 155–196). Cambridge, England/New York, NY: Cambridge University Press.

Schegloff, E. (1984). On some gestures' relation to speech. In. J. M. Atkinson & J. Heritage (Eds.), *Structure of social action: Studies in conversational analysis* (266-296). Cambridge University Press.

Schegloff, E. A. (1988). On an Actual Virtual Servo-Mechanism for Guessing Bad News: A Single Case Conjecture. Social Problems, 35(4), 442–457.

Schegloff, E. A. (1989). Harvey Sacks-Lectures 1964-1965: An Introduction/Memoir, *Human Studies*, 12(3/4), 185-209.

Schegloff, E. A. (1992), Introduction. In G. Jefferson (Ed). *Harvey Sacks: Lectures on Conversation.* Oxford, Basil Blackwell, vol. 2, pp. ix–lii.

Schegloff, E. A. (1996a). Turn organization: one intersection of grammar and interaction. In *Interaction and Grammar* (pp. 52–133). Cambridge University Press.

Schegloff, E. A. (1996b), Some practices for referring to persons in talk-in-interaction: a partial sketch of a systematics. In B. A. Fox (Ed). *Studies in Anaphora*, (pp. 437-485). Amsterdam, John Benjamins Publishing Company.

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

Schegloff, E. A. (2001). Getting serious: Joke -> serious "no." *Journal of Pragmatics*, *33*(12), 1947-1955.

Schegloff, E. A. (2005). On complainability. *Social Problems*, *524*(4), 449–476.

Schegloff, E. A. (2007). *Sequence organization in interaction: A primer in conversation analysis*. Cambridge, England: Cambridge University Press.

Schegloff, E. A. (2010). Commentary on Stivers and Rossano: "Mobilizing Response." *Research on Language and Social Interaction*, *43*(1), 38–48.

Schegloff, E. A. (2013). Ten operations in self-initiated, same-turn repair. In *Conversational Repair and Human Understanding* (pp. 41–70). Cambridge University Press.

Schegloff, E. A. (2016). Increments. In J. D. Robinson (Ed). *Accountability in Social Interaction* (pp. 239–263). Oxford University Press.

Schegloff, E. A., & Sacks, H. (1973). Opening Up Closings. *Semiotica*, *8*(4), 289–327.

Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The Preference for Self-Correction in the Organization of Repair in Conversation. *Language*, 53(2), 361–382.

Scott, S. (2002). Linguistic feature variation within disagreements: An empirical investigation. *Text - Interdisciplinary Journal for the Study of Discourse*, *22*(2). https://doi.org/10.1515/text.2002.011

Searle, J. R. (1969). Speech acts: An essay in the philosophy of language. Cambridge, England/New York, NY: Cambridge University Press.

Searle, J. (1975). Indirect Speech Acts. In P. Cole, & J. Morgan (Eds.), *Syntax and Semantics 3: Speech Acts* (pp. 59-82). New York: Academic Press.

Selting, M. (1996), Prosody as an activity-type distinctive cue in conversation: the case of so-called "astonished" questions in repair initiation, In E. Couper-Kuhlen, M. Selting, (Eds.), *Prosody in Conversation: Interactional Studies,* pp. 231–270. Cambridge, Cambridge University Press.

Selting, M., Auer, P., Barth-Weingarten, D., Bergmann, J., Bergmann, P., Birkner, K., et al. ... . (2011). A system for transcribing talk-in-interaction: GAT 2 translated and adapted for English by Elizabeth Couper-Kuhlen and Dagmar Barth-Weingarten. *Gesprächsforschung*, *12*, 1–51.

Shreeya, A. (2018). The Self as an Active Agent: Understanding Goffman's Theory of Resistance in Total Institutions through Life-histories. *Sociological Bulletin*, *67*(2), 173–187.

Sidnell, J. (2010). *Conversation Analysis: An Introduction*. Oxford, Wiley-Blackwell.

Sidnell, J. (2017). Action in interaction is conduct under a description. *Language In Society*, *46*(3), 313–337.

Sikveland, R. O. (2019). Failed summons: Phonetic features of persistence and intensification in crisis negotiation. *Journal of Pragmatics*, *150*, 167–179.

Sorjonen, M. & Peräkylä, A. (2012). Introduction, In A. Peräkylä & M. Sorjonen, (Eds.). *Emotion in interaction* Oxford, Oxford University Press.

Sorjonen, M., Raevaara, L., & Couper-Kuhlen, E. (Eds) (2017). *Imperative Turns at Talk The design of directives in action*, (pp. 3-15). John Benjamins Publishing Company.

Stevanovic, M. (2012). Prosodic salience and the emergence of new decisions: On approving responses to proposals in Finnish workplace interaction. *Journal of Pragmatics*, *44*(6), 843–862.

Stevanovic, M. (2017). Managing compliance in violin instruction: The case of the Finnish clitic particles *-pA* and *-pAs* in imperatives and hortatives. In M. Sorjonen, L. Raevaara, & E. Couper-Kuhlen (Eds). *Imperative Turns at Talk The design of directives in action* (pp. 357–380). John Benjamins Publishing Company.

Stevanovic, M., & Peräkylä, A. (2012). Deontic Authority in Interaction: The Right to Announce, Propose, and Decide. *Research on Language and Social Interaction*, *45*(3), 297–321.

Stevens, K. N. (1998). Acoustic Phonetics. Cambridge, MA: MIT Press

Stivers, T. (2005). Modified repeats: One method for asserting primary rights from second position. *Research on Language and Social Interaction*, *38*(2), 131–158.

Stivers, T. (2008). Stance, alignment and affiliation during storytelling: when nodding is a token of affiliation. *Research on Language and Social Interaction*, 41(1), 31–57.

Stivers, T., Mondada, L., & Steensig, J. (2011). Knowledge, morality and affiliation in social interaction. In T. Stivers, L. Mondada, & J. Steensig (Eds.), *The Morality of Knowledge in Conversation,* (pp. 3-24). Cambridge: Cambridge University Press.

Stivers, T., & Robinson, J. D. (2006). A preference for progressivity in interaction. *Language in Society*, *35*(3), 367–392.

Stivers, T., & Rossano, F. (2010). Mobilizing Response. *Research on Language and Social Interaction*, *43*(1), 3–31.

Stivers, T., & Sidnell, J. (Eds.). (2013). *The handbook of conversation analysis.* Wiley-Blackwell.

Streeck, J. (2009a). Forward-gesturing. *Discourse Processes*, 46, 161-179.

Streeck, J. (2009b). *Gesturecraft: the manu-facture of meaning*. Amsterdam; Philadelphia: John Benjamins Pub. Co.

Streeck, J. (2020). Self-Touch as Sociality. *Social Interaction. Video-Based Studies of Human Sociality*, *3*(2). https://doi.org/10.7146/si.v3i2.120854

Sutton-Smith, B. (2001). *The ambiguity of play.* Cambridge, MA: Harvard University Press.

Szczepek Reed, B. (2004). Turn final intonation in English. In E. Couper-Kuhlen & C. E. Ford (Eds.), *Sound patterns in interaction* (pp. 97–117). Amsterdam: John Benjamins.

Szczepek Reed, B. (2012), Prosody, syntax and action formation: intonation phrases as 'action components', In P. Bergmann, J. Brenning, M. Pfeiffer, & E. Reber (Eds.)*, Prosody and Embodiment in Interactional Grammar*, (pp. 142–169). Berlin: De Gruyter.

Szymanski. M. H. (1999). Re-Engaging and Dis-Engaging Talk in Activity. *Language In Society*, *28*(1), 1–23.

Taleghani-Nikazm, C., Drake, V., Golato, A., & Betz, E. (2020), Mobilizing for the next relevant action: Managing progressivity in card game interactions, In C. Taleghani-Nikazm, E. Betz, & P. Golato (Eds). *Mobilizing Others: Grammar and lexis within larger activities* (pp. 47-81). Amsterdam / Philadelphia: John Benjamins Publishing Company.

't Hart, J., Collier, R. & Cohen, A. (1990). A Perceptual Study of Intonation: An Experimental-Phonetic Approach to Speech Melody. Cambridge: Cambridge University Press.

Thompson, S. A., Fox, B. A., & Couper-Kuhlen, E. (2015). *Grammar in Everyday Talk: Building Responsive Actions*. Cambridge University Press.

Tuncer, S., & Haddington, P. (2020). Object transfers: An embodied resource to progress joint activities and build relative agency. *Language In Society*, *49*(1), 61–87.

Vatanen, A. (2018). Responding in Early Overlap: Recognitional Onsets in Assertion Sequences. *Research on Language and Social Interaction*, *51*(2), 107–126.

Walker, G. (2012). Coordination and interpretation of vocal and visible resources: "trail-off" conjunctions. *Language and Speech*, *55*(Pt 1), 141–163.

Walker, T. (2014). Form≠ function: The independence of prosody and action. *Research on Language and Social Interaction, 47(1), 1-16.*

Whitehead, K. A. (2015). Everyday Antiracism in Action: Preference Organization in Responses to Racism. *Journal of Language and Social Psychology*, *34*(4), 374–389.

Whitehead, K., Bowman, B., & Raymond, G. (2018). "Risk factors" in action: the situated constitution of "risk" in violent interactions. *Psychology of Violence*, *8*(3), 329–338.

Wootton, A. J. (Ed.). (1997). Requesting at 12–24 months: An overview. *In Interaction and the Development of Mind* (pp. 27–55). Cambridge University Press.

Wu, R. (2004). *Stance in Talk*. John Benjamins Publishing Company.

Yu, G., Wu, Y., & Drew, P. (2019). Couples bickering: Disaffiliation and discord in Chinese conversation. *Discourse Studies*, *21*(4), 458–480.

Zellers, M., & Ogden, R. (2014). Exploring Interactional Features with Prosodic Patterns. *Language and Speech*, *57*(3). https://doi.org/10.1177/0023830913504568

Zinken, J. & Deppermann, A. (2017). A cline of visible commitment in the situated design of imperative turns: Evidence from German and Polish. In M. Sorjonen, L. Raevaara, & E. Couper-Kuhlen (Eds). *Imperative Turns at Talk The design of directives in action* (pp. 27-63). John Benjamins Publishing Company.

Zinken, J., & Ogiermann, E. (2013). Responsibility and Action: Invariants and Diversity in Requests for Objects in British English and Polish Interaction. *Research on Language and Social Interaction*, *46*(3), 256–276.