

## Sequence Facilitation: Grandparents Engineering Parent–Child Interactions in Video Calls

Yumei Gan<sup>a</sup>, Christian Greiffenhagen<sup>b</sup>, Kobin H. Kendrick<sup>c</sup>

<sup>a</sup>*School of Media and Communication, Shanghai Jiao Tong University, China*

<sup>b</sup>*Department of Applied Social Sciences, The Hong Kong Polytechnic University, Hong Kong SAR, China*

<sup>c</sup>*Department of Language and Linguistic Science, University of York, United Kingdom*

### ORCID:

Yumei Gan: <https://orcid.org/0000-0001-8121-3988>

Christian Greiffenhagen: <https://orcid.org/0000-0002-7544-9246>

Kobin H. Kendrick: <https://orcid.org/0000-0002-6656-1439>

### \*CONTACT:

Yumei Gan, [yumei.gan@sjtu.edu.cn](mailto:yumei.gan@sjtu.edu.cn), School of Media and Communication, 800 Dongchuan Road, Shanghai Jiao Tong University, China. 200240.

## *Abstract*

Completing a sequence of actions is a basic problem of social organisation for participants. When a first pair-part is addressed to a not yet fully competent member, such as a young child, a third party can facilitate the completion of the sequence through diverse linguistic, embodied, and material practices. In this article, we examine such *sequence facilitation* in a perspicuous setting, namely grandparent-mediated video calls between migrant parents and their left-behind children in China. The analysis showed that the practices of sequence facilitation can have a retrospective or prospective orientation and involve not only linguistic practices, such as repeating the parent's first pair-part or formulating its action, but also embodied and material practices, such as positioning the camera or physically animating the child's body. The results shed light on the organisation of adjacency pairs in adult-child interactions and the embodied and material circumstances of their production in video-mediated communication. The data were in the Chinese dialects of Sichuan and Guizhou.

The most basic and ubiquitous type of sequence is the adjacency pair, in which a first pair-part (FPP) makes conditionally relevant a corresponding second pair-part (SPP). An important feature of these sequences lies in the fact that they are expandable and can expand recursively beyond the minimal unit, which can occur before, between, or after the base adjacency pair. Such expansions involve additional participation through additional turns over and above the adjacency pair-based sequence (Levinson, 2013; Schegloff, 2007), resulting in complex interactional structures (Schegloff, 1990).

## The Phenomenon: Sequence Facilitation

### Extract 1

3

youngest aunt, you say  
 grf     \$looks at the child→  
 chi                     →\*move gaze up to look at phone→  
 fig     **#fig.1b**             **#fig.1c**  
 04             (0.7)  
 05     CHI:     °小             嬢嬢°  
                  °xiao             niangniang°  
                  °youngest aunt°  
                  °youngest aunt°



**Figure 1: Grandfather looks at the child and directs the child to respond to the mother**

The grandfather's statement ("youngest aunt, you say," line 03) expands the adjacency pair initiated by the mother's question at line 1 and is completed by the child's answer at line 05. Occurring between the base pair parts of the sequence, the grandfather's action constitutes a form of expansion (Goodwin & Goodwin, 1990; Schegloff, 2007), albeit one that involves the participation of a third party and only a single turn rather than an adjacency pair. The minimal expansion is designed to elicit the conditionally relevant SPP from the addressed recipient of the base FPP and in this sense facilitates the completion of the sequence. We term the grandfather's action *sequence facilitation* and note several important features of it.

The grandfather's facilitation is responsive to and occasioned by the locally situated contingences of the interaction. Before the base FPP reaches possible completion, the child averts her gaze (line 01), which violates the normative expectation of the mutual gaze in question–response sequences (Kendon, 1967; Kendrick & Holler, 2017). At the question's completion, when the addressed recipient should respond (Sacks et al., 1974), the child does not answer, and a gap of 0.3 seconds emerges (line 02). This sequential environment, after the addressed recipient fails to respond (Lerner, 2019; Stivers & Robinson, 2006), leads to sequence facilitation.

The grandfather's verbal turn includes two linguistic components. The first provides the conditionally relevant answer to the mother's question ("youngest aunt"), while the second employs an imperative to instruct the child to respond ("you say," line 03). The turn is recognisably produced *for* the child, not on the child's behalf. Thus, it does not complete the sequence. The turn solves several practical problems for the child, in that it reallocates the next turn to her, instructs her to speak, and gives her the words to use. The work required for her to respond is thus lessened. In response, the child directs her gaze to the phone, re-establishing the participation framework; after a further delay (line 04), she repeats the answer provided by her grandfather, closing the sequence.

Furthermore, the spatial, material, and embodied arrangement of the participants becomes critical as the sequence unfolds. The grandfather holds the child on his lap and positions the smartphone in front of her face. The way the grandfather positions the technological artifact (the smartphone camera) matter as it enables a participation framework (Goodwin &

Goodwin, 2004) between the mother and child in which the adjacency pair occurs. From such a perspective, sequence facilitation involves not only linguistic practices, such as those that comprise the grandfather's forms of talk, but also multimodal practices, including arranging the child's body and technological objects that establish the necessary conditions for a normatively organised sequence of action.

## Background

Our development of the notion of sequence facilitation builds on the following three areas of inquiry: (i) sequence facilitation; (ii) embodied sequences with children; and (iii) sequences in video-mediated communication.

### *Sequence Facilitation*

Although researchers have not used the term “sequence facilitation”, the phenomenon has been observed sporadically in the literature. Lerner (2019), for example, analysed a sequence in which a caregiver intervenes to elicit a conditionally relevant response from a child.

#### **Extract 2**

[VYC: 990413A1-FP1] (Lerner, 2019, p. 393)

```
1 Dad:      BYE BYE An:(.)tonio.  
2           (.)  
3 CG1: → BYE::Mitch,  
4           (.)  
5 Dad:      [(Bye bye)  
6 CG1: → [Say bye to your dad?
```

The organisation of the sequence here is similar to our phenomenon in Extract 1. After the parent addresses an FPP to the child, the caregiver employs practices designed to elicit a corresponding SPP. Lerner (2019) presented this case to exemplify an unaddressed recipient's *intervention* in a sequence and to make the point that young children belong to a class of participants “for whom the question of their ability to respond has an ongoing relevance for their coparticipants” (p. 393). According to Lerner (2019), the caregiver first models the relevant response for the child (line 3) and then instructs the child on how to respond (line 6). These actions constitute sequence facilitation as we define it, namely, an unaddressed recipient of an FPP produces a turn or action that expands the sequence and has the opportunity to elicit a relevant SPP from the addressed recipient.

Building on Lerner's (2019) analysis, we observed the linguistic practices that participants employ to accomplish a sequence facilitation. “BYE::Mitch” (line 3) includes an address term appropriate for the adult caregiver, but not for the child, who might address the father as “dad” or “daddy.” The selection of the address term thus designs the response as independent of the child (see again “youngest aunt” in Extract 1, which was designed from a child's perspective). Furthermore, the design of the instruction (line 6) concisely addresses three practical problems for the child: the imperative format directs the child to speak next, the object “bye” provides the word to use, and the phrase “to your dad” indicates the appropriate addressee. Practices such as these are also common in our dataset.

However, while this case comes from a video recording of a co-present interaction, the transcript and analysis do not address relevant features of participants' embodied conduct, such as their spatial arrangement or the direction of their gaze. We thus do not know whether the caregiver gazes at the father or at the child as they produce the actions in question, and we are left to wonder whether the child is disengaged from the interaction. For us, these

embodied details matter, as they enable a holistic account of sequence facilitation as a multimodal phenomenon—one that involves not only turns at talk but also their situated spatial, material, and embodied circumstances.

In the existing literature, these practices of sequence facilitation have been described under a variety of names. Schiffelin and Ochs (1986) observed that caregivers in different cultures use practices such as “prompting” and “providing explicit instruction” to form interactional routines that benefit young children’s language acquisition and development. Goodwin and Goodwin (1990) identified “piggybacking,” in which an unaddressed recipient of an FPP reiterates the FPP before the addressed recipient responds. Stivers (2001) showed that in pediatric medical consultations, parents “prompt” their children to answer doctors’ questions if they display difficulty (p. 272). In a study of bilingual interactions, Bolden (2012) examined practices of repair that unaddressed recipients employ to render FPPs intelligible to addressed recipients, thereby “brokering” the sequence. Furthermore, Kendrick and colleagues (2020) documented a common practice in Tzeltal, a Mayan language of Mexico, wherein unaddressed recipients “pursue” a response, especially from young children (*jak’a laj*, “answer, they say”). In each of these cases, an unaddressed recipient of an FPP employs a practice designed to elicit a fitted SPP from the addressed recipient and thus acts to facilitate the completion of the sequence.

Moreover, sequence facilitation has been repeatedly observed when there are not fully competent (or less than fully competent) participants involved, such as in interactions with young children (Lerner, 2019; Stivers, 2001), patients with cognitive impairments (Doehring, 2018), or migrants who have less linguistic competence (Bolden, 2012). Stivers and Robinson (2006) revealed that situations in which the addressed recipient of an FPP fails to respond promptly, claims an inability to respond, or displays difficulty in responding, often lead to others’ participation. However, in interactions involving children, Lerner (2019) observed that interventions by third parties “may not result from demonstrated inability on that occasion alone but can be oriented to by coparticipants as an omnirelevant possibility” (p. 393). Nonetheless, Liu (2022) found that, adults tend to prioritize children’s participation in conversations over conversational progressivity.

### ***Embodied Sequences with Children***

The involvement of young children in interactions can increase the complexity and challenges of interaction and thereby lead to distinctive interactional structures. For example, Schegloff, Jefferson, and Sacks (1977, pp. 380-381) mentioned that adult–child interactions are an apparent exception to the preference for self-repair, as children are not yet fully competent members. Conversation analytic studies on sequences involving children have compellingly shown that embodied actions constitute recurrent resources in adult–child interactions to control and monitor the children’s responding actions (e.g., Cekaite, 2010, 2015; Goodwin & Cekaite, 2013, 2018). Cekaite (2010) demonstrated adult–child interaction as both a dialogic and an embodied act, for which parents deploy embodied directives, coordinated with verbal, physical and spatial practices to shepherd the child into compliance with the requested actions.

Among many other embodied resources, touch has been seen as a critical resource for getting children to complete a sequence or accomplish an activity in face-to-face interactions (Goodwin & Cekaite, 2018). Touch is not only used to guide the child’s response, but also employed to enforce the child’s attention, especially when the child is engaged in competing activities. For instance, Cekaite (2010) demonstrated that shepherding moves usually begin with an embodied action, which aims to terminate and bound off a child’s competing activity

by re-orientating the child's body and changing their attentional focus in the immediate situational context. In practice, an adult's bodily interventions can of course be accepted or rejected by the child. A child recipient may resist the parental embodied directives by remaining immobile or physically resisting (Goodwin & Cekaite, 2013). Consequently, children's non-compliance may lead to adults' upgrading of their actions.

### ***Sequences in Video-Mediated Communication***

Interactions with young children in video-mediated environments can pose even more challenges as children may wander away or fail to pay attention to the people on screen (Gan, 2021). To explore the complexity of such interactions, our study builds on previous work on the organisation of sequence in video-mediated communication.

Scholars have shown that constituting a sequence in video calls is in fact a practical achievement that implements spontaneously the technical conditions, arrangement of bodies, establishment of mutual gaze, and dialogic speech (Licoppe & Morel, 2012; Mondada, 2010). In video calls, the initiation of a sequence appears to be different compared with telephone calls. In telephone calls, it is often the call taker who speaks first (Schegloff, 1986). However, studies have found that it is not clear who speaks first during an opening of a video call (Licoppe & Morel, 2012) because people may need to deal with a weak video connection and conduct "technology check" (Relieu, 2007). The faces or the images that one shows on screen are also available resources for the recipient to monitor for subsequent actions.

Techno-material arrangements such as camera layout matter significantly for constituting a smooth sequence in video-mediated interactions. Sunakawa (2012) described the video-mediated configuration as a "face-to-machine space" (drawn from Keating, 2006, p. 335) wherein remote speakers have to establish their encounters under the proxy of technology. To achieve this visual space, videographic practices such as camera positioning are carefully coordinated. For example, when adults aim to display a baby, they manoeuvre the camera and discuss the best way to position it (Sunakawa, 2012, p. 61). Consequently, cameras are not simply perceived as technical artifacts necessary for video-mediated interaction, but as meaningful techno-material arrangements that "lead to the opportunity for speakers to respond and to interpret the foregrounded scene" (Sunakawa, 2012, p. 63).

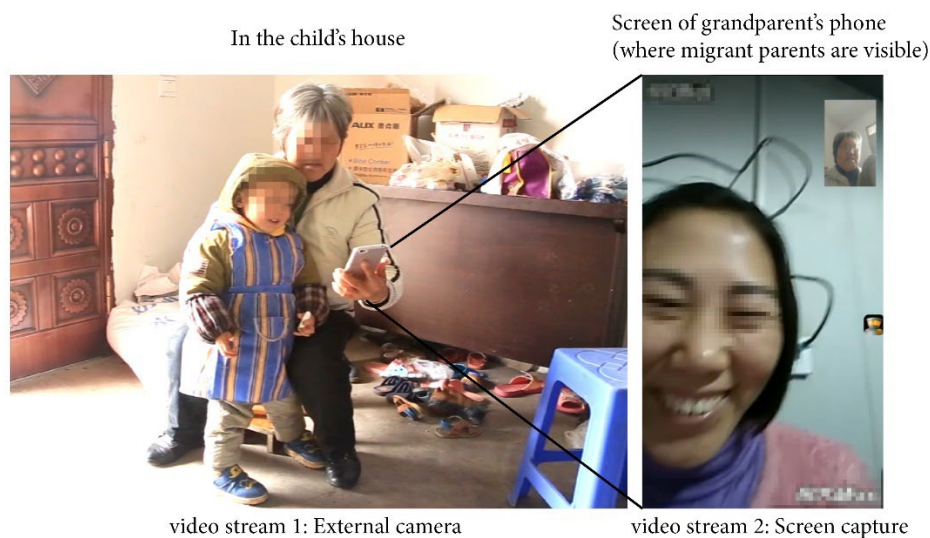
Drawing on these studies, we introduce the concept of sequence facilitation to amplify previous research by identifying and describing the full array of linguistic, embodied, and material practices by which sequence facilitation occurs in a complex environment in which young children are interacting with their parents in a video-mediated setting. We argue that this arrangement constitutes a "perspicuous setting" (Garfinkel, 2002, pp. 181–182) for studying sequence facilitation, given the involvement of not-yet-fully-competent participants. We also demonstrate that the mediated nature of the interactions adds new layers to the discussion of sequence facilitation by bringing the focus of camera and material arrangement into the interactions.

### **Data and Method**

This study adopted the methodology of conversation analysis (Sacks, 1992). Data were drawn from 75 video recorded, naturally occurring video calls between migrant workers and their left-behind children in the Sichuan and Guizhou provinces of China. The video calls lasted between 15 and 65 minutes, and averaged 22 minutes per call. In the selected families, both parents were migrant workers. The children were left behind by their parents and raised in rural areas by their grandparents. The family members were familiar with video-mediated communication technologies and used the Chinese instant messaging app WeChat on their

smartphones to conduct video calls. The children in this study were aged from eight months to three years old (some of them had elder siblings in the video recordings), and thus were always accompanied by caregivers on calls. No instructions were given to the participants regarding their call activities. Data collection received ethical approval from the Research Ethics Committee of Chinese University of Hong Kong, and participants (including the children's legal guardians) provided informed consent for the video recordings and for using the data (including the talk and images) in this article. All names appeared have been changed to pseudonyms.

The data consisted of a screen capture of the participants' smartphones and a traditional camera recording of the interaction in front of the smartphones. This method made it possible to capture all three parties in the conversation, namely the remote parents, caregivers (grandparents), and the left-behind children. Both video streams were subsequently synchronised (Figure 2). The data were in the Chinese dialects of Sichuan and Guizhou.



**Figure 2: Video-recording method**

We assembled a collection of instances ( $N = 230$ ) in which sequence facilitation occurred—that is, when a remote parent selected the child to respond, but the next turn was taken by the grandparent to facilitate the completion of the parent-child sequence. These instances were transcribed using the multimodal transcription convention created by Mondada (2014). To represent the meanings as precisely as possible, our transcription of spoken Chinese Sichuan and Guizhou dialects involved four steps in practice. First, we transcribed the Chinese characters to present the original language orthographically. Second, we created Romanised transcripts by adding Pinyin to show the spelling of each word. Third, we added a literal English translation and grammatical gloss for each word and discourse particle. Fourth, we included an English translation to communicate the meaning of the utterance to non-native readers.

## Analysis

Our analysis of sequence facilitation proceeded with a focus on its multimodal organisation in the unfolding environment. Building on Schegloff's (2007) analysis of the two orientations of insert sequences (i.e., post-first and pre-second), we divided our analysis into the following two parts: retrospective practices oriented toward FPPs and prospective practices oriented toward SPPs. We document a wide range of linguistic, embodied, and material practices



employed by the grandparents to elicit conditionally relevant responses from the children and thereby facilitate the completion of the sequence.

### Practices Oriented Toward FPPs

After a remote parent's FPP comes to a possible conclusion and a response from the child becomes conditionally relevant, the grandparent might intervene in the sequence. Such facilitation practices can have a retrospective orientation, operating on the parent's FPP by repeating it or formulating its action, and can also incorporate various embodied actions such as moving the smartphone camera and manipulating the child's body to establish a necessary participation framework for a successful parent-child sequence. These facilitations often occur when a child is not attentive or when a parent produces a complicated FPP that posed difficulties for the child.

#### *Repeating the FPP*

Perhaps the simplest way in which grandparents initiate sequence facilitation is by repeating the FPP produced by the remote parents. In Extract 3, a grandfather (GRF) and two children, Kai (the child in the middle, aged 2.5 years) and Xin (the child on the right, aged 5 years) are chatting with their mother (MOM). Kai and Xin are brothers. The mother initiates a question in line 01, addressing Kai:

#### **Extract 3**

(LET29\_00.54)

- 01 MOM: \*#王 凯, 你 在 屋头 耍起 舒\*服 不\*#?  
           Wang Kai,ni zai wutou shuaqi shufu bu?  
           Wang Kai,you be home play comfortable not?  
           Wang Kai, are you enjoying playing at home?  
       kai \*gazes at sweet and unwraps it→ \*moves gaze to Xin\*#  
       fig **#fig.3a** **#fig.3b**
- 02 (0.3)+#(0.3)  
       grf +moves gaze from the phone to the child→  
       fig **#fig.3c**
- 03 >> GRF: 问 哈儿 你 在 屋头 耍起 舒服 不?  
           wen haer ni zai wutou shuaqi shufu bu?  
           ask PRT you be home play comfortable not?  
           (she) is asking that "are you enjoying playing at home"?
- 04 (0.2)
- 05 KAI: 舒服  
           shufu  
           comfortable  
           enjoying



**Figure 3: The child looks at the sweet and then moves his gaze toward the other child. After the mother's question, the grandfather gazes toward the child.**

Before the mother produces her question, she uses the child's full name, "Wang Kai", so that the question "are you enjoying playing at home" (line 01) is addressed to him. While the mother summons the child, the child is involved in a competing activity. He is unwrapping a sweet in his hand (Figure 3a). Moreover, after he has unwrapped the sweet, he still does not orient himself to his mother on the screen. Instead, he turns his gaze to his brother, Xin (Figure 3b). After a 0.3 second silence, the grandfather moves his gaze up to look at Kai (from Figure 3b to 3c). The grandfather then repeats the mother's FPP, "(she) is asking that 'are you enjoying playing at home?'" (Line 03). The grandfather produces his repetition as an indirect reported speech by adding a turn-initial *wen haer* "[she] is asking that". As Holt and Clift (2006, p. 51) argued, indirect reported speech is commonly preceded by "that" and can convey the speakers' point of view as not the original speaker. By so doing, the grandfather illustrates that he is reporting an action from others.

The child's inattentiveness and involvements in competing activities are crucial contingencies leading to the non-response. The non-response is followed by the grandfather's facilitation, which provides by reporting and repeating the mother's FPP as a way of pursuing a response (Pomerantz, 1984) *from the child for the mother*. The simultaneous embodied practices of the grandfather, including the shift of his head and gaze to look at the child, show that facilitation is not only done by repeating the FPP but also through concurrent embodied orientation and embodied pursuit (Rossano, 2006).

### ***Formulating the FPP's Action***

The grandparents may also formulate the parent's FPP to manage the contingency of non-response by the child or in response to a parent's complaint about it. By formulating, they interpret the actions implemented in the parent's FPP and thereby provide more resources for the child to produce an answer. As Garfinkel and Sacks (1970) noted, to formulate is "to explain it, or characterise it, or explicate, or translate, or summarise, or furnish the gist of it" (p. 350). In Extract 4, the father (DAD) conducts a video call with two children: Jie (the boy, full name Li Xiaojie, aged 4 years) and Wen (the girl, aged 15 months). The father calls Jie's full name three times (lines 04, 06, and 08). Although the child gazes at the screen attentively, he does not respond. Then the grandmother formulates the father's FPP for the child.

#### **Extract 4**

(LBC52\_7.51)

01     GRA:    来\*            >喊 他< 咬    一    口  
              lai            >han ta< yao yi    kou  
              come        >ask he< bite one   CL  
              come on, >ask him< to have a bite  
              wen        \*Wen and Jie are eating ice-cream→

02 DAD: 看 你 才 \*好吃  
kan ni cai haochi  
look you so greedy  
look you are so greedy  
wen →\*feeds ice-cream to phone/dad→

03 (1.0)

04 DAD: 李 小杰@#  
Li Xiaojie  
Li Xiaojie [full name of the boy]  
jie @gazes at his father who is on screen→  
**#fig.4a**

05 (1.3)

06 DAD: 李 小杰 啊  
Li Xiaojie ah  
Li Xiaojie PRT  
Li Xiaojie ah/huh

07 (1.0)@+(1.0)  
jie →@moves gaze away from the phone→

08 DAD: 李::小::杰  
Li::Xiao::jie  
Li::Xiao::jie  
Li::Xiao::jie

09 (0.9)

10 DAD: 你 不得 给 \$#我 两个 谈 的? \$#是 不\$#?  
ni bude gei wo liangge tan de? shi bu?  
you not with me two talk PRT? be not?  
you are not talking to me? aren't you?  
gra \$gazes at the phone→ \$moves gaze to Jie→  
fig **#fig.4b** **#fig.4c**

11 >> GRA: 李 小杰, 你 爸爸 在 喊 你  
Li Xiaojie, ni baba zai han ni  
Li Xiaojie, you dad be call you  
Li Xiaojie, your dad is calling you

12 JIE: 爸@#爸  
BABA  
DAD  
DAD  
→@moves gaze back to screen and leans his body to screen  
**#fig.4d**



**Figure 4: Jie looks at the phone when the father summons him, but does not respond to the father.**

To fully appreciate the importance of sequence facilitation, it is helpful first to look at the complex unfolding interactions in this piece of data, where two children and two grandparents are chatting with the remote father. While everyone is involved in the conversation, it is possible that one child becomes a focused participant at a specific moment. In line 01, the grandmother talks to Wen and instructs her to feed ice cream to her father. The grandmother says, “Come on, >ask him< to have a bite”. The father comments to the girl “You are so greedy” (line 02). Wen complies with her grandmother’s instruction, feeding the ice cream to the smartphone (line 02). The father does not take the pretend feeding further. In line 04, the father switches his participation framework and now summons the boy by saying his full name, Li Xiaojie. The boy seems to gaze at the screen attentively (Figure 4a), but he does not respond to his father. After a silence (line 05), the father pursues the child’s response by summoning him again (line 06). This time, the father upgrades the summons slightly by adding a particle, “ah/huh” (*a*), to elicit a response (Schegloff, 1997). Although the child’s head is displayed on the screen and his gaze follows his father, he still does not respond to the summons. After these two summoning attempts, Jie moves his gaze away from the phone. A relatively longer gap occurs in line 07. The father then upgrades his summons for a third time by stretching out the child’s name, “Li::Xiao::jie” (line 08). The upgraded summons once again fails to draw a response from the child. The father treats the non-response as problematic, which leads to two questions in line 10: “You are not talking to me? Aren’t you?” He first formulates the child’s non-response as a negative declarative question (“You are not talking to me?”), then produces a negative interrogative (“aren’t you”) that evaluates the recipient’s conduct in problematic terms (Heritage, 2002, p. 1429). As Schegloff (1988, p. 124) argued, speakers can formulate a failure and a negative consequence to issue a complaint. Here, the father’s treatment of the child’s non-response as negative can be heard as a way of complaining about the lack of response.

The father’s turn is followed by the grandmother’s facilitation utterance, “Li Xiaojie, your dad is calling you.” Here, the grandmother formulates the father’s various actions as one, “calling.” By doing so, she also minimises the father’s complaint and prompts the child to produce an affect-laden response shouting, “DAD,” with noticeably increased volume and rapid tempo. As shown in Figure 4d, the child opens his mouth wide and visibly smiles. The child’s exclamation not only answers the grandmother’s prompt but also his father’s summons and various actions from prior turns. By formulating the father’s three turns as “dad is calling you,” the grandmother clarifies the actions implemented in the series of summonses and thereby enables the child to respond. The grandmother once again conducts sequence facilitation here after a series of interactional troubles, namely non-response while Jie is looking at his father, non-response while Jie moves his gaze away, and the father’s complaint about Jie. The delay in facilitation shows that the grandmother orients herself toward the child’s right and obligation to respond to the father on his own (Stivers, 2001), leaving open opportunities for the child to react.

### *Positioning the Camera for Mutual Gaze*

While our first two examples (Extracts 3 and 4) centered on the grandparents' linguistic practices for sequence facilitation, our next two examples (Extracts 5 and 6) demonstrate that sequence facilitation in practice involves extensive multimodal activities. In the following, we illustrate that sequence facilitation in this setting is made possible because the facilitator coordinates the technological artifacts, smartphone cameras, children's bodies, and ongoing participation framework in a complex ecology. The on-site material–body arrangement significantly reshapes the occurrence of sequence facilitation.

Extract 5 exemplifies how the grandmother manipulates a smartphone camera to conduct sequence facilitation. After the mother issues an FPP, the grandmother moves the camera to position the two-year-old child's face on screen to establish a mutual gaze for a parent-child sequence.

### Extract 5

Extract 3  
(LBC37 10.37)

- 01      GRA:    °得，     拿+去°  
             °dei,    naqu°  
             °here, take°  
             °here,take it°  
         grm                  +passes blocks→
- 02      MOM:    #王       夕月  
             WANG   Xiyue  
             WANG   Xiyue [CHI's full name]  
**#fig.3a**
- 03                (0.4)
- 04      MOM:    王       夕月    (.)你    想       吃    不+#?(.) 拿去- 吃  
         Wang   Xiyue (.)ni   xiang chi   bu? (.) naqu- chi  
         Wang   Xiyue (.)you want eat   not?(.) take- eat  
         Wang   Xiyue(.)do you want to eat? (.)take to- eat  
    >> gra    +moves phone fr L to R hand→  
     fig    **#fig.3b-3c**
- 05      GRA:    >来+# 么儿<  
             >lai yao er<  
             >come sweetheart<  
             >come on sweetheart<  
         gra                  →+moves the phone toward the child's face→  
     fig                          **#fig.3d**
- 06      GRA:    妈妈+    又       喊    你    吃    么儿  
         Mama   you       han   ni    chi   yaoer  
         Mom    again ask   you   eat   sweetheart  
         Mom is asking you again to eat sweetheart  
         →+holds phone in front of child's face→



**Figure 5: The grandmother moves the phone from the right to left hand, then moves it even closer toward the child's face.**

As shown in previous cases, children are often involved in a competing activity. At the beginning of this extract, the girl is playing with building blocks, and the grandmother passes blocks to her (line 01). In line 02, the mother summons the child by calling out her full name, “Wang Xiyue.” Note that in Figure 5a, when the mother’s FPP occurs, the child is still playing with building blocks. Moreover, the phone is in the grandmother’s right hand and is not facing the child when the FPP is produced. The mother’s summons does not get an answer (line 03). She then repeats the child’s name and adds an invitation of pretend eating, “Wang Xiyue (.) do you want to eat?” At that moment, the grandmother begins to manipulate the phone, moving it from her right hand to her left hand (Figure 5b to 5c). The mother continues talking, inviting the child to “virtually” eat the food (“take to- eat”). The grandmother does not stop her movement of the smartphone; she continues moving her left hand closer to the child (Figure 5d). By so doing, she positions the current speaker (i.e., the mother) and the selected next speaker (i.e., the child) on the screen (Licoppe & Morel, 2012). Once the participation framework has been established, the grandparent uses the practice of repeating the FPP, which was discussed earlier.

The grandmother’s re-arrangement of the camera position attempts to establish a necessary participation framework (Goodwin & Goodwin, 2004) for a parent–child sequence. This creates a direct line of sight between the parent and the child and thus establish the necessary conditions for mutual gaze. In so doing, the facilitator orients herself toward a normative organisation of adjacency pair sequences similar to face-to-face interaction, wherein FPPs are produced with the gaze directed to the addressed recipient and the selected next speaker; the addressed recipient should gaze at the FPP speaker, such that a mutual gaze is achieved (Kendon, 1967; Rossano et al., 2009). Our data also showed other cases of grandparents moving smartphone cameras as a contingency due to a child’s inattentiveness, or when parents were not visible to the child. By dynamically repositioning the smartphone, the grandparents would manage the material affordances of the technology to deal with the troubles of interacting through a video-mediated channel, where the gaze has to be mediated through a smartphone screen and cameras.

### ***Orienting the Child’s Attention to Camera***

However, sometimes putting the child’s face on screen (i.e., making the selected next speaker visible) is treated by the grandmother as insufficient for establishing a participation framework. Gan and colleagues (2020) showed that establishing a talking head configuration in video calls normally involves the following two perspectives: the head of the speaker is visible on screen and the speaker should look at the screen. As seen in Extract 6 (continuing from Extract 5), the grandmother manipulates the child’s head to re-arrange the child’s attentional focus in the course of an ongoing sequence.

### **Extract 6**

(LBC37 continued\_11.01)

- 06 GRA: 妈妈+ 又 喊 你 吃 么儿  
Mama you han ni chi yaoer  
Mom again ask you eat sweetheart  
*Mom is asking you again to eat sweetheart*  
→+holds phone in front of child's face→
- 07 GRA: 快点+# 妈妈 喊 你 吃 (.) [喂  
KUAIDian mama han ni chi(.) [dei  
COME on mom ask you eat [here  
*Come on, mom is asking you to eat(.) [here*
- 08 MOM: [来 吃\*#不 ?  
[lai chi bu?  
[come eat not?  
[come to eat?
- >> gra →+uses hand to move the child's head up  
chi \*gazes up&opens mouth  
fig #fig.6a #fig6b



Figure 6: The grandmother touches the child's head and moves it up

The child continues playing with the blocks as the grandmother moves the phone in front of her face but is still not attentive to the mother's invitation for pretend eating (Figure 6a). The grandmother uses her right hand to touch the child's head and move it up (Figure 6a). As Cekaite (2016) showed, haptic intervention is often employed by adults to solicit and monitor a child's attention, especially in situations in which the child is engaged in multi-activity. The grandmother's touch thus effectively addresses the child's inattentiveness and elicits her participation and cooperation (Cekaite, 2016, p. 37). In the meantime, the mother again issues the invitation for pretend eating (line 08) to solicit the child's response. The child finally looks at the screen and opens her mouth wide (Figure 6b). Opening her mouth not only exhibits her understanding of the mother's invitation, but also accepts the invitation to pretend to eat the food.

Thus far, by looking backward to the parents' FPPs, we have shown that the grandparents deploy a broad palette of linguistic resources (e.g., repeating and reformulating), haptic interventions (e.g., touching the child's head), and camera practices (e.g., moving the cameras) as they orient to the selected recipient (the child) to respond to their parents on their own. In the following section, we continue to explore the phenomenon of sequence facilitation by looking forward to the expected SPPs.<sup>1</sup>

<sup>1</sup> In our collection, we found that most cases (N=133) were related to facilitating SPPs and relatively fewer cases (N=97) were oriented toward the FPP produced by parents. This distribution of cases seems to be meaningful in the sense that the main objective of sequence facilitation was to get a response from the child.



## Practices Oriented Toward SPPs

Prospective practices that facilitate a child's SPPs include formulating the SPP's action, providing the child with an appropriate response, assisting the child's embodied response, and animating the child's body in response, which we consider in turn.

### *Formulating the SPP's Action*

The grandparents sometimes formulate an appropriate responding action. By doing so, they clarify the type of action that should be performed by the child but left it to the child to select and produce a precise response. The grandparents use this strategy to manage contingencies, such as the absence of a response or a child's inattentiveness. In Extract 7, the mother summons the child (aged 2); after a 0.5 second gap, the grandmother says "answer".

#### **Extract 7**

(LET22\_08.50)

- 01 MOM: 小:::~\*#涛::  
Xiao:::~tao::  
Xiao:::~Tao::[child's nickname]  
chi \*gazes down (not at screen)→  
fig #fig.7a
- 02 (0.4)+\*#(0.1)  
chi →\*gazes away→  
fig #fig.7b
- 03 >> GRA: 应 嘛  
ying ma  
answer PRT  
answer MA
- 04 (0.2)+\*#(0.2)  
chi →\*gazes at screen→  
fig #fig.7c
- 05 CHI: 妈妈  
Mama  
mom  
mom



**Figure 7: The child gazes down and away. After the grandmother's facilitation, the child gazes at the screen.**

In line 01, the mother summons the child with his nickname. The child gazes down at the toy in his hand. After a gap, the child is still gazing away, not at the screen. He does not respond to his mother (line 02). In line 03, his grandmother says "answer" using the *ma* particle. She does not provide an exact answer for the child, simply naming the responding action for the child. Furthermore, the grandmother's facilitation utterance includes the action ("answer") and a non-question particle *ma* (嘛). As Kendrick (2010) argued, the use of this Chinese



particle *ma* encodes the epistemic claim that the recipient has the right and obligation to confirm the information in the *ma*-marked turn. With this particle, the grandmother formulates the action as something that the child should know how to perform, thereby challenging the child's non-response. She thus provides information to the child who has apparently failed to perform the next action correctly. After a pause, the child moves his gaze to look at the screen. Then he says "mom" (line 05), which suffices as an answer to the mother's summons.

### *Providing an SPP*

Extract 8 illustrates how the grandparents facilitate SPPs by explicitly providing the child an SPP. Such practices can be employed before any interactional trouble emerges in order to pre-empt contingencies. The grandparents often do so by employing the instructional format, "you say 'X'".

#### **Extract 8**

(LBC26\_00.00)

14 MOM: \*+#你 的 饭:饭 吃 完 了 啦:?  
           ni de fan :fan chi wan le la: ?  
           your PRT rice: rice eat finish PFV PRT:?  
           *have you finish::ed your meal:?*  
 chi \*gazes at phone/mom→  
 ucl +holds the phone→  
 fig **#fig.8a**

15 >> GRA: 你 说 >吃 完 了<  
           ni shuo >chi wan le<  
           you say >eat finish PFV<  
           *you say >finished eating<*

16 UNC: 嘢+#  
           Dei  
           Here  
           *here you are*  
 unc →+passes the phone to grandmother→  
 fig **#fig.8b**

17 CHI: 吃 完 了  
           chi wan le  
           eat finish PFV  
           *finished eating*

18 (1.0)

19 MOM: 你 在\*# 干 嘛\*?  
           ni zai gan ma?  
           you be do what?  
           *what are you doing?*  
 chi →\*leans body to screen\*  
 fig **#fig.8c**

20 GRA: 你 说&# 我 在 耍  
           ni shuo wo zai shua  
           you say I be play  
           *you say I am playing*  
 gra &gazes at the child→

#fig. 8d

21 CHI: >耍↑<  
 >shua↑<  
 >play↑<  
 >playing↑<



**Figure 8: The uncle passes the phone to the grandmother, and the grandmother provides an answer for the child.**

This conversation occurs at the opening of a video call. Before line 14, the participants were connected on the video, and exchanging greetings. At in line 14, the mother asks the two-year-old child a question: “have you finished: your meal:?”. Here, the mother uses childlike language to describe the meal: *fan:fan* (“rice: rice”). The reduplication of the word *fan* in Chinese operates as motherese, and its use clarifies that the question is addressed to the child, not the grandmother. Without any silence, the grandmother instructs the child: “You say >finished eating<” (line 15). Again, in line 20, after the mother’s second question—“what are you doing?”—the grandmother immediately provides an answer for the child without giving him any opportunity to do so himself. She again employs the prompting format “you say”. As Demuth (1986, p. 55) proved, the use of “say” is often found in question routines in interactions involving children. By prompting the child to “say,” the adults orient the child to respond with a certain verbal form. The second “you say” in Extract 8 is used in direct reported speech (“you say ‘I am playing’” rather than “you say ‘you are playing’”). The grandmother uses the first-person deictic “I” to formulate the SPP for the child, designing it from his perspective. Of course, she could instruct the child with “you say *you* are playing”. However, by using “I”, she reduces the interactional work necessary for the child to produce an SPP. The child can just repeat the SPP provided by his grandmother and complete the expected responding actions.

Importantly, when the grandmother enacts this formulation, the child looks attentively at his mother. His head is directed toward the phone, and he gazes at her during her production of the whole question—no trouble is observed, unlike with previous cases in which the children were involved in competing activities or the grandparents’ facilitation was occasioned by the absence of a response. It is possible that such non-contingent facilitation is a consequence of what Lerner (2019, p. 393) called children’s omni-relevant incompetence, including their not yet fully developed linguistic competence.

### ***Assisting the Child’s Embodied Response***

Sometimes a relevant responding action is embodied. In adult-child interactions, this is salient since children often engage in goal-oriented activity, either playfully or seriously keyed (see examples of getting children to brush their teeth in Goodwin & Cekaite, 2018). In such situations, sequence facilitation is oriented toward the practical accomplishment of an SPP, and grandparents can be recruited to offer assistance (Kendrick & Drew, 2016). Extract

9 illustrates this analysis. The father asks the 2.5-year-old child to put on a schoolbag, and the grandmother voluntarily stands up to assist the child in putting it on.

### **Extract 9**

(LBC08\_33.50)

- 05 DAD: 吴 子凡  
Wu Zifan  
Wu Zifan [child's full name]
- 06 DAD: 我 不 跟 你 好 了  
wo bu gen ni hao le  
I not with you good PFV  
I will be unhappy with you
- 07 DAD: 你 不 背 (.)书\*#包 啊  
Ni bu bei (.)shubao a  
you not carry (.)schoolbag PRT  
If you do not put on(.)schoolbag  
    >> gra \*stands up and moves tw child→  
    fig #fig.9a
- 08 CHI: 唉嘎\*#  
Eiga  
Opps  
Opps  
    gra →\*moves hand to the child's schoolbag→  
    fig #fig.9b
- 09 (0.5)\*#+(1.3)  
    gra \*holds and takes the child's schoolbag in her hand→  
    fig #fig.9c
- 10 GRA: 来 过来  
lai guolai  
come come here  
come on come here
- 11 (1.0)&#+(1.5) &+\*(0.3)  
chi &walks tw GRA&  
    gra →\*helps the child put on schoolbag→  
    fig #fig.9d
- 12 CHI: 我 不 (要)-  
wo bu (yao)-  
I not (want)-  
I don't (want)-
- 13 CHI: 我 自己 整=  
wo ziji zheng=  
I myself do  
I'll do it myself=
- 14 CHI: =我 自::己 \*整:::  
=wo ZI::JI ZHENG::  
=I MY::SELF DO::  
=I'LL DO:: IT BY MYSELF::  
    gra →\*moves hand away from schoolbag\*

```
&puts on the schoolbag by herself→
```

15      GRA: 哦, 好 好 好  
             O, hao hao hao  
             ah, ok ok ok  
             ah, ok ok ok



**Figure 9: The grandmother stands up and takes the schoolbag in her hand and helps the child to put it on.**

Prior to this extract, the mother has been chatting with the child. In line 05, the father starts to interact with the child. To do so, he first summons the child by using her full name. He then challenges the child to put on the school bag, saying that “if” the child does “not” do so (line 07), he will be unhappy with her. With this “if not” structure, the father formulates both the problem action (not putting on the schoolbag) and the negative consequences (he will be unhappy with her). This structure is similar to what Hepburn and Potter (2011) described as parents’ use of the “if–then” format when threatening children in English data. Adults encode their power to affect negative consequences when issuing such challenges and threats, thereby reinforcing the recipient’s responsibility to comply. The father’s challenge forms a request for action from the child, and putting on the schoolbag becomes a relevant SPP to respond to his request.

The grandmother stands up from the sofa before the father completely finishes his utterance (line 07). As shown in Kendrick and Drew (2016), others can be recruited to offer assistance when they anticipate a projectable trouble in a course of action. Here, the grandmother may anticipate that the child is not able to put on the bag by herself, and she assists pre-emptively. She stands up (Figure 9a), moves her body toward the child (Figure 9b), and takes the schoolbag in her hand (Figure 9c) to help the child put it on (Figure 9d). In this series of assistive steps taken to put on the bag, the grandmother is recruited to do the following two things: helping the father solicit an embodied response from the child, and assisting the child in putting on the bag. However, as the extract continues, we see that the child complains about grandmother's assistance. The child says "I'll do it by myself" (line 13) and then continues with a louder cry, "=I'LL DO:: IT BY MYSELF::" (line 14). The grandmother moves her hand away from the bag and says, "ah, ok ok ok". In this case, from the perspective of the grandmother, putting on the schoolbag projects possible difficulty for the child, she voluntarily offers assistance, not only for the child but also for the father. However, complaints are drawn from the child, since she does not have a opportunity to try doing the task on her own, thereby showing an orientation toward a preference for self-remediation over assistance in the interaction (Kendrick, 2017).

### *Animating the Child's Body in Response*

In the previous example, we saw the grandparent assisting the embodied accomplishment of the SPP, which was rejected by the child who displayed her willingness to “do it by myself”, a response that orients to and defends the child’s autonomy over her own actions. The final

fragment shows an even greater involvement by the grandparent and even less autonomy on the part of the child (aged 3). As the video call closes, the grandparent takes control of the child's body, animating it to perform the relevant embodied response: waving to the child's mother.

### Extract 10

(LET18\_closing)

```

01    MOM:    *#好    拜拜
           hao  baibai
           okay  byebye
           okay  byebye
    chi    *child's head is on screen but she gazes away from phone→
    fig    #fig.10a

02    (0.1)$#+(0.1)
    >> gra    $grandmother holds the child's hand→
    fig    #fig.10b

03    GRA:    好$#拜拜
           hao baibai
           okay byebye
           okay byebye
    gra    $lifts the child's hand up→
    fig    #fig.10c

04    (0.1)$+(0.1)
    gra    →$waves the child's hand to the phone screen→

05    GRA:    给    你    妈妈    拜拜$#
           gei ni    mama  baibai
           to  your mom  byebye
           byebye to your mom
           →$puts down the child's hand.
    fig    #fig.10d

06    MOM:    拜拜    哈
           Baibai ha
           Byebye PRT
           Byebye

07    GRA:    拜拜    哈, 我 帮$#    你 抬 手, 你 这 个 人
           Baibai ha, wo bang ni tai shou, ni zhege ren
           Byebye PRT, I help you lift hand, you this person
           Byebye, I am helping you lift your hand up, you girl
           $gra again lifts up the child's hand→
    fig    #fig.10e

08    GRA:    自己 拿 起来
           ziji na qilai
           self lift up
           Do it yourself

09    (1.0)

10    GRA:    哈哈
           Hahaha
           Hah-hah-hah

```

*Hah-hah-hah*

11 (1.6)

12 MOM: 好, 挂了  
hao, gua le  
okay, close PRT  
okay, hang up

13 (0.4)

14 GRA: 挂 了 挂 了  
gua le gua le  
close PRT close PRT  
hang up hang up

15 ((hanging up sound))



**Figure 10: The grandmother holds the child's hand and lifts it up to demonstrate a waving goodbye to the mother.**

In line 01, the mother initiates the closing by uttering a pre-closing boundary marker term (“okay”) (Beach, 1993), which projects a transition or a termination of an activity. At the moment of her initiation of the terminal exchange (Goodwin & Cekaite, 2018), the child’s head is displayed on screen, but she gazes away. It is therefore understandable that the child is held accountable for producing a return “goodbye”. In line 02, the grandmother starts holding the child’s right hand, and lifts it up (Figure 10b-10c). The child’s hand is now made visible on the phone screen (Figure 10c). The grandmother then waves the child’s hand to animate a “waving goodbye” performed by the child to the mother. While doing so, the grandmother also verbally aligns with the mother’s terminal exchanges by saying “okay byebye” (line 03). In this sense, the grandmother acts as what Goffman (1981) termed an “animator,” who produces utterances as a “sounding box” of the person responsible for the utterance (i.e., the child). Similar to Extract 6 (grandmother moves the child’s head to solicit gaze), we see that the grandmother touches the child’s body. However, this extract exhibits not only a haptic control of the child’s body, but also animation of it. While the animation results in involving the child in the parent-child sequence in an embodied way, the child is oriented with a lesser extent of autonomy than in the previous case.

When the conversation goes on, in line 05, the grandmother formulates this animation as an act to respond to the mother’s goodbye. She reaffirms that the child’s animated waving hand is *for* the mother (“bye bye to your mom”). In line 06, the mother reciprocates with a second “bye bye” suffixed with a particle to index the termination. Similarly, the grandmother aligns with a second “bye bye” while lifting the child’s hand up again and waving her hand. Interestingly, the grandmother now formulates her act of animating the child’s body as explicit facilitation (“I am helping you lift your hand up, you girl”, line 07). By this, the grandmother displays a sense of complaint and orients toward the child’s responsibility to say goodbye. She also enforces this orientation by adding to the child “Do it yourself” (line 08). This analysis has demonstrated that although the grandparents typically provide spaces and

opportunities for the child to respond, they also intervene and facilitate the non-response by treating the child as passive. However, if they do so, the grandparents may complain about the child and still orient toward the child's right and responsibility to respond.

## Discussion

This article has examined a form of sequence expansion, which we term *sequence facilitation*, in video-mediated interactions between migrant parents and the children they have left behind in China. Our proposal of the concept of *sequence facilitation* captures the practices employed by the grandparents that, in various ways, promote the completion of sequences initiated by the remote parents to their children. Sequence facilitation brings together previous studies into the discussions of this phenomenon, which provides a more holistic way of conceptualizing and studying how an unaddressed recipient of a first pair-part deploys practices to encourage the addressed recipients to provide a relevant second pair-part. Sequence facilitation, we have shown, includes both retrospective practices oriented toward FPPs, and prospective practices oriented toward SPPs. This concept encompasses the linguistic, embodied, and material practices through which a third party helps the addressed party to respond and to deal with troubles when one interlocutor is not fully competent.

The cases of sequence facilitation that we have identified reveal that the problem of responding to an FPP can, in practice, be broken down into more basic elements. The orientations of the grandparents in our data show that the production of an adjacency pair calls for a state of mutual attention and orientation (Extracts 5 and 6), that the addressed recipient must be able to hear and understand the FPP (Extracts 3 and 4), and that they must also both recognise that they should respond and know what that response should be (Extracts 7–10). Such practices expose and therefore make available for analysis features of the normative organisation of adjacency pair sequences and their local adaptations that might otherwise go unnoticed.

Furthermore, an exploration of the basic elements for accomplishing a sequence allows us to better understand the precise division of labour between the addressed and unaddressed recipients. While some facilitation practices provide minimal assistance and orient to the child's autonomy, others necessitate a greater involvement in the sequence. When a grandparent repeats a parent's FPP for a child, for example, they manage several issues on the child's behalf, such as whether the child has heard the turn and whether they have recognised it was addressed to them. However, the repetition of an FPP does not instruct the child to respond or provide them with a response to produce. It is left to the child to manage these issues independently. In contrast, when a grandparent animates a child's body and performs the embodied response on the child's behalf, the child needs neither to have heard/understood the FPP, nor to know what the expected SPP was. Sequence facilitation is thus an interactional process through which the requirements for adequate participation in an adjacency pair sequence that would normally fall to one participant are variously redistributed across participants using specific practices.

Our analysis has highlighted that embodied actions are an integral part of sequence facilitation. While the phenomenon has been previously observed in the literature, the embodied conduct was mainly neglected (see our discussions again in the background section). Apart from linguistic facilitation, our data show that grandparents' gaze movement (e.g., Extract 3), manipulation of the mobile phone (e.g., Extract 5), coordination with material objects (e.g., Extract 9), and animation of the child's body (e.g., Extract 10) are essential for establishing a possible participation framework and for facilitating the interactions between children and their parents. This may be especially salient in interacting

with young children. Prior research related to sequences with children has identified how embodied conduct is crucial to get children to accomplish sequential activities (Goodwin & Cekaite, 2018). What is even more challenging in our setting is that the speaker of the FPP relies primarily on verbal resources and cannot, for example, touch their child. The co-present grandparents have more resources at their disposal. Our findings illustrate that when grandparents facilitate sequences, they gaze toward the child, teach the child to recognise and construct the context of the interaction, and provide the body-spatial resources and cues necessary for the child to understand “when” and “how” to respond appropriately. By employing sequence facilitation, the caregiver also facilitates the child’s participation in verbal and embodied social activities, through which the child becomes an active agent in their own socialisation process. The coordination of such multimodal resources in the real world builds up what Vygotsky’s (1987) termed “semiotic mediation”, arguing that the central forms of children’s development are by the means of a mediation process. It is in the performance of the different dimensions of sequence facilitation (e.g., form of talk, body, and camera practices) that caregivers provide a semiotic scaffold device (Bruner, 1978) through which the child understands what is going on and what they are expected to perform in the situated context.

The findings also advance our understanding of sequence organisation in video-mediated communication. Existing studies on video calls have shown that visual and bodily practices are critical for people to accomplish actions (e.g., Licoppe & Morel, 2012), and our study extends them by adding new layers and complexity to the organisation of sequences in such an environment. While the affordance of visibility provided by video cameras enables the ability of participants to see and be seen, video cameras have a limited angle which requires people to be in the camera frame if they want to see and be seen (Gan & DeSouza, 2022). As we have shown, framing a mobile child is a challenging interactional task. Given the young age of the children, most of them are not physically able to hold and position the phone in an appropriate way for participation, and they are often involved in competing activities. As a result, the realisation of a sequence in video calls involving children becomes a three-party issue, and the work of facilitation becomes relevant when a visual participation framework needs to be established as a condition for accomplishing a sequence. Our analysis reveals how grandparents attempt to do this through various techno-material practices (e.g., moving the camera, maintaining the camera in front of the child’s face), and allow the visual field of video calls to turn into a ground for a parent–child sequence.

Finally, it is in the situated moments of sequence facilitation that we note how grandparents’ facilitation makes a significant contribution to these families. When a grandmother moves the phone toward a child, she not only helps establish a parent–child participation framework but also moves the mother’s on-screen body toward the child. Mother and child can therefore be “physically” and “virtually” closer. It seems that the grandparents act as an “extended body” (McLuhan, 1964) of the remote parent. While video-mediated communication is considered “co-presence by proxy” (Baldassar, 2008), our study suggests that this proxy can include not only the mediated platform (the technology) but also a vivid third-party participant (the grandparents). With so many constraints on interacting at a distance, it is hard to imagine what a remote interaction with children or other less than fully competent participants would be like without sequence facilitation.

## **Disclosure Statement**

All authors declare that they have no conflicts interest.



## Funding

The first author disclosed receipt of the financial support from Shanghai Philosophy and Social Science Foundation for Young Scholars (No. 2021EXW002). The second author disclosed receipt of the support from the Research Grants Council of the Hong Kong SAR (Project No. CUHK 14600218).

## References

- Baldassar, L. (2008). Missing kin and longing to be together: Emotions and the construction of co-presence in transnational relationships. *Journal of Intercultural Studies*, 29(3), 247-266.
- Beach, W. (1993). Transitional regularities for casual 'Okay' usages. *Journal of Pragmatics*, 19, 325-32.
- Bolden, G. B. (2012). Across languages and cultures: Brokering problems of understanding in conversational repair. *Language in Society*, 41(1), 97-121.
- Bruner, J. (1978). Early social interaction and language acquisition. In H. R. Schaffer (ed.) *Studies in mother-infant interaction* (pp. 217-289). Academic Press.
- Cekaite, A. (2010). Shepherding the child: Embodied directive sequences in parent-child interactions. *Text & Talk*, 30(1), 1-25
- Cekaite, A. (2015). The coordination of talk and touch in adults' directives to children: Touch and social control. *Research on Language and Social Interaction*, 48(2), 152-175.
- Cekaite, A. (2016). Touch as social control: Haptic organization of attention in adult-child interactions. *Journal of Pragmatics*, 92, 30-42.
- Goodwin, M. H., & Cekaite, A. (2013). Calibration in directive/response sequences in family interaction. *Journal of Pragmatics*, 46(1), 122-138.
- Goodwin, M. H., & Cekaite, A. (2018). *Embodied family choreography: Practices of control, care, and mundane creativity*. Routledge.
- Doehring, A. (2018). *Three-party interactions between neurologists, patients and their companions in the seizure clinic* [PhD Thesis]. Loughborough University.
- Demuth, K. (1986). Prompting routines in the language socialization of Basotho children. In B. B. Schieffelin, & E. E. Ochs (Eds.), *Language socialization across cultures* (pp. 51-79). Cambridge University Press.
- Gan, Y. (2021). 'Awww, she is feeding you': Young children's (dis) engagement in video calls with their migrant parents. *Research on Children and Social Interaction*, 5(1), 80-102.
- Gan, Y., & DeSouza, D. K. (2022). Moving out of view: Managing participation and visibility in family video-mediated communication. *Social Interaction. Video-Based Studies of Human Sociality*, 5(2).
- Gan, Y., Greiffenhagen, C., & Licoppe, C. (2020). Orchestrated openings in video calls: Getting young left-behind children to greet their migrant parents. *Journal of Pragmatics*, 170, 364-380.
- Garfinkel, H. (2002). *Ethnomethodology's program*. Rowman & Littlefield Publishers.
- Garfinkel, H., & Sacks, H. (1970). On formal structures of practical actions. In J. C. McKinney, & E. A. Tiryakian (Eds.), *Theoretical sociology perspectives and developments* (pp. 337-366). Appleton-Century-Crofts.

- Goffman, E. (1981). *Forms of talk*. University of Pennsylvania Press.
- Goodwin, C., & Goodwin, M. H. (1990). Interstitial argument. In A. Grimshaw (Ed.), *Conflict talk* (pp. 85–117). Cambridge University Press.
- Goodwin, C., & Goodwin, M. H. (2004). Participation. In *A Companion to linguistic anthropology* (pp. 222–244). John Wiley.
- Hepburn, A., & Potter, J. (2011). Threats: Power, family mealtimes, and social influence. *British Journal of Social Psychology*, 50(1), 99–120.
- Heritage, J. (2002). The limits of questioning: Negative interrogatives and hostile question content. *Journal of Pragmatics*, 34(10-11), 1427-1446.
- Holt, E., & Clift, R. (Eds.). (2006). *Reporting talk: Reported speech in interaction*. Cambridge University Press.
- Keating, Elizabeth (2006). Habits and innovations: Designing language for new, technologically mediated sociality. In Nick J. Enfield and Stephen C. Levinson (Eds.), *Roots of human sociality* (pp. 329-350). Berg.
- Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica*, 26, 22-63.
- Kendrick, K. H. (2010). *Epistemics and action formation in Mandarin Chinese* [Doctoral dissertation]. University of California, Santa Barbara.
- Kendrick, K. H. (2017). The preference for self-remediation over assistance in interaction. Paper presented at the *15th Annual International Pragmatics Association Conference*, Belfast, UK.
- 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., Brown, P., Dingemanse, M., Floyd, S., Gipper, S., Hayano, K., ... & Levinson, S. C. (2020). Sequence organization: A universal infrastructure for social action. *Journal of Pragmatics*, 168, 119-138.
- Lerner, G. H. (2019). When someone other than the addressed recipient speaks next: Three kinds of intervening action after the selection of next speaker. *Research on Language and Social Interaction*, 52(4), 388–405.
- Levinson, S. C. (2013). Action formation and ascription. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 103-130). Wiley-Blackwell.
- Licoppe, C., & Morel, J. (2012). Video-in-interaction: “Talking heads” and the multimodal organization of mobile and Skype video calls. *Research on Language and Social Interaction*, 45(4), 399–429.
- Liu, R. Y. (2022). Guiding children to respond: Prioritizing children’s participation over interaction progression. *Research on Language and Social Interaction*, 55(2), 184–202.
- McLuhan, M. (1964). *Understanding media: The extensions of man*. McGraw-Hill.
- Mondada, L. (2010). Eröffnungen und Prä-Eröffnungen in medienermittelter Interaktion: Das Beispiel Videokonferenzen. In L. Mondada & R. Schmitt (Eds.), *Situationseröffnungen* (pp. 277–334). Narr.

- Mondada, L. (2014). Conventions for multimodal transcription. [https://franz.unibas.ch/fileadmin/franz/user\\_upload/redaktion/Mondada\\_conv\\_multimodality.pdf](https://franz.unibas.ch/fileadmin/franz/user_upload/redaktion/Mondada_conv_multimodality.pdf).
- Pomerantz, A. (1984). Pursuing a response. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 152–164). Cambridge University Press.
- Relieu, M. (2007). La téléprésence, ou l'autre visiophonie. *Réseaux*, 144, 183–224.
- Rossano, F. (2006). When the eyes meet: Using gaze to mobilize response. Presented at The International Conference on Conversation Analysis, Helsinki, Finland.
- Rossano, F., Brown, P., & Levinson, S. C. (2009). Gaze, questioning, and culture. In J. Sidnell (Ed.), *Conversation analysis: Comparative perspectives* (pp. 197–249). Cambridge University Press.
- Sacks, H. (1992). *Lectures on conversation*. Blackwell.
- Sacks H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50, 696–735.
- Schegloff, E. A. (1988). Goffman and the analysis of conversation. In E. Goffman (Ed.), *Exploring the interaction order* (pp. 89–135). Polity.
- Schegloff, E.A. (1990). On the organization of sequences as a source of “coherence” in talk-in-interaction, in B. Dorval (ed.) *Conversational organization and its development*, (pp. 51–77). Ablex.
- Schegloff, E. A. (1997). Practices and actions: Boundary cases of other-initiated repair. *Discourse Processes*, 23(3), 499–545.
- Schegloff, E. A. (2006). Interaction: The infrastructure for social institutions, the natural ecological niche for language, and the arena in which culture is enacted. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of human sociality* (pp. 70–96). Berg.
- Schegloff, E. A. (2007). *Sequence organization in interaction: A primer in conversation analysis*. Cambridge University Press.
- 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.
- Schieffelin, B. B., & Ochs, E. (1986). Language Socialization. *Annual review of Anthropology*, 15, 163–191.
- Stivers, T. (2001). Negotiating who presents the problem: Next speaker selection in pediatric encounters. *Journal of Communication*, 51(2), 252–282.
- Stivers, T. (2013). Sequence organization. In J. Sidnell & T. Stivers (Eds.), *The Handbook of conversation analysis* (pp. 191–209). Wiley-Blackwell.
- Stivers, T., & Robinson, J. D. (2006). A preference for progressivity in interaction. *Language in Society*, 35(3), 367–392.
- Sunakawa, C. (2012). *Virtual “Ie” household: Transnational family interactions in Japan and the United States* [Doctoral dissertation]. The University of Texas at Austin.
- Vygotsky, L. (1987). Thinking and speech. In R.W. Rieber & A. S. Carton (Eds.), *The collected works of L.S. Vygotsky. Vol. 1 Problems of general psychology* (pp. 39–285). Plenum.