A CONVERSATION ANALYTIC STUDY OF ONLINE VIDEO-MEDIATED L2 CLASSROOMS: THE AFFORDANCES OF AUDIO ACTIVATION/DEACTIVATION FEATURES IN INTERACTION MANAGEMENT

ALI MOHAMMAD A ALGHAMDI

Thesis submitted for the degree of
Doctor of Philosophy
Integrated PhD in Educational and Applied Linguistics

Newcastle University
School of Education, Communication and Language Sciences

March 2023

Abstract

Recent years have witnessed a rapid growth in using videoconferencing in L2 classrooms and managing interaction in these classrooms requires participants to create novel practices to ensure its success. Such a rapid growth warrants a close examination of individuals' practices to manage these interactions in these classrooms. A review of the literature reveals the need for more work to investigate the management of talk-in-interaction in online computer-mediated-communication (Jenks, 2014), the use of the technology mediums' affordances to organise social interaction (Arminen et al., 2016); and what teachers do in online, synchronous, video-mediated classrooms (Moorhouse et al., 2022). Against this background, the current study investigates the teachers' and learners' use of audio activation/deactivation features to manage turn-taking and repair in small-group online synchronous video-mediated L2 speaking classes. Using conversation analysis, this study examines 32 hours of screen-recorded L2 speaking classes on Zoom.

The analysis shows how the on-mute learners project self-selection for the next turn using the audio activation feature. In addition, the analysis reveals the on-mute learners' fine coordination of their audio activation/deactivation with the ongoing talk. The analysis also demonstrates the participants' maintenance of boundaries between their physical space and the Zoom room using the audio activation/deactivation features. Moreover, it demonstrates how the participants employ audio activation/deactivation features to repair or pre-empt trouble caused by interference from background noises and to maintain the progressivity of the class activities.

Based on these findings, it can be said that the teachers' and learners' use of the audio activation/deactivation features has a reflexive relationship with the management of turn taking and repair in L2 video-mediated classroom interaction. The study contributes to the knowledge of using audio activation/deactivation features to manage interaction in online synchronous video-mediated L2 speaking classrooms. Broadly, the study adds to the growing field of analysing the organisation of social interaction in online, synchronous video-mediated classrooms. Furthermore, it adds to the knowledge relating to the competencies that teachers and learners require for successful interaction in video-mediated classrooms.

Acknowledgements

Doing this PhD has been a life-changing experience for me, and reaching this far would not have been possible without the support of many people. I want to use this space to express my gratitude to them.

I cannot thank enough my supervisors, Dr Muge Satar and Dr Spencer Hazel, for all their support and encouragement. Thank you for believing in me as a person and in my project's potential. Thank you for being there for me in all my ups and downs. Thank you for encouraging me and lifting my spirit when I had moments of doubt and thank you for praising my efforts when I reached my milestones. I have always revisited these words to recharge my batteries and keep going.

I also would like to thank my examiners Dr Olcay Sert and Dr Paul Seedhouse for the thought-provoking discussion and the constructive feedback that contributed to enhancing the quality of this thesis.

Huge thanks to my parents and siblings for all their support, motivation and patience. It is their unconditional love and care that made doing this possible. I would also like to express my gratitude to my wife, Samah, for all the enormous love and support. I will not forget all the sacrifices that she made to help me keep going. We have been in this together every step all the way and having you on my side made all the difference. Thank you for everything; I will be forever grateful.

A journey like this is never easy without true friends. Thank you Dhaifallah Alzahrani and Saad Albaqami, for the never-ending support and a true friendship. I will surely miss the morning coffee, the quick chat, encouraging each other and heading to the library.

I also would like to extend my thanks to all MARG members for everything I learned from them. The data sessions and seminars gave me the chance to learn from the experts and show my data, and the privilege to have these experts examine my data and think with me. Thank you for all the comments that helped me to explore new areas and expand my knowledge.

Table of Contents

List of Figures

Section…….

Chapter 1. Introduction

1.1 Setting the Scene

As the most commonly used language, speaking English has become an essential skill for people seeking to study advanced sciences, apply for jobs and advance in their careers, personal or professional development, and/or to travel abroad (Mahu, 2012). Furthermore, English is generally the primary language for communication among people of many nationalities and cultural backgrounds. The importance of learning and teaching English, thus, stems from its significance for the abovementioned purposes.

Learning and teaching English entails the engagement of teachers and learners in interaction to acquire necessary skills, such as listening and speaking. According to Walsh (2011), interaction is essential to the success of language teaching and learning. Therefore, a thorough understanding of this interaction will aid teachers in facilitating it, thereby maximising target language learning (ibid). Interaction in English learning and teaching can occur in two modes in relation to co-presence. Firstly, it can occur in in-person classrooms, where both teachers and learners are physically co-present in a physical classroom; or, secondly, teachers and learners can engage in interaction beyond the physical classroom space using the available communication technologies (with or without co-presence).

The advancement of communication technologies made it available for both language learners and teachers to engage in distance language education. Online distance education is characterised by the geographical dispersion of teachers and learners and the use of technology as a means of communication (Keegan, 2013). The availability of videoconferencing platforms such, as Zoom, Microsoft Teams, and others, has made them a preferred choice for language distance learning. As a result, distance education has rapidly grown and become more prevalent, especially during the COVID-19 pandemic (Uçar & Kumtepe, 2021). The pandemic contributed to the increasing the use of videoconferencing (Due & Licoppe, 2020) and triggered a transition to distance language learning and teaching and accelerated the advancements of the communication platforms' features. These developments made data available for more researchers to examine interaction in such settings.

In synchronous modes of distance language classes, teachers and learners engage in real-time interaction as they are co-present on an online platform. Some platforms offer text-based chat

services, while others provide text/audio chats or text, audio and video chats. For interaction to be successful and to run smoothly, it is important for both teachers and learners to acquire the necessary abilities to make good use of the medium's affordances and overcome any constrains they may face in technology-mediated interaction (Hampel & Stickler, 2005, 2015; Moorhouse, Li, & Walsh, 2021). However, managing turn-taking in synchronous video-mediated second language (L2) learning can prove to be a challenging task. Self-selection, in particular, can be challenging for learners, as there are two barriers: the language barrier and the technology barrier (Jenks, 2009; Malabarba et al., 2022). Also, interaction in such settings is vulnerable to interruptions due to aspects of trouble that, in some cases, are caused by technology or by other people or objects in the participants' physical spaces (Brandt & Jenks, 2013; Jenks, 2014). Overcoming such interruptions and undertaking repair to maintain/restore the progressivity of interaction are challenging tasks for both teachers and learners in these environments. Nevertheless, investigating the participants' use of the technology affordances to manage turn-taking and overcome trouble sources threatening the smooth running of video-mediated L2 speaking classes remains underexplored.

The current chapter outlines the central themes of the current study and makes a case for its importance. Section 1.2 presents an overview of the study and introduces its main themes, setting, and the methodology used. Following this, Sections 1.3 and 1.4 present the research objectives and demonstrates the relevance of the study; while Section 1.5 details the context of the study. The chapter concludes by sketching the organisation of the chapters of the study.

1.2 Research Overview

The current study examines the teachers' and learners' use of technology features in managing their interaction. The setting of the study is online, synchronous, video-mediated English language speaking classes. To examine the participants' use of technology affordances, the research uses conversation analysis (CA) as a methodology (Sacks, 1992). The following sub-sections elaborate on the central themes in the study, namely, L2 classroom interaction, L2 video-mediated interaction and, finally, CA.

1.2.1 L2 Classroom Interaction

The interaction in L2 classroom is described as having a unique organisation compared to that of ordinary interaction (Sert, 2015b). Similar to other institutional settings, a significant line of research has investigated the structure of L2 classroom interaction as a primary interest (Gardner, 2019). Early research in classroom interaction investigated the organisation of turn-

taking, sequence and repair in classroom interaction (Mehan, 1979; Mchoul, 1978; McHoul, 1990; Sinclair & Coulthard, 1975). For instance, Sinclair and Coulthard (1975) revealed findings that showed a pervasive pattern in classroom interaction: the sequence of initiation-response-feedback (IRF). Further, findings from early CA studies of classroom interaction revealed interesting results demonstrating its deviance from interaction in ordinary conversation. Not only was it seen to be different in general structure, but it can also have different varieties depending on the context at any particular moment (Seedhouse, 2004b). In addition, the findings of these studies explored the teacher's role in managing interaction and directing turn-taking. While some studies viewed the teacher as the only party able to direct turn-taking in classroom interaction (Mchoul, 1978), subsequent studies benefited from video recordings and argued that learners' displays of bidding for the next turn and showing willingness to be the next speaker using practices, such as gaze and hand raising, could play a role in directing the teacher's nomination of who speaks next (Lauzon & Berger, 2015; Kääntä, 2012; Mortensen, 2008; Sahlström, 2002).

The literature on classroom interaction also revealed valuable findings on the methods by which learners self-select and initiate turns without being selected by the teacher (Dolce & van Compernolle, 2020; Mortensen, 2009; Solem, 2016; Waring, 2009). While some researchers demonstrated that these learners' initiatives were accepted by teachers and created affordances for learning (Sert, 2017; van Balen, Gosen, de Vries, & Koole, 2022), others revealed that teachers and learners treated these initiatives as inappropriate behaviour and resulted in the reproaching of learners (Hazel & Mortensen, 2017; Petitjean, 2014; Rodriguez & Wilstermann, 2018). Research into classroom interaction also revealed interesting observations related to the preference of teachers and learners for progressivity (Lauzon & Berger, 2015; Sert & Walsh, 2013). This preference manifests itself in successive repair initiations (Hırçın Çoban & Sert, 2020) and the management of interactional troubles created by the learners' claims of insufficient knowledge, for instance (Lauzon & Berger, 2015; Sert & Walsh, 2013). However, this preference for progressivity may collide with other preferences and result in halting progressivity momentarily to attend to instances of transgressions to classroom order (Hazel & Mortensen, 2017; Hosoda & Aline, 2013; Margutti, 2011).

The majority of these findings relating to classroom interaction come from in-person classes. A worthy investigation could be conducted to ascertain how interaction in video-mediated L2 teaching and learning settings is managed. The lack of physical co-presence and the enablement and constraints of communication platforms' features can contribute to shaping

social interaction in such settings. The following section presents findings from previous research studies that have focused on L2 video-mediated interaction.

1.2.2 Online Synchronous L2 Video-Mediated Classroom Interaction

Interaction of L2 learners and teachers beyond physical classrooms has been examined by many researchers deploying different methodological standpoints, such as multimodal (inter)action analysis (Satar & Wigham, 2020, 2023; Wigham & Satar, 2021) and social semiotics (Sindoni, 2014). While some researchers have examined areas, such as the effectiveness of teaching and learning L2 using video communication software (McAndrew, Foubister, & Mayes, 1996; Ockey, Timpe-Laughlin, Davis, & Gu, 2019; Yousefi, 2014), others have sought to investigate the structure of interaction in such settings. This is consistent with the focus of the current study, as its interest lies in examining the methods by which teachers and learners manage interaction in L2 video-mediated classrooms. Similar interest has been expressed by a number of scholars using CA. According to González-Lloret (2015), two research strands exist in studying interaction in online language learning and teaching settings using CA. These studies have either described interaction in online language learning audio/video-mediated language learning settings (descriptive studies) or have tracked the development of the participants' practices over a period of time (i.e. longitudinal studies). In both study strands, researchers have examined the organisation of turn-taking (Badem-Korkmaz & Balaman, 2022; Guo & Zhang, 2021; Jenks, 2009; Malabarba et al., 2022); and the repair of troubles that may arise in such settings (Brandt & Jenks, 2013; Jenks, 2014; Rusk & Pörn, 2019).

Features of technology can enable and constrain the participants' production of social actions. This relationship between such features of technology and the structure of social interaction can be examined using the notion of affordances (Hutchby, 2001a, 2003a, 2014) (the notion of affordances will be revisited in Section 2.3.1.1). Following this approach, the interest is not only given to technology or to interaction; instead it is directed to both technology and interaction, as well as to present evidence of how the use of the technology affordances plays a role in how participants structure their interaction (Hutchby, 2001b). The previous literature on video-mediated interaction demonstrates that the participants can use the affordances of the mediating technology to overcome the constraints imposed by the lack of physical co-presence (Due & Licoppe, 2020; Luff et al., 2016; Melander Bowden & Svahn, 2020; Mlynář et al., 2018).

The lack of co-presence and connecting from different physical spaces may create incongruities in the interaction, such as difficulties in recognising gestures, gaze direction, and referential practices since participants have asymmetrical or limited access to one another's immediate environments. This phenomenon is referred to as "fractured ecologies" (Luff et al., 2003, 2016), which means that interaction in these settings is "fractured from the environment in which it is produced and from the environment in which it is received" (Luff et al., 2003, p.55). For example, pointing gestures may lose their significance in online, synchronous video-mediated interaction, as demonstrated by Melander Bowden and Svahn (2020). Indeed, they demonstrate how the tutor and tutee used the mouse cursor to do pointing to parts of the learning materials on the shared screen in videoconferencing. (Fractured ecologies are further discussed with examples in Section 2.3.1.2).

In the study of turn-taking in audio/video-mediated L2 learning and teaching settings, several topics have been investigated, such as managing overlaps (Jenks, 2009; Malabarba et al., 2022), turn allocation and eliciting responses (Badem-Korkmaz & Balaman, 2022; Guo & Zhang, 2021), and displaying recipiency and willingness to take the next turn (Stone & Brinham, 2022). These studies revealed valuable findings regarding the importance of timing turn-initiation (Jenks, 2009) and the methods by which participants resolve overlapping talk (Malabarba et al., 2022). Additionally, the studies cited above provided important findings relating to turn allocation and how participants in video-mediated classes display their willingness to be the next speaker (Badem-Korkmaz & Balaman, 2022).

Interruptions to the smooth flow of interaction in L2 audio/video-mediated learning and teaching settings have also been examined. A number of aspects of trouble have been considered, such as transmission delays (Rusk & Pörn, 2019), difficulties in joining an ongoing conversation (Brandt & Jenks, 2013), and interference from background noises (Alzaidi, 2016; Jenks, 2014). Similarly, Jenks' (2014) study reported interesting findings regarding the consequentiality of interference of such noises to the ongoing interaction in L2 audio chatrooms; such noises can terminate the ongoing conversation or alter its trajectory (ibid). The examination of the consequentiality of interference from background noises is relevant to the current study.

In regards to the repair organisation in online synchronous audio/video-mediated interaction, it is reported to have some differences from that of in-person interaction. These differences stem from the awareness of producing a trouble source due to the lack of access to what others hear or see (Rintel, 2013, 2015), and difficulties in locating the origin of the trouble

source (Ruhleder & Jordan, 2001a). These differences may have an impact on the repair trajectories. For example, Schegloff et al. (1977) stated that the trouble source producer can have the first access to initiating repair in the same turn or in the transition space. This may not be the case in video-mediated settings, as the trouble source producer may not be aware that their turn was distorted or delayed due to connection problems. Thus, the recipient in this case is the first to have access to the initiation of repair (Rintel, 2013). Repair and aspects of trouble in L2 audio/video-mediated interaction are presented in Section 2.3.3.2.

Some scholars have argued that teachers in these settings need to acquire a number of competencies in order to be successful in online synchronous settings (Grammens et al., 2022; Guichon, 2009; Hampel & Stickler, 2005, 2015; Moorhouse et al., 2021, 2022). The aforementioned studies, however, have focused mainly on the teachers' side of the interaction; and less attention has been paid to the learners' side. The importance of learners' digital competence has been stressed in relation to maximising their opportunities to learn and perform in online settings (Gunawardena et al., 2001; Mosa, Mahrin, & Ibrrahim, 2016). In language learning, such competencies enable the learners to cope with the challenges in online learning and, thus, create more space for learning the target language (Del-Moral-Pérez, Villalustre-Martínez, & Neira-Piñeiro, 2019). Nevertheless, a comprehensive analysis of what makes a successful management of L2 video-mediated classroom interaction needs to consider *both* parties (i.e. teachers and learners). The current research has generated findings relevant to what abilities teachers and learners need to acquire to manage interaction in such settings successfully.

Overall, the current study investigates how turn-taking and repair are organised using Zoom communications platform's audio activation/deactivation features (see Chapters 5,6 and 7). To carry out this investigation, CA is employed as a research methodology to guide the entire research process. The following section briefly considers CA and justifies its use in this study.

1.2.3 Conversation Analysis

This section provides a brief presentation of CA (it will be revisited in Chapter 3). Harvey Sacks developed CA during the 1960s and 1970s as an approach concerned with the study of human interaction. Stivers (2015) defined CA as an approach that "identifies and describes the practices that interactants use in talk-in-social interaction and uses these results to understand and describe the underlying structural organisation of social interaction" (p.1). CA was further developed by Sacks' collaborators, Emmanuel Schegloff and Gail Jefferson. Although CA's name suggests that it is concerned only with ordinary conversations,

Schegloff (1984) noted that CA is an appropriate approach for studying different kinds of social interaction. This was proven true, as CA has been deployed to examine social interaction in different institutional settings, such as courtrooms (Komter, 2012) and classrooms (Seedhouse, 2004). In doing CA, practitioners begin by collecting naturally occurring data, which means that the social interaction captured would have occurred, regardless of the researcher's presence.

The scope of CA was broadened from a focus on verbal conduct to include embodied conduct, such as body movement, gesture, gaze shifts, and the manipulation of objects. An important factor that has played a role in this is technological advancements, which have allowed CA practitioners to capture social interactions in videos. A number of principles need to be followed when examining social interaction. Seedhouse (2005, pp.166-167) outlined these principles, starting from Sacks's (1984) argument that social interaction is systematically organised; and there is order at all points. Seedhouse also highlighted CA's understanding of context, mentioning how the participants' contributions shape and renew the context (Heritage, 1984b). This emphasises the importance of studying social interaction in the sequential environments in which it occurs, as turns are produced in response to what preceded them. Their production creates a context for what follows them, which subsequently renews the context. Further, Seedhouse contended that CA uses a transcription system that caters to the tiniest details of social interaction, making interaction open to close examination. Moreover, CA is bottom-up and data-driven, meaning that all claims researchers make must be grounded in the data (Seedhouse, 2005). Such principles and other validity and reliability measures (see Section 3.6) render CA a rigorous approach that produces highly reliable research findings.

CA adopts an emic perspective of analysis, meaning that researchers base their claims on the participants' orientations and do not start from any assumptions. CA aims to investigate why a specific action is produced in a particular manner at a certain position in a sequence (Seedhouse, 2004). It seeks to understand and unpack the methods by which participants achieve mutual understanding. These are the methods by which participants take turns and repair trouble in their talk, for example. The stance of CA on the explication of the participants' methods to take turns and attend to aspects of trouble that hinder the progressivity of their interaction is relevant to the current study, which aspires to investigate teachers' and learners' use of technology features in the management of turn-taking and repair in video-mediated L2 classrooms.

CA has been tested in the examination of video-mediated interaction in various settings. For example, it was employed to examine interaction in mundane settings (Licoppe & Morel, 2014a; Rintel, 2015b), semi-experimental settings (Luff et al., 2003, 2016), and different institutional settings (Badem-Korkmaz & Balaman, 2022; Mondada, 2015; Stommel & Stommel, 2021). Using CA, these studies have revealed interesting findings relating to how social interaction is organised in video-mediated settings. They have also shown that CA is a useful research methodology for examining technology-mediated social interaction.

Having briefly presented the current study's research methodology, the following section outlines the research objectives this study seeks to fulfil.

1.3 Research Objectives

The overriding objective of the current study is to examine participants' use of technology features in organising their interaction in online synchronous video-mediated L2 speaking classes. Specifically, the study investigates the participants' use of audio activation/deactivation features in managing aspects of turn-taking and repair. Firstly, the study seeks to demonstrate how learners use the audio activation feature to project the possible completion of the turn-in progress and indicate that they are the potential next speaker. The analysis will track the sequential positions where the on-mute learners (OMLs) activate their audio and self-select for the next turn. Moreover, the analysis seeks to examine how these learners mark the completion of their participation by deactivating their audio after taking a turn. The study also aims to show the consequentiality of mistiming the audio activation to these learners' ability to self-select and access an ongoing conversation.

The study also seeks to examine the teachers' and learners' use of the audio activation/deactivation features for managing trouble that originates from background noise interference. It also aims to examine how the teachers use these features to locate where the noise originates from when the trouble source producer is unknown. In addition, this study seeks to investigate the use of these features to handle noises from a known participant's background. The following section demonstrates the importance of conducting the current study.

In pursuing the fulfilment of the aforementioned objectives, the study will mainly argue that the teachers' and learners' use of the audio activation/deactivation features has a reflexive relationship with the management of turn taking and repair in L2 video-mediated classroom interaction.

The current study also argues that:

- The successful use of the audio activation/deactivation features enables the participants to participate in class activities while isolating their background noises.
- The unsuccessful use of these features is noticeable and accountable by the other participants and can hinder the OML's participation and potential learning.
- The development of CIC and e-CIC plays a role in advancing the progressivity of interaction and creating more opportunities for participation and learning in L2 video-mediated classrooms.

Having presented the objectives of the study and its main arguments, the following section sheds light on its significance and aspects of originality.

1.4 Significance of the Study

The current study is original as it adds to the knowledge of interaction in L2 video-mediated speaking classes in a number of areas. Firstly, it adds to the knowledge regarding the participants' use of the medium's affordances in L2 video-mediated speaking classes. The notion of affordances (Hutchby, 2001b, 2003, 2014) presents an approach to studying how the mediating technology affords or constrains social interaction. However, as Arminen et al. (2016) stated, there is a need for more work to be undertaken to more fully understand participants' use of the medium's affordances in the organisation of social interaction. In L2 video-mediated learning and teaching settings, various studies have generated findings on the use of different affordances, including those examined in this study (Andrews, 2020; Cheung, 2021; Moorhouse et al., 2021; Moorhouse, Walsh, Li, & Wong, 2022b; NurSürüç Şen, 2022). However, none of these studies has presented a detailed multimodal microanalysis of the teacher's and learners' use of audio activation/deactivation features in organising aspects related to turn-taking and maintaining or restoring the progressivity of the video-mediated L2 classroom interaction.

Relevant to the use of these features, another important area to explore is the timing of using them and coordinating their use with the actions of other participants. As Jenks (2009a) points out, an essential skill that the participants in L2 audio chats need to acquire is good timing of launching their turns, as mistiming can result in overlaps. Andrews (2020) reported a similar difficulty in video-mediated classes, as the muted learners faced a challenge in knowing when to join an ongoing conversation without risking overlapping with the current speaker's talk. Being muted adds more complexity to the situation. The current study presents a moment-by-

moment analysis that demonstrates how the participants time and finely coordinate their audio activation with the turn in progress which projects its possible completion and signals their incipient speakership. Hence, this study not only adds to the literature on turn-taking organisation in video-mediated L2 learning and teaching settings, but also adds to our knowledge of how participants interact in fractured ecologies (Luff et al., 2003, 2016) maintain the boundaries between their physical space and the class-shared environment (Fornel, 1996), as they have potential sources that can produce background noises.

The absence of this maintenance of boundaries can result in the interference of background noises in ongoing class activity. This, in turn, can hinder the progressivity of classroom activities. Previous researchers have noted that background noises are one of the aspects that cause trouble and can be consequential for the progressivity of the ongoing interaction in both voice-based (Brandt & Jenks, 2013; Jenks, 2014) and video-mediated L2 learning and teaching interaction (Cheung, 2021; NurSürüç Şen, 2022). However, as Jenks (2014) points out, there is a need to examine how the participants repair such troubles and manage their interaction. Even though the studies cited above have referred to the challenges caused by background noises in audio/video-mediated L2 interaction, the body of knowledge in the field lacks a moment-by-moment multimodal analysis of how the occurrence of background noise interference is managed by teachers and learners in L2 video-mediated speaking classes using the audio activation/deactivation features.

1.5 Context of the Study

The context of the current study is small-group, online, synchronous, video-mediated L2 speaking classes. Each class lasted for about 50 minutes, during which the participants discussed different topics in each class. The classes took place on the Zoom communication platform, and in some cases were supported by using other platforms, such as Google Slides and Menitmeter. A total of 30 L2 learners connecting from different countries and five teachers from the United States of America (USA) and Brazil participated in this study. There are two groups of data collected for this study: the first was collected between September 2019 and February 2021; while the second was collected between October and November 2021. A more detailed description of the context will be presented in Chapter 4.

1.5.1 Zoom

As noted, the classes recorded in this study took place on the Zoom communication platform (more details in Section 4.2.1). It is a platform that allows the participants to communicate in real time using text, audio, or video. In addition, it enables the participants to share their screens and share feedback using the different reactions icons available. Participants can remain in one room, or the host (i.e. the teacher in this study) can divide them into breakout rooms. In the classrooms, teachers may combine the use of Zoom with other supporting tools, such as Google slides or Google documents, which can enable collaborative work on tasks or text co-construction in small groups (Kohnke & Moorhouse, 2022). To manage interaction, Zoom affords both the host (teacher) and hostess (learners) a number of features, such as the ability to turn on/off their audio, video, or both. However, only the hosts and co-hosts have exclusive access to deactivating the audio or video of themselves and others. As noted, investigating the participants' use of the audio activation/deactivation features to manage turn-taking and repair in Zoom is important to the current study.

1.5.2 Audio Activation/Deactivation Features

The reason these features are labelled as audio activation/deactivation in the current study and not mute/unmute as named by the Zoom developers is in consideration of what using these features does. By clicking the mute button, the user deactivates their audio and vice versa. The microphone icon next to their names on the screen represents the participants' audio status. If the audio is activated, only the participant's name appears at the bottom left corner of their video frame. However, a microphone icon with a red slash crossing it or a red microphone in some versions of Zoom next to the participant's name indicates that his/her audio is deactivated, either by themselves or by the host (see Figure 1.1). Participants can activate/deactivate their audio by clicking the 'mute' button in the control bar. Once clicked, it changes to the 'unmute' button, through which audio can be activated again (see Figure 1.1). Zoom automatically places a green box around the current speaker's video frame, as demonstrated in Figure 1.2. In other cases, the green box appears around the video frame of the participant with a louder sound, such as in the occurrence of background noises. Should the teacher deactivate a learner's audio, a notification appears on the latter's screen to alert them of this change to their audio status. This notification, in turn, makes the host's decision to deactivate a learner's audio available to them and other participants, as the microphone icon will appear next to their names (Section 4.2.1 offers further details of Zoom features).

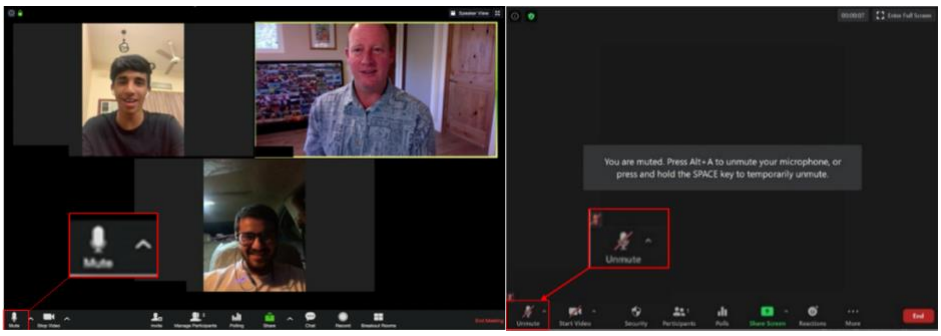


Figure 1.1: The Mute/Unmute Buttons as they Appear on the User's Screen.

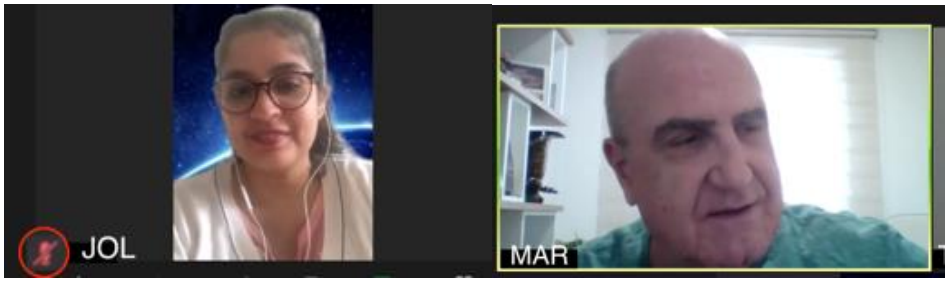


Figure 1.2: Audio Status as it Appears on Screen (off on the left and on in the right picture).

The following section sheds light on the organisation of the chapters in the current study.

1.6 Organisation of the Thesis

This opening chapter has set the scene for the current study and provided an overview of the research by briefly presenting its central themes in the current study. It has also presented the research objectives and demonstrated the relevance and significance of the current study by identifying the research gaps it aims to fill.

The second chapter introduces in more detail the literature relevant to the main focuses of the thesis. It begins by reviewing the literature on the organisation of L2 classroom interaction, including turn-taking and aspects of interactional troubles that can cause a halt to the progressivity of interaction. The chapter then narrows the discussion to the literature relevant to online synchronous video-mediated interaction. It provides an overview of computer-mediated communication (CMC) and the specific mode examined in this study – computer-mediated spoken interaction (CMSI) in L2 learning and teaching settings. In this section, the notions of affordances and fractured ecologies are presented before the chapter moves on to review the literature on the organisation of turn-taking, repair and aspects of trouble in video-mediated L2 teaching and learning classrooms. Also, the necessary competencies for teachers in such settings are presented. The chapter then concludes by presenting the research gaps the study aims to fill.

The third chapter presents CA as the research methodology and a theory of analysing social interaction employed in this study. The chapter begins by identifying the purpose of the study before presenting an introduction to CA. The epistemological foundations of CA - ethnomethodology – are then considered. Following that, the interactional machineries of sequence, turn-taking, and repair are presented. This is then followed by a section on the measures taken to ensure the validity and reliability of the study. The chapter concludes by presenting some of the limitations and criticisms of CA and how they have been addressed in this study.

Chapter 4 offers more details regarding the research setting, including the platform examined in the current study, the classes recorded, and the study's participants. Additionally, the chapter explains the process of data collection and the tools used for that purpose. Next, the details relevant to the data collected and how they were transcribed are provided. The Discovery of the phenomena examined in this study and how they are analysed in the analysis chapter are then presented. The chapter closes by setting out the ethical considerations and measures to ensure the participants' privacy/anonymity and how the data will be stored.

Chapter 5 offers an analysis of the learners' use of the audio activation features of Zoom to project the possible completion of the current speaker's turn and signal self-selection to speak next. The analysis tracks the sequential positions of the audio activation by learners who have their audio deactivated. This chapter also offers an analysis of the importance of the timing of audio activation and the consequences of mistiming it for those learners. The chapter then moves on to consider the learners' audio deactivation after activating it to participate in the ongoing activity. The sequential positions of audio deactivation are also identified and analysed. Moreover, the chapter's final section presents an analysis of the interactional consequences of the absence or delay of audio activation.

Chapter 6 provides an analysis of the teachers' and learners' use of the audio activation/deactivation features of Zoom to search for the trouble source. It shows the teachers' use of these features to identify the participant feeding their background noises in the shared Zoom space. In addition, the analysis shows the collaborative work by the teacher and the learners towards identifying the trouble source producer, deactivating their audio and restoring the progressivity of the suspended class activity. The chapter presents an analysis of the non-minimal other-repair initiated sequences to tackle the interference of disruptive noises.

Chapter 7 presents an analysis of the teachers' use of the audio activation/deactivation features of Zoom to tackle trouble caused by background noises coming from a known source. The analysis demonstrates how knowing the source of the disruptive noise can minimise the length of the repair sequences. Such minimal repair sequences are demonstrated in the chapter. The chapter also demonstrates the teachers' use of the audio deactivation affordance as a pre-emptive procedure to potential trouble.

Chapter 8 is an in-depth discussion of the research findings in light of the results from the literature. Finally, Chapter 9 offers a conclusion of the study in addition to the methodological considerations, limitations of the study, implications, and potential directions for future research.

1.7 Summary

This opening chapter has provided an overview of the research in the current study. Section 2 briefly introduced classroom interaction and L2 video-mediated classroom interaction as the setting of the current study. CA as a research methodology and a theory of interaction was also presented. Section 3 set out the research objectives, followed by demonstrating the relevance and importance of the current study in Section 4. Next, Section 5 briefly presented the context of the study. Section 6 sketched the organisation of the thesis and briefly presented the content of the following chapters. As noted, the next chapter details the existing research on managing classroom interaction in in-person and online video-mediated L2 learning and teaching settings.

Chapter 2. Literature Review

2.1 Introduction

The current study seeks to examine the participants' strategic use of audio activation/deactivation features of Zoom to organise their turn-taking and repair in online, synchronous, small groups, video-mediated L2 speaking classrooms. Relevant to the research focus are a number of central key concepts, namely, classroom interaction, turn-taking overcoming interactional troubles in classroom, and video-mediated L2 classroom interaction. The current chapter aims to introduce these key concepts and review findings from the previous research relevant to these concepts. Therefore, it is important to begin with a brief discussion of the characteristics of classroom interaction (Section 2.2). This is followed by a discussion of the organisation of the turn-taking system in classroom interaction (Section 2.2.1) and a review of the relevant literature on turn allocation and students' self-selection for the next turn (Section 2.2.1.1). Next, the relevant literature on the aspects of trouble that can hinder the progressivity of classroom interaction and methods to repair them is reviewed (Section 2.2.2).

Having presented these concepts, the second section in this chapter presents a review of the previous research on online language learning and teaching settings (Section 2.3). To do this, the section begins by introducing Computer-Mediated Communication and Computer-Mediated Spoken Interaction (CMSI). This is then followed by presenting two important concepts for the understanding of interaction in CMSI: affordances (Section 2.3.2.1) and fractured ecologies (Section 2.3.2.2). This enables the possibility of a discussion into how the use of technology features while interacting in fractured ecologies can impact the organisation of turn-taking (Section 2.3.3.1) and aspects of trouble and repair (Section 2.3.3.2) in video-mediated L2 classrooms. Section 2.4 sheds light on the competencies necessary for language teachers in L2 video-mediated classrooms. Finally, the chapter concludes by outlining the research gaps, which the current study aims to bridge.

2.2 Classroom Interaction

In its simplest definition, classroom interaction refers to the interaction between teachers and learners or learners with their fellow learners in a classroom (Tsui, 2001). Generally speaking, interaction in institutional settings is described to differ from ordinary conversations. One finding is that interaction in such institutional settings is *goal-oriented,* and these institutional

goals impact how the participants' interaction is shaped (Kurhila, 2004). In addition to being goal-oriented, Heritage (2005, p.106) points out two more characteristics of interaction in institutional settings; first, its link to the participants' identities, such as teacher-learner, and how this may shape turn-taking and the overall structure of interaction. Secondly, interaction in such settings can be constrained in terms of what is considered acceptable or not in the activity in progress. Also, interaction is produced in association with procedures specific to a particular institutional context. The institutionality of these settings is brought into interaction and is demonstrated in the participants' conduct. It should be noted that all the above differences are not to be attributed to a particular context unless they are oriented to by the participants during interaction (Heritage, 2005; Schegloff, 1991).

In relation to the language classrooms, Seedhouse (2004) suggests three main characteristics of English language teaching classrooms, which can apply to all language learning classroom interaction:

1. Language is both the vehicle and object of instruction.
2. There is a reflexive relationship between pedagogy and interaction and interactants constantly display their analyses of the evolving relationship between pedagogy and interaction.
3. The linguistic forms and patterns of interaction which the learners produce in the L2 are potentially subject to evaluation by the teacher in some way. (pp.183–184)

Classroom interaction has been a subject of study by researchers following different methodologies. For instance, Sinclair and Coulthard (1975) used a discourse analytic approach in their study of the initiation-response-feedback IRF sequences. Others have used Systemic Functional Grammar (Gibbons, 1998); and Dynamic social network (Bokhove, 2018). CA has also been used by previous researchers to analyse classroom interaction (Kääntä, 2010, 2012b; Mortensen, 2008, 2009; Sahlström, 2002; Seedhouse, 2004). As Gardner (2019) notes, a major strand in the research examining classroom interaction using CA has taken a primary interest in the analysis of how turn-taking, sequence, and repair are organised. An understanding of turn-taking management in classroom and aspects of trouble that can hinder the progressivity of interaction is useful in the understanding of the analysis of the interactional phenomena examined in the current study. Next, a review of such topics is provided before studies on the online context are considered.

2.2.1 Turn-taking in Classroom Interaction

The study of turn-taking is one of the main interests of CA. Sacks et al. (1974) illustrated that two components characterise the organisation of turn-taking in talk. The first component is what they termed the 'Turn Constructional Unit' (TCU), which is concerned with the units that comprise a turn at talk and how these units project their completion (more details on projection are presented in Chapter 3). The completion of a turn opens up a transition relevance place (TRP) which creates a possibility for a speaker change to occur (i.e. another participant speaks next). The techniques by which speakership transfers between participants is termed the 'Turn Allocation' component (ibid). In ordinary conversations, Sacks et al. (1974) note that a general 'rule' in conversation is that "one party talks at a time", and there is a tendency towards no gaps between speakers (p.700). Further, overlaps are kept short using resolution devices (Schegloff, 2000). Moreover, Sacks et al. (1974) presented a set of locally managed practices by which the interactants take turns in a conversation. These practices are grouped in terms of the next speaker selected by the current one or the next speaker self-selects and take the next turn.

The study of turn-taking system in classroom context has been of interest to many researchers. As Sacks et al. (1974) state, there are differences between turn transition trajectories in ordinary conversations and other settings, depending on the type of activities being undertaken. This relationship between the different activities and the organisation of turn-taking in these settings was illustrated in the early works by analysts who studied classroom interaction, such as McHoul (1978) and Seedhouse (1996, 2004).

Before moving to discuss CA work on the organisation of turn-taking in classroom interaction, an important finding from non-CA work needs to be considered. These findings concern the prevailing pattern of turn-taking in teacher-led classroom interaction: the initiation response feedback (IRF), or initiation-response-evaluation (IRE), as presented by Sinclair and Coulthard (1975) and then later in Mehan (1979). These patterns consist of three turns, the teacher initiates the turn (a question, for example), with the learner providing a response, and this response is then followed by a turn in which the teacher provides feedback or an evaluation. Following a CA perspective, Ingram and Elliot (2014) claim, however, that the IRF is not a single sequence type; rather, it is composed of an adjacency pair, consisting of question and answer, followed by a third turn that can close the sequence, or instead can pave the way for a wide range of class activities to ensue (Lee, 2007).

McHoul (1978) investigated the turn-taking system in formal classroom talk and the 'rules' that govern it. Adopted a conversation analytic perspective, McHoul's study compared the turn allocation rules in the classroom data to the rules suggested by Sacks et al. (1974). Mchoul (1978, p.188) suggested several modifications to the system that are specific to formal classroom talk. If the teacher selects a student to speak next, then that particular student has the right and obligation to take that turn, while the others do not have such a right or obligation. When the teacher is the current speaker and does not select a student for the next turn, then the teacher must continue. When the student is the current speaker and designs their turn to select the next speaker, the right to take the next turn is for the teacher and no other. If the student does not select the next speaker, then self-selection may be instituted with the teacher as the first to take the next turn. Alternatively, the student may hold the turn, unless the teacher reclaims it. Additionally, McHoul points out that it is "only teachers can direct speakership in any creative way" (Mchoul, 1978, p.188).

According to Ingram and Elliott (2014), McHoul's findings clearly show how the turn-taking 'rules' in formal classrooms can deviate from those that govern turn-taking in ordinary conversations. According to the turn allocation trajectories presented in McHoul's study, they deviate from those in ordinary conversation (Sacks et al., 1974) in the sense that gaps and pauses can be maximised. Teachers are able to pause during their turn in progress with no potential risk of interruption, as the students do not seem to self-select. Moreover, as Ingram and Elliott (2014) pointed out, the possibility of overlaps is lower in McHoul's findings, as the teacher has the first right to the next turn and also to determine when the right to the turn is allocated to the specific student selected by the teacher.

Turn-taking in classroom is dynamic and can vary depending on the pedagogical focus. Seedhouse (2004) investigated the organisation of turn-taking in second language classroom interaction. He argues that "there is a reflexive relationship between pedagogy and interaction" (p.178). This means that the pedagogical focus of the class can have its fingerprint on the way turn-taking, sequence, and repair are organised. Further, Seedhouse proposed four classroom contexts based on the various pedagogical foci, in which the organisation of turn-taking, sequence, and repair also varies. These particular contexts are form and accuracy, meaning and fluency, task-oriented, and procedural contexts. For example, in the meaning and fluency contexts, the focus is on offering the learners more opportunities to talk about personal meanings and to express themselves. Contrary to the accuracy and form context, the teacher does not place much focus on the learner's production

of accurate linguistic forms. Seedhouse (2004) argues that each pedagogical focus requires a special organisation of interaction.

Similar questions to the above regarding the management of turn-taking in classroom have been explored by many researchers, covering two main areas: how the teachers allocate turns to the students in the classroom (Kääntä, 2012b; Lee, 2017; Mortensen, 2008; Sahlström, 2002; Van Lier, 1988; H. Waring, 2013), and how learners self-select for the next turn (Lee, 2017; Sahlström, 2002; Takahashi, 2016). Regarding the teachers' allocation of turns to learners, Van Lier (1988) contended that teachers can verbally nominate learners or can otherwise deploy non-verbal practices, such as gazing at the selected learner. Verbal nomination is usually done by using address terms, such as the student's name before or after a question (Lauzon & Berger, 2015; Amar, 2020). For instance, Amar (2020) examined teachers' turn allocation practices in English classrooms and found that teachers used summons (i.e. use of the first or full name) following a question. In their data, the teacher used summons to nominate a particular student or a small group of students to speak next. Following the nomination of a small group, the selection of who takes the answering slot becomes a negotiable matter for the students. Using talk and gesture, the small group selects the next speaker to provide an answer to the teacher's question.

However, as noted, teachers' nomination of the next speakers is not always verbal. A great deal of the literature has investigated the embodied turn allocation by teachers (Mortensen, 2008; McHoul, 1978; Van Lier, 1994; Sert, 2011; Kääntä, 2012). Kääntä (2012), for example, examined the teachers' use of embodied means, such as head nods, gaze, and pointing to select the next speaker in instructional interaction and how this is well-coordinated with the emerging interaction and the ongoing participation frameworks. Kääntä notes that the co-presence of the participants is essential to the success of such embodied turn allocation practices. Moreover, Kääntä showed that teachers allocated turns to the bidding students using embodied means only or with minimal talk. Furthermore, the successful allocation of turns using the pointing gesture, for instance, is due to the interactional work undertaken by the students, as the bidding students recognise the gesture and orient to the next turn as allocated to them and launch it, while the other students do not claim the next turn.

This observation by Kääntä highlights the importance of considering the learners' displays of availability and willingness to take the next turn (Lauzon & Berger, 2015; Mortensen, 2008). This also entails an examination of the practices which the learners use to bid for the next turn. One of the practices that was found to be prevalent and efficient in classrooms to bid for

the turn and to gain the teacher's attention is hand raising (Kääntä, 2012; Mondada, 2009; Mortensen, 2008; Sahlström, 1999, 2002). In their study of hand raising to bid for the next turn, Sahlström (2002) indicated that the learners finely tune their hand raising with the possible completion of the turn-in-progress in that they raise their hands at the possible transition places. Sahlström also reports that hand raising may not occur on its own; rather, it is accompanied by gazing at the teacher. This sequential position of the hand raising gesture by learners functions as a display of bidding to speak at the next slot. In teacher-fronted classes, the multiple biddings by learners for the next turn provide the teacher with different options to choose from for the next turn. In younger age groups, in addition to hand raising, summons (i.e. calling the teacher's name or other address terms such as Mr. or Miss) is another resource that learners use to draw the teachers' attention and, thus, be selected for the next turn (Maroni, 2012).

Mortensen (2008) noted that teachers scan the class in search for those students who display a willingness to take the next turn by establishing mutual gaze with the teacher. The sequential place of the student's gaze to the teacher is oriented to by the teacher as a display of willingness to be the next speaker, which is then followed by the teacher allocating the turn to that student (ibid). However, next speaker selection by the teacher is not always directed to learners who display willingness to participate by establishing mutual gaze. On the contrary, learners may display unwillingness to participate and still be selected by the teacher to be the next speakers. Previous studies showed that learners utilised gaze direction to display unwillingness to participate (Sert, 2013, 2015, 2019). For instance, Sert (2015) demonstrated how learners display their unwillingness to participate using gaze withdrawal or in some cases cover their faces to avoid mutual gaze. In spite of such avoidance of establishing mutual gaze by the learners, they were selected by the teacher to take the next turn (ibid). In a similar vein, Mortensen and Hazel (2011) pointed out that the establishment of mutual gaze between the teacher and the next speaker may not occur. Instead, the teachers may use the address terms (students' names) followed by requests without establishing mutual gaze in round robins. In these cases, the gaze of both parties can be directed towards the whiteboard or to the textbook where the relevant task requested to be done is located.

A common feature in the studies cited above is that it is always the teacher who allocates turns to the students. This is in agreement with McHoul's (1978) argument that it is only teachers who control turn allocation in the classroom. However, as noted, later researchers argued that turn allocation by the teacher is a jointly coordinated process as the learners can play a role in it by displaying willingness to participate and by bidding for the next turn

(Kääntä, 2012b; Mortensen, 2008; Sahlström, 2002). Learners' turns, however, are not always initiated following a teacher's nomination. Learners can self-select for the next turn without being nominated by the teacher. Nevertheless, learners' self-selection can entail some interactional work prior to launching the turn, which is relevant to the current study, as the first analysis chapter examines episodes where learners project self-selection for the next turn using the audio activation feature on Zoom. In the following section, a review of the literature on learners' self-selection is provided.

2.2.1.1 Learners' Self-selection in Classroom

Next speaker self-selection has been studied by a number of CA practitioners in both ordinary conversations (Sacks et al., 1974) and in institutional settings such as work meetings (Mondada, 2007) and classrooms (Ingram & Elliott, 2014; Jordan, 1990; Mortensen, 2009; Solem, 2016). According to Sacks et al. (1974), self-selection can be done using *pre-starts* or *turn-entry devices*, such as 'well', 'but', 'and', or 'so' (Sacks et al., 1974, p.719). Other than the verbal resources, such as recycling turn-beginnings (Schegloff, 1987, p.80) and overlaps (Jefferson, 1984), participants can employ non-verbal resources such as clearing their throats, gaze, facial expressions, and head and body movements (Schegloff, 1996). Such resources are used to project the possible completion of the current speaker's turn and self-selection to take the conversational floor next.

In the broader context, an influential study is Mondada (2007), which focused on the participants' multimodal practices for self-selection in work meetings while sitting around a table. Mondada examined the use of pointing gestures by the participants as a method to project the possible completion of the turn in progress and to project that they are the possible next speakers. Mondada tracked the sequential positions of the pointing practices for turn-taking in terms of their beginning and ending. Such practices can be initiated before the possible completion of the turn in progress or at the incipient speaker's turn. Such a sequential position reveals the participants' close monitoring of the turn in progress and such a public display of it makes it possible for the other participants to adjust their conduct accordingly. Moreover, Mondada's findings show how the incipient speakers display their engagement in the current speaker's turn using such pointing gestures. The participants managed to use the specificities of the setting to accomplish their self-selection (i.e. by pointing with their fingers or pens to the materials on the table where other participants were sat).

In classroom interaction, similar observations have been reported regarding the students' practices for self-selection. For instance, Jordan (1990) indicated that students' used the

oppositional turn entry-device '*pero*' ('but') in Spanish speaking classes to self-select for the next turn. Using such an entry-device projected the content of the coming turn, as it displayed a disagreement to the preceding turn. Moreover, these disagreements changed the course of the conversation and moved it in new directions.

Mortensen (2009) examined students' self-selection in L2 classrooms and the interactional work they undertake to claim incipient speakership and establish recipiency with a fellow participant before they fully initiate their turn. His study examined episodes of classroom interaction where the teacher's turn did not specify a next speaker and observed the students' displays of engagement and setting up the participation framework, even before the verbal initiation of the turn (i.e. in a pre-beginning position). The students in Mortensen's study used visual resources to claim incipient speakership and seek the others' displays of recipiency. Indeed, Mortensen demonstrated that the students used gaze, body movement, and in-breaths to claim incipient speakership as the teacher's turn was reaching a possible completion. Moreover, using these visual resources by the possible incipient speakers was recognised by the teachers, as they also displayed recipiency by gazing at the student.

Furthermore, learners' opportunities for self-selection can be enabled by the teachers. For instance, Waring (2009) examined an ESL learner's methods to self-select and depart from a series of successive IRF sequences. The self-selecting learner's-initiated turns were done "in a careful close coordination and cooperation with the teacher" (Waring, 2009, p. 796). Moreover, Waring illustrated that while working on a task, the teacher produced a successive series of IRFs and, once finished, the teacher's production of a question ">◦Does anybody◦-<" opened a space for students' self-selection. Another position for self-selection reported by Waring is right after a turn that had been produced by the same student. According to Waring, this can be explained in part as due to the preference for the next turn to be initiated by the last speaker.

Students' self-selection is sometimes referred to in the literature as the learner's initiative or the learner's unsolicited turn (Dolce & van Compernolle, 2020a; Duran & Sert, 2021; Petitjean, 2014; Solem, 2016; van Balen et al., 2022; Waring, 2011). Both terms describe the learners' uninvited contribution(s) to the talk in progress. According to Waring (2011), the description of 'uninvited' refers to the contribution of a learner who was not selected by the teacher to be the next speaker, or a learner who provides a response other than the one which they were selected to provide.

Learners' initiatives examined in the studies above conveyed several actions, such as requests for clarifications, asking questions, questioning the teacher's prior turn, or commenting on a turn that had preceded the initiative. These findings offer valuable insights for the current study as the learners self-select to produce similar actions, as will be demonstrated later in the Analysis Chapters. These initiatives are reported to have different degrees of legitimacy; in some cases, they are taken as and oriented to as legitimate turns. It is argued that acceptance and successful management of learners' initiatives can result in creating language learning opportunities (Sert, 2017; van Balen et al., 2022). In contrast, learners' initiatives can be treated as illegitimate and reprimanded by teachers (Petitjean, 2014).

This section discussed turn-taking in classroom interaction and the methods by which teachers allocate turns to students. Moreover, the section discussed students' self-selection and the practices by which they project the possible completion of the turn in progress and their incipient speakership. Also, the social actions conveyed by the learners' initiated turns and how teachers treat them were briefly presented. The next section presents the aspects of in-person classroom interaction that can cause trouble for the progressivity of class activities.

2.2.2 Managing Progressivity in Classroom Interaction

During interaction, turn-taking and sequence can be compromised by experiencing interactional troubles causing difficulty in hearing, speaking, or understanding, resulting in the temporary suspension of the ongoing activity by the participants to repair these trouble sources using the available resources. Repair is defined as the range of practices by which the participants locally manage problems in talk related to hearing, speaking or understanding (Schegloff et al., 1977). There are three basic components of a repair sequence, namely: the trouble source, repair initiation and the repair itself. When a trouble source is produced or occurs, this may be detected and indicated by the trouble source producer or by the recipient(s) (self vs other) and initiate repair using a wide range of practices both verbal and embodied. Repair initiation is then typically followed by the repair proper, either by self or by other(s), which could be a repetition of a misheard word, a rewording of the trouble source or replacing a lexical item with another (More details regarding repair as an interactional phenomenon can be found in Chapter 3). In classroom interaction, Sert (2015) defines interactional trouble as "the emergence of a temporary misalignment in the unfolding of an interactional and pedagogical activity, which is oriented to by the participants as such through verbal and nonverbal means" (p.58). This orientation by the participants to the trouble source

and the engagement in repair sequences is an orientation to their preference for progressivity (Schegloff, 1979).

Before moving to discuss studies that have examined interactional troubles and how the participants resolved them, a definition of what entails progressivity is provided. It can manifest in the successive production of turns; each turn follows what preceded it.

As Schegloff (2007) states:

> Moving from some element to a hearably-next-one with nothing intervening is the embodiment of, and the measure of, progressivity. Should something intervene between some element and what is hearable as a/the next one due – should something violate or interfere with their contiguity, whether next sound, next word, or next turn – it will be heard as qualifying the progressivity of the talk, and will be examined for its import, for what understanding should be accorded it. Each next element of such a progression can be inspected to find how it reaffirms the understanding-so-far of what has preceded, or favors one or more of the several such understandings that are being entertained, or how it requires reconfiguration of that understanding. (p.15)

This means that turns at talk are produced in relevance to what preceded them. The production of a turn makes conditionally relevant a next, and when that next is not produced, the participants orient to its absence. When a violation occurs, mutual understanding is disrupted, and therefore, the progression of interaction falters. This is then followed by repair by the participants to reestablish mutual understanding, and restore and advance the progressivity of the interaction. Thus, as the above quote states, progressivity is "moving from some element to a hearably-next-one with nothing intervening" (Schegloff, 2007, p.15).

In classroom context, Hırçın Çoban and Sert (2020) examined the interactional resources used by paired L2 learners to maintain progressivity in speaking assessment settings. During the speaking assessment activity, the two interactants faced interactional troubles that compromised the progressivity of their assessment activity. Hırçın Çoban and Sert (2020) pointed out that progressivity in their study is seen as "the resolution of interactional trouble and producing subsequent talk in an assessment setting" (p.68). The progression of the participants' talk is central to the assessment project, which without teachers, cannot assess test-takers. The current study does not examine test-takers' interaction; however, Hırçın Çoban and Sert's (2020) description of what entails progressivity is still applicable, as the participants in the current study need to handle interactional troubles caused by interference from disruptive noises and restore/maintain the progressivity of the ongoing class activities.

Progressivity of classroom interaction can experience minor or major disruptions. One of the issues that can cause these disruptions is related to the other's selection (i.e. by teachers) of

the next speaker (Lauzon & Berger, 2015). In their study of verbal nominations of the next speaker during a class activity, Lauzon and Berger provided examples where the teacher nominated a student who displayed unavailability to participate (i.e. by engaging in a parallel activity). In these examples, when the selected student complied with the nomination, he/she took the time to disengage from the parallel activity and get in the work mode (i.e. looks at the book on the table where the exercise was), which in turn did not necessarily lead to problematic consequences, but did cause a minor disruption to progressivity. Major disruptions to progressivity occur when the nominated student does not comply with the nomination, such as by producing a claim of insufficient knowledge. Following this occurrence, the participants needed to carry out extra work to pursue the progression of the class activity.

Lauzon & Berger (2015) demonstrated that the participants were actively working to address any interruption to the progressivity of the class activity. For instance, following a claim of insufficient knowledge, either another student proposed to replace their classmate, or the teacher kept the nominated student involved and that student made another attempt to answer the question addressed to them. Additionally, Lauzon and Berger indicated that both minor and major disruptions to the progressivity of classroom interaction were overcome by the collaborative work of the participants.

Similarly, Sert and Walsh (2013) reported on the interactional consequences of claims of insufficient knowledge. In their study, they focused on the interactional unfolding of students' claims of insufficient knowledge and how such claims were managed. Their findings revealed that such claims could result in the momentary halting of the progressivity of an ongoing class activity. Following the teacher's question, the learner may claim insufficient knowledge by using utterances, such as 'I don't know'. The teacher then may not present the correct answer in the next turn. Instead, the teacher pursues the activity progression by offering the answering slot to another learner, thereby co-constructing intersubjectivity.

Learners' unsolicited turns can also cause ruptures to the progressivity of classroom interaction. For example, Rodriguez and Wilstermann (2018) examined the interactional consequences of Spanish as a foreign language learner's initiative in teacher-student interaction in a round robin. In one of the examples they presented, a learner self-selected to ask a question related to a previously discussed topic, thereby halting the progressivity of the current topic. The teacher responded by providing an answer to the learner's question. Following this, the learner self-selected again and re-established joint attention to the new

topic by asking a question related to it. This example demonstrates that, even though the learner halted the progressivity of the new topic, they resumed it, which is a task usually undertaken by the teacher.

A common theme in the above studies is the preference for progressivity displayed by both teachers and students in classroom interaction. Such preference is viewed as endemic to how conversations are organised at both levels of turn-construction and the structure of sequences (Sacks, 1987/2020; Schegloff, 1979, 2007). There is always a preference for the successiveness of turns and for producing the next turn in interaction. In Lauzon and Berger (2015), for instance, the collaborative work of the insufficient knowledge claim producer's and their attempts to provide the correct answer contributed to the resumption of the classroom activity's progressivity. Also, other learners stepped in to take the answering slot following their classmate's claim of no knowledge, thereby contributing to the progression of the talk. Similarly, offering another learner the answering slot instead of providing the correct answer following a claim of insufficient knowledge by the selected learner is an orientation towards the preference for progressivity.

The studies cited above in this section have demonstrated that claims of insufficient knowledge could halt the progressivity of classroom interaction. However, it is worth noting that these claims are only one example of aspects that can cause interruptions to the progressivity of interaction in classroom. Other examples, such as delayed responses to questions (Amar, Nanbu, & Greer, 2022), a teacher's lack of knowledge pertaining to a learner's personal experiences (Batlle & Deal, 2021), word searches (Pekarek Doehler & Berger, 2019), and displays of understanding/misunderstanding (Antaki, 2012) can also interrupt the progressivity of classroom interaction.

On some occasions in classroom interaction, as Hosoda and Aline (2013) note, other types of preferences may momentarily suspend this preference for progressivity. Research in teacher-fronted classes has indicated that preference for progressivity might be contravened as the teachers halt the progressivity of the classroom interaction to address what is oriented to as a breach to the classroom order. In what follows, some studies discussing this issue are presented.

2.2.2.1 Orienting to a Behaviour as Inappropriate in Classroom Interaction

Relevant to the current study is literature on one of the important aspects of classroom interaction: the management of inappropriate behaviour (Hazel & Mortensen, 2017; Jakonen, 2016; Klattenberg, 2021; Macbeth, 1990; Margutti, 2011). The relevance stems from how teachers and learners identify and flag a certain behaviour as a trouble source or as transgressive to the moral order of the classroom, and in response momentarily halt the progressivity to address that behaviour before resuming the suspended activity.

Klattenberg (2021) examined teachers' question-formatted reproaches to parallel activities initiated by the students which ran alongside the teacher's programme of actions in German EFL classrooms. Examples of parallel activities in Klattenberg's data are talking and laughing with a fellow student. The teacher's programmes of action refer, for example, to when the teachers address a question to one of the students. The programme, in this case, is the rest of the class listening to the dyadic interaction between the teacher and the addressed student. Klatternberg's findings show that these question-formatted reproaches had two results. The first is that the student(s) ceased their parallel activity (i.e. they stopped talking). The teacher's question, in this case, is met by ceasing the parallel activity, which means that it is understood by the students as a reproach. The mutual epistemic access between the teacher and the students to the parallel activity led to stopping it following the teacher's question instead of providing an answer (Klattenberg, 2021). The second result showed that teachers designed their questions to both highlight the inappropriate behaviour and encourage the students to realign with the normative behaviour at that particular moment (i.e. work on the task). These reproaches held the students accountable for not adhering to behavioural expectations.

Similarly, Hazel and Mortensen (2017) examined the orientations of participants in an L2 classroom to some or other forms of transgression to classroom conduct. In an example of transgression related to speaker selection, the students were engaged in an activity where the turn allocation followed a round robin organisation. This means that each student was taking a turn to make a contribution and pass the turn to the next student. As the teacher addressed a question to a student, another student self-selected to answer and stepped on the turn assigned to her classmate. The teacher oriented to this as transgression by leaning toward the student and asking her 'if her name is the selected student's name?'. Similar to the questions in Klattenberg's study (2021), the question does not seek an answer; it is an admonishment. The transgressing student recognised this as such, as in the next action what she produced was not

an answer but an apology. Overall, Hazel and Mortensen (2017) and Klattenberg (2021) studies revealed important findings on the moral order of the classroom (i.e. what is acceptable and what is not), as well as the design of admonishment and recognition of it by the 'offending' student.

Margutti's (2011) analysis of conditional-formatted reproaches in video-recorded sixth and seventh grade classes revealed interesting insights into the design of the teacher's reproaches and their relevance to dispreference. In the conditional-formatted reproaches in Margutti's data, teachers either held the inappropriate behaviour accountable but the student doing it was unaware, which then became a subject for instruction; otherwise, the misconduct was deemed deserving of punishment and was, therefore, reproachable. Margutti highlighted the resistance displayed by the teacher to describing the students' misconduct in an explicit manner. There is always a tendency to design reproaches in an indirect way (i.e. when teachers do not explicitly articulate the problem). In one example Margutti provided, the explicit articulation of the problem only came after nine attempts to obtain the students' attention at the beginning of the class, when the teacher was trying to get the class ready for the lesson. When the students did not seem to share the epistemic access to what the problem was, the teacher oriented explicitly to the students not paying attention as an appropriate behaviour.

Interesting observations from these studies are related to the design of the turns in which participants orient to a certain behaviour as inappropriate. One feature is the indirect flagging of inappropriate behaviour. For example, in exemplifying the orientation to speaking out of turn as a transgressive behaviour, Hazel and Mortensen (2017) reported that the teacher smiled and asked a female student, 'is your name Andre?' to indicate that it was Andre's turn to speak. Similarly, the teachers in the studies by Klattenberg (2021) and Margutti (2011) designed their reproaches with the assumption that the student producing the reproachable behaviour had the same knowledge that their behaviour was unacceptable. Teachers can also design their turns not to flag the unacceptability of a certain behaviour, but to direct to students as to how to align to what is expected from them in such a situation (Klattenberg, 2021). Moreover, the sequential position of producing an explicit description of the misconduct (i.e. delaying it) can be seen as producing the reproach as a dispreferred action (Margutti, 2011).

Another common observation is the placement of the responsibility for ending the misconduct. Participants' orientations to transgression seem to impose an obligation on the producer of the misconduct to cease it. In the examples mentioned above by Hazel and

Mortensen (2017) and Klattenberg (2021), the teacher's orientations were designed in a way that resulted in the termination of the misconduct by the transgressors. The teachers' turns were designed in a way to indicate that this transgression belonged to certain participants and a cessation of it was relevant next. This is in a way similar to the case of some practices in other-repair initiations, as they can indicate whom the trouble source belongs to (Robinson, 2006). Carrying out self-repair by the indicated participant then becomes conditionally relevant when an orientation to a trouble source has been initiated. Not conducting the repair can then trigger the production of another or multiple repair initiations.

Hitherto the discussion has aimed at presenting some aspects of classroom interaction that are relevant to the current study, namely turn-taking and aspects in interaction that can halt the progressivity of classroom interaction and methods by which they were flagged and addressed. The current study examines the participants' use of the audio activation/deactivation features of the videoconferencing platform to manage turn-taking and overcome interactional troubles caused by interference of background noises. It is worth noting that although the current study investigates interaction in L2 video-mediated settings, the research findings from in-person classrooms are valuable and can provide useful insights. The following sections present the online context in the current study, the affordances approach, and fractured ecologies. Moreover, the organisation of turn-taking in video-mediated interaction and what can be problematic to the progressivity of interaction in such settings will be presented.

2.3 Computer-Mediated Communication

This section offers an overview of computer-mediated communication (CMC) and its different modes and forms leading to videoconferencing. CMC has been defined as "the use of computer systems and networks for the transfer, storage and retrieval of information among humans…the computer /network system is primarily a mediator rather than a processor" (Berge & Collins, 1995, p.11). Similarly, Herring (1996) described CMC as "communication that takes place between human beings via the instrumentality of computers" (p.1). This can be used as an umbrella term that includes the social interaction mediated by communication technologies (Meier & Reinecke, 2020). Initially, CMC was textual in the main, but then texts were supported by graphics, audio and video for the purpose of communication, and became available not only on stationary computers but on websites and applications on a wide range of devices, such as tablets and mobile phones (Herring, 2019). CMC can occur either

synchronously or asynchronously. One way to differentiate between these two modes is by the simultaneity of interaction.

In asynchronous communication, the participants do not engage in simultaneous or real time communication, instead using email exchanges, posting threads on learning management systems, discussion forums, web blogs or other social media platforms, where there is usually longer lags (Graham, 2006). In this mode of online communication, it is not necessary for the participants to be present at the same time (Amiti, 2020). Additionally, in asynchronous communication the participants are able to access the content or other people's contributions and interact with them at their convenience (Chun, 1994). Contrastingly, synchronous communication occurs when the participants meet at a specific time on an online platform to work on certain activities or lesson(s) (Ajabshir, 2019). This means that they communicate or talk to each other in a more direct manner. This can occur either in written/audio chats, or video calls.

Another way to differentiate between the two modes is the requirement of joint attention. O'Rourke and Stickler (2017) defined synchronous communication as the "dialogic communication that proceeds under conditions of simultaneous presence (co-presence) in a shared communicative space, which may be physical or virtual" (p.2). This definition is sufficiently general to include in-person talk, text chats, audio, and video calls. They argue that both synchronous and asynchronous modes share the feature that participants' conversational contributions are produced in response to their interlocutors. However, a distinction can be made in relevance to the mutual attention by the users to the unfolding meaning, which is a property that is evident in synchronous communication. In contrast, the asynchronous mode of communication seems to lack such a feature (O'Rourke & Stickler, 2017). The context examined in the current study is a synchronous video-mediated L2 speaking classes. Videoconferencing is considered as a mode of Computer-Mediated Spoken Interaction (CMSI) in addition to audio chats (Jenks, 2014).

2.3.1 Computer-Mediated Spoken Interaction (CMSI)

Interaction management in CMSI has been examined from different methodological stances. For instance, multimodal (inter)action analysis was employed to examine how instructions were multimodally constructed and delivered in a video-mediated language learning setting (Satar & Wigham, 2020, 2023; Wigham & Satar, 2021). Another methodological stance - social semiotics – was used to examine participants switching between writing and speech, gaze management, and difficulties in establishing eye contact (Sindoni, 2014). In addition,

CA was also used to examine overlapping talk in multi-party audio chats (Jenks, 2009); aspects of trouble in audio chats (Brandt & Jenks, 2013); and latency and turn-taking management (Seuren, Wherton, Greenhalgh, & Shaw, 2021). The current study employs an ethnomethodological conversation analysis methodological stance (a detailed description of which is provided in Chapter 3).

A fair amount of interest by researchers who have examined CMC in language learning settings using CA was directed towards text-based encounters, as they are a common form of communication (González-Lloret, 2008, 2011; Negretti, 1999; Tsai & Kinginger, 2014; Tudini, 2002). However, the current study examines the participants' use of audio activation/deactivation features in a video-mediated interaction. This also provides relevance to the literature of studies that have examined interaction in audio chats. The following section mainly addresses video-mediated interaction, which is the context for the current study. The research findings from audio-based interactions, however, will be reported when relevant.

2.3.2 Video-Mediated Interaction

This section presents a summary of the EMCA research on video-mediated interaction. It begins with a brief history of the development of video-mediating technology, then continues with a discussion of the contexts examined by earlier researchers, and follows with a discussion of the significant findings relevant to the present study. In addition, this section demonstrates how previous EMCA researchers have addressed the study of turn-taking and repair in video-mediated contexts.

Video-mediated interaction can be defined as "interaction conducted in and through a specific type of technology (e.g. Skype, Teams, Zoom and the like) that enables synchronous communication via a video link" (Due & Licoppe, 2021, p. 2). It was only in the early 2000s that the ability to connect through video gained in popularity and started to be commonly used. The development of mobile phones and other devices, the widespread availability of faster, cheaper, and a more reliable internet connection, and the development of applications, such as Skype, FaceTime, Zoom, and others, contributed to the widespread utilisation of video communication as is found nowadays (Harper, Watson, & Licoppe, 2019). This was partly driven by the need for this form of communication by corporations, family, and friends located in different geographical spots (ibid). Recently, an important factor that has played a significant role, not only in somehow forcing a transition to video-mediated settings, but also in accelerating the development of such technologies, is the COVID-19 pandemic. The Zoom

technical support page showed a long list of updates and new functions added to the platform almost every week during the COVID-19 pandemic. It is safe to say that even though video mediating technology has been in use for nearly a century, it was only available to a small number of users until quite recently. Indeed, as noted above, its widespread availability and use took off only in the last 10 to 15 years (Mlynář et al., 2018).

EMCA researchers have always shown an interest in technology-mediated interaction, including written chats, audio chats, and video-mediated interaction. According to Mlynář et al. (2018), the vast majority of EMCA studies on video-mediated interaction have been undertaken from 2010 onwards due to the widespread integration of video cameras into daily used devices, such as laptops and mobile devices. This, and the transition to online venues following the COVID-19 lockdowns, has played a role in making data available for more researchers to investigate interaction on platforms that support both synchronous and asynchronous communication.

Researchers using EMCA have investigated video-mediated interaction in a variety of everyday, institutional, experimental or quasi-experimental settings. For instance, in mundane settings, Licoppe and Morel (2014) analysed the organisation of showing sequences (i.e. when a participant showed an object in their local environment). Their findings stated that launching the showing sequences is preceded by some prefatory work, such as announcements or requests by the potential recipients of that showing. Such prefatory work allows the participant to deviate from the standard image in which their face appears on the screen. The authors claim that on other occasions, such deviation from the 'talking heads configuration' can be noticeable and accountable by fellow participants (Licoppe & Morel, 2014). Other researchers have carried out studies in experimental or semi-experimental settings. For example, Luff et al. (2003, 2013, 2016) examined referential activities, such as pointing to an object in the local environment for a remote participant in video-mediated interaction. Their findings on what they termed 'fractured ecologies' and how this may have an impact on the recognition of gestures, have been influential in the study of the organisation of interaction in video-mediated interaction. The findings from their work are relevant to the current study in terms of how the participants orient to the impact of the setting on their conduct (more detail on fractured ecologies is found in Section 2.3.2.1).

In institutional settings, many EMCA scholars have shown an interest in examining interaction in healthcare settings by looking at video-mediated consultations, follow-ups, or meetings between doctors (Mondada, 2015; Seuren et al., 2021; Stommel & Stommel, 2021).

Moreover, many scholars have documented classroom video-mediated interaction; for instance, Hjulstad (2016) analysed participants' practices to organise the built space in sign language classes. Balaman and Sert (2017) tracked the development of participants' interactional resources to accomplish collaborative tasks; Badem-Korkmaz and Balaman (2022) examined how the teacher pursued students' responses; and Malabarba et al. (2022) examined multimodal practices to resolve overlapping talk in tutor-tutee classroom interaction. Even though the field is growing, there is a need for more studies to understand the interactants' practices to manage interaction in different settings.

Research using EMCA on video-mediated interaction in the aforementioned contexts has generated valuable findings, which have provided significant insights into the different practices people use to interact in different settings (Due & Licoppe, 2021; Mlynář et al., 2018). The current study is interested in examining the participants' use of the video communication platform's affordances to interact with each other in language teaching interaction. More specifically, this study is interested in examining the participants' use of audio activation/deactivation features to manage their turn-taking and deal with aspects of trouble that threaten the progressivity of their interaction on Zoom platform. Following this research focus, the literature in the remainder of the section will be curated in relation to the themes of affordances, fractured ecologies, the organisation of turn-taking, aspects that can interrupt the progressivity of interaction in video-mediated settings, and methods for flagging and repairing them.

In what follows, the two concepts of affordances and fractured ecologies will be presented before what previous researchers have reported regarding the organisation of turn-taking and repair in video-mediated settings is presented.

2.3.2.1 Affordances

The current study draws on the concept of affordances as an approach to examine the relationship between technology (i.e. Zoom video-mediation in this case) and the organisation of social interaction in L2 classrooms. To this end, it is essential to highlight the participants' orientations to the reflexive relationship between the organisation of their social interaction and the features of the technology that mediates it (Arminen et al., 2016; Hutchby, 2001b). Hence, it is important at this point to define what is meant by affordances in the current study. Initially, affordances as a concept was first coined by Gibson (1979) in his work on the psychology of perception. From a realist and anti-dualistic perspective, affordances for Gibson, refers to the notion that objects in the environment can afford an animal the

possibilities for action. In his study, Gibson claims that humans, animals, and insects perceive objects in their surrounding environments in relation to the actions these objects afford them to accomplish. For some insects, the water surface can be perceived as a walkable space, while it is not perceived as such by a larger animal, such as a tiger, as it does not afford them the same possibility. This, subsequently, means that the possibilities for the action to be afforded by an object differ according to species. Gibson (1979) indicates that there is a relationship between the object's physical qualities and social norms and rules. The concept of affordances refers to the idea that objects have their own properties, but these properties emerge as actors interact with the objects (ibid).

Drawing on Gibson's notion of affordances, Hutchby (2001b, 2001a) contributed to the work on the concept by using it to analyse technology-mediated social interaction using a conversation analytic perspective. Hutchby (2001b) sought to develop a stance that stands in-between determinism and constructivism (i.e. technology vs social). The interest is neither solely in technology nor in conversation, but in both, where it becomes evident that affordances of technology play a role in the structure of turns of talk and the actions they accomplish (ibid). Hutchby's stance calls for a 'change in empirical footing', which suggests that determining the affordances of technology is best examined through the participants' perspective as they engage in technology-mediated interaction. This contrasts with other views in sociology that suggest it is the job of sociologists to determine what is an accurate description of an affordance and what is a misrepresentation (Hutchby, 2003b).

Hutchby (2001a) argues that there is a "complex interplay between the normative structure of conversational interaction and the communicative affordances offered by different forms of technology" (p.13). By normative structures, Hutchby refers to those underpinning the production of talk in social interaction in any given situation. As his work draws heavily on CA, such normative structures are those oriented to by the participants, which as a result make them available to analysts as well. For instance, the formal opening of a meeting mediated by video is either preceded by informal pre-openings where the participants check the functionality of their audio and video, or in some cases skip it (Licoppe & Dumoulin, 2010). Hutchby (2014) describes these communicative affordances as both "functional in that they enable (and also constrain) the engagement in some activity; they shape the conditions of possibility associated with an action. Relational, in that, they may differ for one object in different contexts, or between different species" (p. 87). This demonstrates an agreement between Gibson and Hutchby on the relational dimension of the affordances, as they can differ according to species and context. However, Hutchby (2003) calls for the focus to be

extended, as well to examine the manifestation of enablement or constraints of a technological artifact, such as video-mediation in the production of such normative structures of conversation. Such enablement and constraints manifest in the participants' orientations to them in their interaction.

The focus in Hutchby's work is directed towards the 'communicative affordances', which describe how the features of technology shape social interaction as it occurs in situ and how they can facilitate or hinder the participants' perception of the ongoing interaction. The notion of communicative affordances (Hutchby, 2001a) is of special interest to the current study, as its main interest revolves around an investigation of the participants' use of the technology features at hand in the organisation of their turn-taking and repair in video-mediated L2 classroom interaction. Moreover, Hutchby (2001) differentiates between what the developers design certain technological features for (design features), and how users actually utilise these features (features in use) (p.129). This means that participants can use technologies beyond their intended use by the developers, which then adds to the functionalities and meanings associated with these features. The participants in video-mediated interaction can create novel practices to accomplish their social actions using the affordances of the platform to overcome the limitations posed by a lack of physical co-presence. In video-mediated sign language, for example, Hjulstad (2016) reported on teachers' use of establishing a spatial location of the learners according to where they appear on his screen by using the learner's name in addition to pointing to their location on the screen. By doing this, the learners could recognise which pointing direction was addressed to them and respond accordingly.

2.3.2.2 Fractured Ecologies

In video-mediated settings, the interactants are located in different physical spaces. Such a condition is described as fractured ecologies, which means that the participants have limited or asymmetrical access to each other's immediate environments (Luff et al., 2003). Their interaction is, therefore, "fractured from the environment in which it is produced and from the environment in which it is received" (Luff et al., 2003, p.55). The fractuturedness of these ecologies is found to pose constraints on the participants' access to resources, such as pointing gestures and visual conduct in interaction (ibid). Being located in different physical spaces can place some constraints on the recognition of such conduct (Luff et al., 2003). The participants may, however, modify their interaction in accordance with the affordances of the technology in use to overcome such constraints (Due & Licoppe, 2020; Luff et al., 2016). This sub-section shows how the lack of physical co-presence can affect video-mediated

interaction and the methods by which participants manage such effects and maintain intersubjectivity.

Previous studies indicated that interaction in fractured ecologies can pose challenges to the participants' ability to reach intersubjectivity and showed how participants modified their conduct and used the technology affordances to manage such challenges and were able to ensure intersubjectivity in video-mediated interaction (Due et al., 2019; Ilomäki et al., 2021; Licoppe, 2017; Licoppe & Morel, 2014; Luff et al., 2016; Melander Bowden & Svahn, 2020). For instance, in their study of interaction between a tutor and a tutee, Melander Bowden & Svahn (2020) showed how the tutor and tutee used referential practices, such as pointing, but to ensure intersubjectivity such practices were mainly conducted using the online cursor to highlight written or drawn parts on a shared worksheet and to highlight some functions of the platform.

Similarly, in their examination of showing sequences in mundane video-mediated interaction, Licoppe and Morel (2014) pointed out that camera mobility by the participants played a role in ensuring intersubjectivity. To compensate for the lack of physical co-presence, participants moved the camera around their physical spaces to show objects to their fellow participants. Following the announcement of an upcoming showing, the shower used camera movement in combination with speech in order to make images that are understandable and relevant to the viewer. The viewers produced displays of understanding either via talk or in combination with facial expressions which are treated as such by the showers.

Another challenge to intersubjectivity resulting from interacting in fractured ecologies manifested in remote instruction-giving as demonstrated by Due et al. (2019). Their study examined how professionals gave instructions to clients regarding handling an object in their physical proximity (i.e., a printer in their case). As referential practices can be challenging to recognise since participants are located in different physical locations, professionals used what the authors termed "mimicable embodied demonstrations" (Due et al., 2019, p. 13). They found out that participants broke down their instructions into smaller simplified mimicable steps to overcome the complexity of instruction-giving in video-mediated interaction. The practice of decomposing instructions enabled both instructors and instructees to maintain intersubjectivity despite the challenges in recognising referential practices in such a setting.

Similarly, in a recent study of teachers' elicitations of students' answers in a video-mediated language learning class, Badem-Korkmaz and Balaman (2022) demonstrated teachers'

successful use of the medium's affordances. Following the teacher's multiple attempts to elicit a response from the students, the teacher drew on the learners' multimodal actions by monitoring the list of learners' video frames on the Zoom interface, where their displays of willingness to answer are publicly available. By monitoring the video frames, the teacher was able to identify a potential next speaker, and the student's multimodal actions within the video frame enabled him to be selected for the answer slot. Despite being geographically dispersed, the authors revealed that monitoring other participants' conduct is still at least partially possible. Accordingly, the student's display of their willingness to participate was recognised, and thus, the next turn was allocated by the teacher to that student.

While connecting from different physical places, an intersection between private and public activities can occur. Private activities refer to those either undertaken by the participants themselves or by others in their physical proximity; usually, participants keep them hidden. On the other hand, public activities refer to joint activities between the participants on the communication platform (Rosenbaun et al., 2016; Ruhleder & Jordan, 2001; Tutt, Hindmarsh, Shaukat, & Fraser, 2007). According to Ruhleder and Jordan (2001), participants in video-mediated interaction often engage in side or parallel activities, which may be designed to remain hidden from the other participants. However, as will be demonstrated later in the current study, some of the parallel activities can intersect with the ongoing classroom activities and may be topicalised by the other participants. Previous researchers have called for the need to consider this intersection between private and public activities, as they can reshape the ongoing interaction between the participants (Rosenbaun et al., 2016; Tutt et al., 2007).

However, not all private activities are designed to remain hidden. Participants are sometimes found to make their screen activities visible, as such activities are inaccessible to their fellow participants. Balaman and Doehler (2022) revealed that the participants in their data employed practices - such as 'let me/let's x' - to account for an upcoming suspension of the ongoing activity to carry out some screen activities. By doing this, they made some aspects of their screen activities available to their fellow interactants, in turn making such suspensions of talk task-related and not a threat to the progression of task work.

Interaction in fractured ecologies results in a lack of access to what others see or hear. This can be oriented to by the participants as they question their (or objects in their local spaces) visibility or audibility. Fornel (1994) indicated that it is common to see participants regularly check if they are seen or heard by their co-participants. Similarly, participants may ask each

other to position themselves in a specific spot to ensure their visibility (Mondada, 2015). Additionally, participants may begin their meetings by ensuring that their video and audio are in working order (ibid). In telepresence robot-mediated interaction, the participants used visibility checks, such as "can you see?" to ensure that the remote participant could see the whiteboard in the classroom (Jakonen & Jauni, 2021). This lack of mutual access to what is being heard or seen by others can cause trouble to the ongoing interaction, as will be shown in the analysis chapters.

To sum up, interacting in fractured ecologies may constrain the participants' conduct in such environments. The participants do not have access to what others hear or see as they are geographically dispersed. However, as noted by Due and Licoppe (2020), they can overcome such constraints by either fitting their conduct into the available affordances of the medium or adjust their conduct by changing some aspects of the setting. This is relevant to the current study, as it shows the participants' use of technology affordances to overcome the constraints faced while interacting on Zoom.

Having presented the concepts of affordances and fractured ecologies, the next section presents an overview of what is reported in the literature regarding the study of the organisation of turn-taking and what aspects can cause trouble to the progressivity of interaction in L2 video-mediated classroom interaction.

2.3.3 Video-mediated L2 Interaction

A number of researchers have explored interaction in video-mediated settings using EMCA. Previous research falls into two main categories: descriptive studies and developmental studies (González-Lloret, 2015). In the former, the focus was on investigating the interactional structure of the technology (i.e. describing sequence, turn-taking, and repair). Moreover, these studies investigated how the medium's affordances may impact the structure of the interaction. Research in developmental studies investigates the tracking of how social practices develop over a certain period of time. González-Lloret (2015) indicated that the majority of studies that used CA to study CMSI fall into the first category; this is also the case for the current study. Thus, the following sections review the literature on the organisation of turn-taking and repair in audio/video-mediated L2 learning and teaching settings.

2.3.3.1 Turn-taking in L2 Video-mediated Interaction

The study of turn-taking organisation in video-mediated L2 learning and teaching contexts has received a great deal of scholarly attention. Turn-taking in video-mediated contexts can be different from face-to-face interaction due to the lack of co-presence. Indeed, it is argued that it is a unique turn-taking system (Arminen et al., 2016; Licoppe & Morel, 2018). In addition, turn-taking in video-mediated interaction is described to be a challenging task for both speakers of first and second languages (Payne, 2020). Researchers who have examined the organisation of turn-taking in L2 video-mediated settings have reported valuable findings on a set of topics, including management overlapping talk, turn-allocation, displaying recipiency, the use of medium affordances to manage turn-taking and technical issues that may impact the turn-taking system.

Managing overlapping talk in audio/video-mediated L2 learning contexts was examined by previous researchers (Jenks, 2009a; Malabarba et al., 2022; Stone & Brinham, 2022). For example, Jenks (2009a) examined the management of overlapping talk in multiparty learner-learner audio-based chats. His analysis shows that the participants employed pauses strategically to manage contiguous or simultaneous talk. The pauses were significant as they were strategically placed immediately after the contiguous/simultaneous utterance, which then momentarily opened the floor for the fellow participants. According to Jenks (2009a), this strategic use of the pauses is an orientation to the norm, stating that only one participant can speak at a particular moment. Moreover, Jenks pointed out that turn-beginning overlaps draw attention to a difficulty facing participants in audio-based chats, namely the lack of non-verbal cues as the participants need to time and project their biddings for the conversational floor. Jenks emphasised the importance of this basic skill for learners in audio chats, as they need to launch their turns in a well-timed manner in such environments where they lack access to the embodied displays of others.

Timing of contributions for participants in video-mediated classes seems also to be an important skill. Andrews (2020) examined the challenges of teaching and learning in large classes on Zoom in American universities and K12 classes. One of the challenges the author reported is related to the 'muted learners' ability to join the ongoing conversation. Andrews indicated that it is difficult, especially when there is a connection lag, for learners to know where to join the ongoing conversation without the risk of overlapping with others' talk. Being muted is another layer of difficulty, which is presented by the time that unmuting takes from the learners, which can form a sort of deterrent to participation (ibid).

Overlapping talk is not exclusive to multiparty interaction, as it is also reported to occur in dyadic video-mediated L2 interaction. The methods by which overlaps are resolved, however, can be different, due to the ability to use non-verbal conduct. Malabarba et al. (2022) investigated the multimodal resolutions of overlapping start-ups in video-mediated L2 one-to-one tutoring sessions. They reported that the teacher used a set of practices, such as the termination of verbal talk, lip pressing, the 'go ahead' utterance, or both, to display the withdrawal of competing for the floor and securing the recipient mode. Such a combination of both verbal and non-verbal practices is referred to by the authors as 'enhanced explicitness'. Other cases in their data, though, show that the overlap resolution can be simply done by the termination of talk accompanied by nodding or a smile. Compared to the findings by Jenks (2009a), the participants in Malabarba et al. (2022) had access to more embodied resources afforded by Zoom, which are not available to participants in audio-based chats. Regardless of this availability, turn-beginning overlaps still occurred.

Another aspect examined in the literature on turn-taking organisation in both teacher-learner and learner-learner video-mediated L2 settings is the turn allocation and eliciting learners' responses practices (Badem-Korkmaz & Balaman, 2022; Guo & Zhang, 2021). Teachers' practices for turn allocation in video-mediated classes are reported by Guo and Zhang (2021) to be different from those in in-person classes. They examined teacher-initiated turns in both in-person and video-mediated German as a second language classes. They compared the practices by which teachers allocated turns in both contexts; either teachers addressed questions to the whole class and students bid to answer, or teachers allocated the next turn to a specific student in both contexts. The number of instances where the teacher addressed the whole class was significantly lower in video-mediated classes than in in-person classes. The teachers' nomination of the next speaker was found to occur more in video-mediated classes. According to the authors, this is because the mediated context reduces the availability of multimodal and embodied resources, which impacts the participants' ability to achieve mutual orientation and coordination of actions. Also, in the video-mediated classes, the nominated students for next speakership had to verbally confirm their status as the current speaker, rendering the organisation of turn-taking in such a context unsmooth and fragile (ibid). It is worth noting that the students had their cameras off during the class, which according to the authors disabled the teachers' ability to scan the students' video frames to check their willingness to be the next speaker.

Scanning the learners' video frames to check for displays of willingness to take the next turn is evident in Badem-Korkmaz and Balaman's (2022) study of teacher's practices to pursue

responses from the students after addressing a question to the whole class. Such questions do not specify who the next speaker is; rather, they indicate an open floor for the learners to bid for. The lack of students' responses then led the teacher to pursue responses. One of the practices reported is the teacher's monitoring of learners' video frames and identifying a potential next speaker based on their embodied action. This in turn shows the successful use of the medium's affordances to identify the next speaker and maintain the progressivity of the class activity.

Relevant to the multimodal displays of recipiency and willingness to take the conversational floor next is Stone and Brinham's (2022) examination of Japanese students' organisation of turn-taking in English Communication course video-mediated discussions in Zoom breakout rooms without the teacher's presence. They indicated that the participants displayed recipiency using embodied resources. The participants in their data positioned their bodies and gaze towards the screens during the turn in progress to display their recipiency.

This section has briefly surveyed studies that have investigated the organisation of turn-taking in L2 teaching and learning video-mediated interaction. The literature revealed that there was a great deal of attention directed towards factors that affect the organisation of turn-taking, such as delays in transmitting the participants' actions to each other. Such an effect was mainly measured in dyadic interaction, both in the classroom and in other contexts. Other studies have investigated how participants displayed recipiency and how they managed instances of overlaps and mistimed turn beginnings. The study of turn-taking in video-mediated interaction is intertwined to an extent with the examination of aspects of trouble that can hinder the progressivity of the interaction. The following section presents some other examples of trouble sources that can halt the progressivity of interaction in L2 video-mediated settings and how participants addressed them, as reported in the literature.

2.3.3.2 Repair and Trouble Sources in L2 Video-Mediated Interaction

This section reviews the literature on the organisation of repair in video-mediated interaction, which begins by showing how repair in video-mediated interaction can differ from face-to-face interaction in terms of what is considered trouble and how the technology affords or constrains the participants' abilities to carry out repair. The section also surveys the findings of research on aspects of trouble in L2 learning and teaching video-mediated settings.

Repair in video-mediated interaction can differ in a number of ways from that in in-person interaction. First, researchers have reported that participants may not be aware that they are

producing the trouble source. The lack of access to what others hear or see due to connecting from different physical spaces can result in the participants not knowing how their turn is received by their fellow participants and, thus, not knowing that they have produced a trouble source. For instance, Rintel's (2013, 2015) examination of distortions in video-mediated conversations suggests that the distorted turn producers did not know about the distortion. They could only have known if they had been alerted by the recipient. Otherwise, the absence of a proper response may alert the trouble source producers that their turn may not have fully arrived to the recipient. As a result, the distorted turn producer carries out self-repair. Hence, other-repair was more evident due to the lack of access to how turns are received by the other participant. Also, to the speakers, their turn was produced without any problems but arrived distorted to the conversational partner. This may have had an impact on the repair trajectories, as the recipient of the distorted turn was "materially afforded the first access to the agenda of repair, while speakers are materially constrained to a position of second response" (Rintel, 2013, p.9). This may explain the prevalence of other-repair in the majority of the cases in Rintel's data.

Another characteristic that may be unique to repair in video-mediated interaction is that in some cases the interactants may face difficulty locating the trouble and, thus, become unable to repair it as "its origin is obscured" (Ruhleder & Jordan, 2001b, p.132). Ruhleder and Jordan (2001a) indicated that transmission delays troubled the turn-taking system in videoconference meetings, but the participants were unable to identify what was causing this trouble, although they knew that something was not right. Therefore, interacting in fractured ecologies may impair the participants' ability to locate and repair the trouble source (ibid).

There were a number of findings on aspects of trouble that can disrupt the progressivity of interaction in both audio and video-mediated interaction, such as delays of transmission or (latency), video/audio distortions, and background noises (Brandt & Jenks, 2013; Jenks, 2014; Rusk & Pörn, 2019). These are all variables that can be consequential to the ongoing interaction in such settings. As noted, the findings from previous research on online audio chats can also be relevant to the current study, as it focuses on the participants' use of audio activation/deactivation features to manage trouble originating from disruptive noises. As one of the objectives of the current study is to examine the participants' use of audio activation/deactivation features to handle trouble caused by disruptive noises, the remainder of this section is dedicated to reviewing the literature on the interference of such noises.

Interference of background noises is reported to be treated as trouble by the participants in both audio and video-mediated interaction (Alzaidi, 2016; Cheung, 2021; De Fornel, 1996; Jenks, 2014). For instance, De Fornel (1996) investigated participants' strategies to cope with a video mediating device for the creation of what he described as an 'adequate interactional frame' in multiparty interaction (p.47). He argues that the participants are responsible for controlling their immediate environment and for maintaining some kind of boundaries between their physical spaces and the shared space with the coparticipants. He demonstrates this by providing an example, where the current speaker produced hesitation markers and leaned towards the video device in an orientation to a noise coming from the background of one of the participants. The recipient recognises this noticing and shuts her door to prevent the noise outside her physical space from disrupting the ongoing interaction. The example also shows the verbal accounting for closing the door, as the person uttered "wait: i'm closing the door because::" (p. 53). This is a good example because it shows that participants in video mediate interaction hold themselves and others accountable for trouble sources that they do not personally produce, as their occurrences can threaten the progressivity of ongoing interaction. The noise in this case was not produced by the participant, but was due to something happening outside the room; however, it was treated as trouble and as his/her responsibility to handle it.

Background noises can be consequential to the progressivity of the ongoing conversation and participants may treat it as a trouble source (Jenks, 2014). Even though Jenks' (2014) data comes from interaction in online L2 voice-based chat rooms, its results still provide valuable insights into the consequentiality of background noises and their interactional relevance. Jenks distinguishes the different kinds of background noises, pointing out that they can be 'white noises' that cover verbal interactions, such as typing sounds, or can be 'ambient noises' resulting from emphatic typing noises, people, pets, cars, and TVs in the participants' physical space. He also notes that such ambient noises may not mask verbal communication but are still audible. Jenks (2014) demonstrated that the participants accounted for these noises and that they had a choice as to whether to topicalise them and make them interactionally relevant.

Background noises in Jenks' (2014) data were found to have "interactional and/ or sequential consequences" (p.98), as they can hinder the ongoing conversation by stopping it momentarily or permanently. For instance, a white noise halted the progressivity of participants' activity of getting acquainted for 9 seconds before the activity was resumed following the disappearance of the noise. Another example from Jenks (2014) demonstrates

the termination of an ongoing conversation following the engagement of one of the participants in a parallel conversation with a non-participant in her physical proximity. In both of these examples, the background noises held the conversational floor as the other participants did not produce any turns while the parallel conversation was going on. Following the disappearance of such noises, pauses occur before a mutual orientation is re-established by the participants (i.e. talking again). Similarly, Alzaidi (2016) noted that background noises were among the trouble sources that occasioned multiple other-repair initiations in L2 video-mediated interaction on Google Hangouts.

In whole-class interaction, background noises are also reported to be a common source for interruptions. As part of an examination of the interaction patterns in teacher-led video-mediated 6th grade EFL lessons on Zoom, Cheung (2021) investigated the time that teachers spent on classroom management using the technological affordances at hand. Cheung's findings revealed that a common disruption to these classes was due to the interference of background noises. This is despite the fact that students in the whole-class situations were allowed to deactivate their audio when not speaking, in order to prevent background noises from interfering with class activities. To handle this interference, Cheung noted that teachers deactivated the learners' audio. Moreover, Cheung observed that learners who chose to have their microphones switched off could turn them on to answer a question before turning them off again. Similar findings regarding the teachers' use of audio deactivation features as a practice to maintain classroom order are presented by NurSürüç Şen (2022), who noted that teachers on some occasions used the feature of 'mute all' to restore classroom order when there are competing sounds. A reason behind such competing sounds is that children join from home, where there are several sources of background noise.

Overall, trouble sources in audio/video-mediated interaction and the methods by which they are flagged and resolved can differ from in-person settings. As noted, this is due to the fracturedness of the interaction (Luff et al., 2003), and the affordances and constraints of the mediums (Hutchby, 2001a, 2014). However, as aforementioned, participants can create novel practices and use the technology's features to overcome such troubles. The current study investigates the participants' use of the audio activation/deactivation features of the Zoom platform in managing turn-taking and handling trouble caused by the interference of disruptive noises. Successful management in online language teaching and learning settings requires mastering a number of competencies. In what follows, the competencies that previous research has suggested teachers need to acquire in order to succeed in managing online language teaching are reviewed.

2.4 Competencies in Online Teaching and Learning

For teachers to be successful in managing interaction and fostering language learning in online settings, researchers argue that a number of competencies need to be acquired in addition to classroom interactional competence (Grammens et al., 2022; Hampel & Stickler, 2005, 2015; Moorhouse et al., 2021). It should be noted that it is not a primary aim for the current study to investigate the classroom interactional competencies for teachers and learners in video-mediated L2 classroom interaction. However, an understanding of such competencies can help to develop the understanding of certain interactional decisions made by both teachers and learners. In the first analytical chapter, the learners' use of the audio activation/deactivation features to project the completion of the turn in progress and self-select for the next one is analysed. In the discussion chapter, this is linked to the teachers' and learners' classroom interactional and technological competencies. Therefore, an introduction to classroom interactional competence (CIC) may be useful at this stage.

CIC is defined as "teachers' and learners' ability to use interaction as a tool for mediating and assisting learning" (Walsh, 2011, p.158). It emphasises the essential role of interaction in language teaching and learning, and that a good understanding of this interaction will help teachers to facilitate it and, therefore, contribute to maximising the learning of the target language. As Walsh (2012) pointed out, CIC is concerned with understanding how interactional decisions made by teachers and learners in the classroom are made and how the actions that follow these decisions can increase the potential for language learning. According to Walsh (2012), there are three features where CIC becomes evident. First, CIC is context-sensitive, which means that different pedagogical goals may need different language use. A pedagogical task that focuses on eliciting the learners' opinions and experiences regarding a certain topic requires that teachers use wh-questions rather than polar questions. Second, CIC creates 'space for learning' as the teacher provides more opportunities for learners to participate and contribute to the ongoing class activity by increasing the wait time, encouraging extended learners' turns, and by providing an opportunity for learners to plan their contributions. Third, CIC manifests itself in what the teacher is able to do with the learners' contributions. Instead of only accepting the answers, the teacher can paraphrase them and perhaps extend them to introduce a new learnable such as new vocabulary, phrases, or grammatical structures.

As noted above, CIC is important and transferable to online language learning and teaching. However, Moorhouse et al. (2021) argued that additional competencies are needed for

teachers to be successful in synchronous online lessons. Based on a thematic analysis of self-reported survey responses by the teachers, Moorhouse et al. (2021) suggested three competencies: *technological competencies, online environment management, and online teacher interactional competencies* (pp.7-10). Technological competencies are important to overcome technical problems, such as connection troubles, that can result in frozen videos, muting delays by the students or overlapping talks. Teachers need such technological competencies to be able to solve technical problems and carry out some work around connection troubles and train the learners on using the platforms. For instance, Moorhouse et al. (2021) mentioned that teachers needed to set up some backup channels for communication, especially for students from China, who need VPN software to use platforms such as Zoom or Microsoft Teams. Accordingly, should teachers lack such technological competencies, overcoming such troubles can be challenging.

Second, Moorhouse et al. (2021) stressed the significance of *online environment management competencies,* as such environments can be dissimilar to in-person classes. Teachers need the ability to adapt aspects of their lessons' design to the affordances and constraints of the platform used. For example, lessons need to be simple and instructions need to be explicit and clear. Moreover, it is important for teachers to manage expectations and inform the learners of how they are expected to participate in these classes, such as whether or not to ask for permission to speak (Moorhouse et al., 2021).

Third, Moorhouse et al. also state that teachers need to acquire the *online teacher interactional competencies*. The teachers need to overcome the challenges of teaching in such environments and be able to facilitate interaction. This includes enabling the learners to use the different channels to communicate, such as text-chats in videoconferencing software. Such a strategy can help overcome the challenge of shifting between muting/unmuting for the learners. Thus, the learners become able to participate more and, as a result, learn. Also, Moorhouse et al. suggest that teachers can increase their wait time in such environments to provide more opportunities for learners to answer.

Building on the findings from their survey, Moorhouse et al. (2022) conducted an empirical study to explore the professional practices by the teachers in synchronous online lessons via videoconferencing and to identify the competencies required; a refined e-CIC framework was proposed. Unlike their previous study, the refined framework points out that these three key competences are intertwined, but that they have a hierarchal relationship. Figure 2.1 below demonstrates that the technological competence encompass the other two competences, as all

actions and interactions in these lessons are mediated by technology. The authors highlighted the need to create context for interaction in these lessons by the teachers. Following the establishment and maintenance of the context, teachers need to be able to help the learners maximise their opportunities to learn by offering them more opportunities to interact (Moorhouse et al., 2022). For successful teaching in synchronous online lessons, teachers need to use all these competencies (ibid).

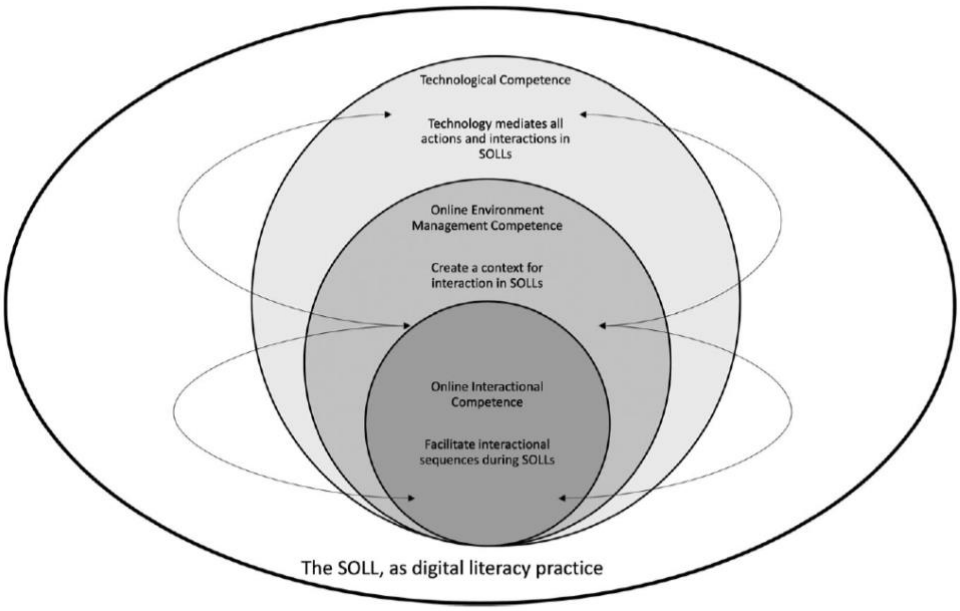


Figure 2.1: e-CIC framework (Moorhouse et al., 2022, p.965)

Similarly, a set of skills was suggested by Hampel and Stickler (2015), originally in Hampel and Stickler (2005), for teachers to be successful in online teaching (see Figure 2. **2**). This was different from Moorhouse et al. (2021, 2022), who examined synchronous lessons via videoconferencing during COVID-19 pandemic. Hampel and Stickler (2005, 2015) propose their set of skills for teachers in all different modes of online teaching. In their skills pyramid, teachers need to acquire the lower level of skills in order to be able to achieve and master the higher ones.

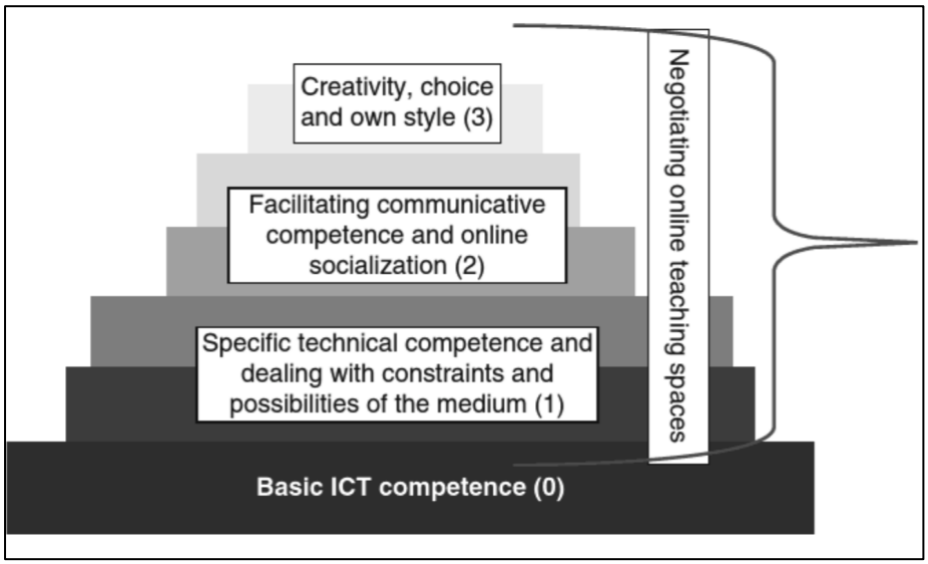


Figure 2. 2: Online Tutors' Skills Pyramid (Hampel & Stickler, 2015, p.66)

Hampel and Stickler's (2015) skills pyramid consists of four levels. The first level includes the ability to use basic knowledge and the ability to use technology, such as the use of the Internet, sending emails and using word processors (level 0). Once this basic competence is acquired, the teacher needs to acquire sufficient knowledge of the platform used, to be able to deal with its affordances and constraints (level 1). This familiarity will allow the teachers to design then the learning materials to be compatible with the main and/or supporting platforms. Placing these two levels at the base of the skills pyramid by Hampel and Stickler shows their importance to effective online teaching. Level 2 encompasses the skills needed by the teachers to promote socialisation, build rapport both with and between the learners, and maintain a sense of community that is similar to what is found in in-person classes. Teachers need to set the etiquette framework in these online environments and encourage the learners to follow and respect them. Level 3 is concerned with the skills needed by the teachers to be creative and to devise their own style of teaching in online settings, which is the highest level in this pyramid (Hampel & Stickler, 2005, 2015).

A similar set of roles and competencies required from the teachers' side in video-mediated settings are reported by Grammens et al. (2022). In their systematic review of previous research on synchronous teaching using videoconferencing systems, they identified five main roles for teachers: instructional, technical, social, managerial, and communicational. To enact each of these roles, teachers are required to have a set of competencies. In the technical role, for instance, teachers need to be able to provide technical support, have sufficient knowledge about using the different tools in the platform, and be familiar with the technical tools' affordances and constraints.

The focus in the studies cited above is on the competencies that teachers need to acquire in online synchronous classes. However, interaction in online synchronous classes includes both parties – teacher and learners as individuals or as groups (Lerner, 2002). Therefore, a question to be asked at this stage concerns the relevance of acquiring such competencies to the learners in order to successfully interact and maximise their participation and learning opportunities in such L2 video-mediated classrooms. In the Discussion Chapter, an argument will be made that learners also need to acquire online interactional and technological competencies in order to participate successfully in L2 video-mediated classes.

2.5 Research Gap

A growing field has developed in recent years that is interested in investigating interaction in video-mediated interaction. A considerable amount of literature has grown around a central theme, such as affordances (Hutchby, 2001b, 2003a). This is in an attempt to understand how technology features can afford or constrain the structure of social interaction. Despite the growing field, Arminen et al. (2016) called for more work to explore and unpack the participants' use of the medium's affordances for the organisation of mediated interaction. A central theme in the current study is to examine the participants' systematic use of the audio activation/deactivation features to manage their turn-taking and repair speaking/hearing trouble caused by disruptive noises in synchronous video-mediated L2 classrooms. Even though some studies cited in this chapter (Andrews, 2020; Cheung, 2021; Moorhouse et al., 2021; NurSürüç Şen, 2022) have mentioned the teachers' and learners' use of these features in the management of their interaction in video-mediated L2 learning, none of these studies provides a multimodal fine-grained analysis of the unfolding of their use by the teachers and learners to organise turn-taking and handle trouble caused by the interference of disruptive noises. The strategic use of these features by the participants in terms of their sequential positions and the reflexive relationship between their use and the organisation of turn-taking and repair in L2 video-mediated classrooms remains an unexplored area.

Similarly, the studies that examined turn-taking in audio/video-mediated L2 learning and teaching settings provided valuable results in terms of turn allocation, regarding managing overlaps and turn allocation. Jenks (2009) revealed that timing contributions by participants can be a challenging skill in audio chats, due to the lack of embodied displays. This challenge can gain more complexity when the participant has their audio deactivated, even with the availability of embodied resources in video-mediated classes (Andrews, 2020). This complexity and the learners' use of the audio activation feature to project the possible

completion of the turn in progress and indicate that their incipient speakership is an area that needs further investigation. Also, the consequentiality of mistiming the audio activation to self-select for the next turn is worthy of examination. Oittinen (2020) also called for more work on larger datasets to extend the understanding of how participants coordinate their actions to reach the smooth running of interaction in different contexts mediated by technologies.

A number of scholars have reported findings on the interference of background noises in audio/video-mediated interaction (Alzaidi, 2016; Cheung, 2021; De Fornel, 1996; Jenks, 2014). However, an examination of the handling of this interference using the technology affordance is not yet produced. Indeed, Jenks (2014) called for a detailed study of participants' methods to repair the troubles that can be caused by the interference of noises coming from the participants' backgrounds. The current study investigates the moment-by-moment unfolding of trouble caused by such noises beginning with their emergence, the participants' orientation to them, repairing the trouble, and restoring the progressivity of the suspended activity. Previous researchers also examined turn-taking and repair in a variety of CMSI settings, including L2 learners interaction in audio chats (Jenks, 2009, 2014); large teacher-fronted classes on Zoom (Badem-Korkmaz & Balaman, 2022; Cheung, 2021); dyadic interaction between tutor and tutee (Malabarba et al., 2022). However, interaction in L2 teacher-led small group speaking classes remains an under-explored context.

Against those backdrops in the literature, the current study seeks to fill these gaps by providing a multimodal analysis of the moment-by-moment unfolding of the participants' use of audio activation/deactivation features to organise turn-taking and repair in L2 video-mediated classrooms.

2.6 Summary

This chapter has reviewed the concepts relevant to the study's quest to investigate participants' systematic use of the audio activation/deactivation features to manage their turn-taking and manage trouble caused by the interference of disruptive noises and restore/maintain the progressivity of the ongoing activity. To begin with, a brief introduction to classroom interaction and what characterises it was provided in Section 2.2; then the literature related to turn-taking in classroom interaction, which was reviewed in Section 2.2.1. In this section, a review of how turn-taking organisation in classroom interaction deviates from that in ordinary conversations was made. In addition, it discussed the techniques for turn allocation in the classroom by the teacher and the methods by which learners bid for the next

turn and display willingness to participate. Section 2.1.1.1 presented the methods by which the learners project self-selecting in classroom interaction. Next, Section 2.2.2 reviewed the literature related to managing progressivity in classroom interaction and aspects of trouble that can hinder it.

In Section 2.3, the discussion turned to the CMC settings to show the differences between synchronous and asynchronous settings; following this, CMSI was presented in Section 2.3.1 and video-mediated L2 classroom in Section 2.3.2. This section provided a brief historical overview of the study of video-mediated interaction. Subsequently, Section 2.3.1.1 presented the notion of affordances as an approach to understanding the complex relationship between technology and the structure of social interaction. Section 2.3.1.2 then explained how connecting from different geographical spots can affect interaction in video-mediated interaction, known as 'fractured ecologies'. Section 2.3.3 presented a review of the previous research findings in the context of L2 video-mediated classrooms in relation to the organisation of turn-taking (Section 2.3.3.1); and repair, trouble sources and how participants dealt with them were shown in Section 2.3.3.2. The discussion was next directed towards the competencies that teachers need to acquire in order to successfully manage interaction in video-mediated L2 classrooms (Section 2.4). The chapter concluded by presenting the research gaps found in the literature and which the current study seeks to bridge (Section 2.5).

The following chapters introduce CA as the methodology used to examine the participants' use of the audio activation/deactivation features to manage their interaction in video-mediated L2 speaking classrooms. This is followed by the research design Chapter, which details the processes of data collection and analysis in the current study.

Chapter 3. Methodology

3.1 Introduction

The current study employs CA as a methodology to examine the participants' use of audio activation/deactivation features as practices to organise turn-taking and repair in online, synchronous, video-mediated, L2 classes on the Zoom platform. This chapter provides an overview of CA's epistemological and theoretical foundations, as well as its core principles, serving to guide the collection and analysis of the data for the study. Providing such an overview is to support an understanding of the analysis of examined social phenomena in the subsequent chapters.

The organisation of this chapter is as follows. Section 3.2 revisits the purpose of the study. This is followed by an introduction to CA and its main principles in Section 3.3. The next section, Section 3.4, introduces and discusses ethnomethodology, the epistemological foundation of CA and provides an overview of its main principles, as they underpin CA as an approach to study social interaction. Section 3.5 then presents the basic interactional machineries that previous CA studies have revealed, such as sequence organisation, turn-taking, and repair. Next, the chapter discusses the issues of validity, reliability, and generalisability of CA research and how these issues are dealt with (Section 3.6). The chapter concludes with an overview of the main criticisms and limitations of CA, in Section 3.7.

3.2 Purpose of the Study

The main aim of the current study is the investigation of the participants' use of audio activation/deactivation technological features as resources in the organisation of turn-taking and repair in online synchronous video-mediated L2 classes on the Zoom platform. The researcher aims to fulfil the study's objectives by using a sequential, multimodal analysis. Previous literature on classroom discourse and talk-in-interaction has identified a number of contextual and methodological gaps. Firstly, there is a need for more work to better understand the roles of mediating technologies in the construction of social interaction (i.e. facilitating or constraining it) (Arminen et al., 2016; Hutchby, 2001a, 2013, 2014). Secondly, the literature shows that there is a need to examine a large set of data to learn more about the participants' coordination of their actions in video-mediated interaction (Oittinen, 2020). Thirdly, according to Jenks (2014), repair sequences, preference organisation, and other sequential aspects of talk are vital to the unfolding of computer-mediated communication. Jenks calls for addressing the empirical question of how talk in interaction is managed in

online computer-mediated communication. This study seeks to address this question by investigating the participants' use of audio activation/deactivation features as repair devices that manage disruptive background noises. Finally, the current study investigates the participants' deployment of audio activation/deactivation features as resources to organise turn-taking in video-mediated L2 classes. The following section presents CA as the methodology that guided the collection and analysis of the data in this study.

3.3 Introduction to Conversation Analysis

This study uses CA as a methodology to examine interaction in online L2 video-mediated speaking classes. In this section, a brief overview of CA's history and its basic principles will be provided. CA has been defined as "an approach within the social sciences that aims to describe, analyse, and understand talk as a basic and constitutive feature of human social life" (Sidnell, 2010, p.1). It aims to systematically analyse the "talk produced in everyday situations of human interaction: talk in interaction" (Hutchby & Wooffitt, 2008, p.11). It was principally originated by Harvey Sacks in his lectures during the late 1960s and in the work of his collaborators Emmanuel Schegloff and Gail Jefferson in the early 1970s (Schegloff & Sacks, 1973; Sacks, Schegloff & Jefferson, 1974; Sacks & Schegloff, 1979).

According to Seedhouse (2005, pp.166-167), CA has developed its own principles to be used when adopting a conversation analytic way of looking at data. In what follows, I briefly introduce these principles. The first principle is that of Rational, which refers to the notion that conversation as social interaction is systematically organised and follows 'order at all points' (Sacks, 1992). The notion of 'order at all points' is one of the core assumptions of CA and assumes that there is an overwhelming order in conversation (Liddicoat, 2021). Such order is produced by the participants through the coordination of their practices in conversation; it is not a pre-existing condition (Psathas, 1995).

The second principle is the participants' contributions to interaction are context-shaped and context-renewing. To elaborate, this principle indicates that it is difficult to understand the contributions of participants without referring to the sequential environments in which these contributions are produced. The context in which participants speak, and for, shapes what they say, and the subsequent turn in a conversation is understood in relation to what came before it (Heritage, 1984b). Since each of these contributions is produced as a response to what preceded it, their production creates a context for the next turn. Hence, the context is renewed with each contribution by the participants.

Third, CA analysts deploy a transcription system that caters for the tiniest details of interaction. Additionally, this system has a highly empirical orientation in the sense that no order can be excluded or considered a priori, accidental, or irrelevant (Heritage, 1984b, p.241). CA analysts focus on the recording of naturally occurring interaction and then transcribe these recordings in order to make them available for close examination by both themselves and the readers. The transcription system used in this study will be further discussed in the following chapter. Fourth, analysis in CA is bottom-up and data-driven. Seedhouse (2005) asserted that analysts should not have any prior theoretical assumptions or try to cast any background knowledge on the analysis when looking at data, unless this was oriented to by the interactants. This, in fact, is one of CA's main strengths, as it makes a robust method for analysis of social interaction and produces findings that can be tested by the readers.

In its early days, CA was concerned with analysing audio recordings of daily life situations, such as phone call openings and conversations in family settings, as exemplified by the work of Schegloff and Sacks (1973). The use of audio recording technology at that time contributed to the rise of CA in that it enabled researchers to revisit the data multiple times and closely observe any interesting phenomena. This development in audio and, later, video recording technologies was a significant transition from traditional observation and note-taking, in which the observer may miss important details and depends mainly on what they see or remember.

Since the early 1970s, as technology has developed further, CA practitioners began to look at video recordings of interactions that occurred in everyday life and, subsequently, the scope of CA was widened to include the study of interaction elements beyond speech, such as gaze, body movement, and the manipulation of objects within the environment in which the interaction occur (Jewitt, Bezemer, & O'Halloran, 2016). The technology advancements made it possible for researchers to capture naturally occurring interaction on video and examine it.

A number of concepts are central to the study of multimodal interaction by conversation analysts: mutual elaboration of semiotic resources for each other, sequential organisation of actions, coordination of actions, and multiactivity (Jewitt et al., 2016, pp. 91–96). A fundamental notion is that people construct their actions using the different semiotic resources available in the environment. According to Goodwin (2000), the understanding of social actions may be incomplete when one of these semiotic resources is isolated from the others; they mutually elaborate on each other to produce a whole meaningful entity.

The second key concept is that actions are sequentially organised, which refers to how actions in a certain situation are organised in a sequential manner that involves the use of semiotic resources including language, gaze and hand or whole-body movements that interactants use to form their actions. The video recordings of interaction enabled the examination of other methods by which participants perform social actions using gesture (Mortensen, 2016), gaze shifts (Rossano, 2013), artefacts (Mondada, 2007, 2014), or technology affordances in online contexts (Badem-Korkmaz & Balaman, 2022; Malabarba et al., 2022; Melander Bowden & Svahn, 2020), as will be seen in the analysis in this study.

Thirdly, the concept of coordinated actions refers to the study of how people in a given situation coordinate their actions with those of the other interactants. Hindmarsh and Pilnick (2002) pointed out that participants might orient to the beginning of an object manipulation as projecting a specific social action, and "witnessing someone beginning to engage in a particular activity can project a trajectory of actions that routinely follow" (p.151). Moreover, the environment in which these actions are produced plays a role in the participants' coordination of actions. For example, Mondada (2007) notes how interactants make use of the specificity of the environment and available artefacts to coordinate their actions in work meetings. The concept of coordination of actions is of importance to the current study as it seeks to analyse the relationship between the use of technology features and the projection of courses of social actions in video-mediated L2 classroom interaction and how it is coordinated with the other participants' conduct.

Fourthly, the multiactivity concept considers how interactants can be engaged in multiple activities in everyday life interaction. These multiple activities can occur simultaneously (i.e. parallel to each other) or interactants may alternate between activities by temporarily halting the progressivity of one action in favour of another (Haddington, Keisanen, Mondada, & Nevile, 2014). Similarly, it is uncommon for participants to be involved in multiple simultaneous activities in online contexts (Ruhleder & Jordan, 2001b) and orient to a certain action as demanding more immediate attention at a particular moment (Haddington et al., 2014).

Later, CA researchers expanded their use from daily in-person interaction to interactions occurring in video-mediated settings. As established in the previous chapter, interaction in video-mediated settings can differ from in-person interaction. Participants can simultaneously use multiple modalities (Mlynář et al., 2018) - such as speech, typing, and sharing content on a shared screen - that do not exist in in-person interaction, as will be demonstrated in the

analysis chapters. Moreover, interacting in fractured ecologies (Luff et al., 2016) can make certain multimodal behaviours, such as gaze direction, body posture movement, or pointing, less recognizable. Luff et al. (2013) showed that while a participant may appear on the screen to be looking at another participant, he or she might be looking at a colleague or an object in their physical space. Such difficulty in determining gaze direction poses a challenge regarding establishing mutual gaze between participants in video-mediated interaction.

When examining video-mediated interaction, conversation analysts must address various challenges in terms of data collection, transcription, presentation, and analysis. First, researchers need to consider questions regarding the type of data to collect (e.g., screen recordings, platform-generated recordings, or both). Moreover, researchers need to decide on the number of perspectives required (e.g., only teachers' screens or screens of all participants in multi-party interactions). Also, CA researchers need to consider the need to capture off-screen interaction (e.g., interacting with different devices, other people, pets, or objects in the physical space). In doing CA, researchers must efficiently capture social interaction in a given setting to produce a rigorous analysis of a particular phenomenon (more details on the process of data collection and its relevant challenges for the current study are presented in Section 4.4).

Following data collection, researchers may face challenges related to transcription and data presentation. As mentioned, participants can engage in multiple activities simultaneously, both on and off the screen. To transcribe and present these multiple - and in many cases subtle - activities, researchers need to be creative and produce transcripts that account for the relevant details. Previous studies (such as Balaman, 2016, 2019; Meredith, 2016) have produced excellent representations of data that accounted for small multimodal details of interaction, such as clicking, mouse movement, and typing, among others. These data representations also included unanalytical descriptions of details, such as a participant's screen activities, that are not accessible to fellow participants to add clarification for the readers. These are presented along with illustrations on still images from the participants' screens. It can be challenging to transcribe all the details of interaction at a particular moment, but it is a decision that a researcher is required to make regarding the details which need to be included.

Regarding the analysis of multimodality in online video-mediated interaction, CA has proven to be a useful methodology despite the challenges, as evident in recent studies (Badem-Korkmaz & Balaman, 2022; Balaman & Doehler, 2022; Balaman & Sert, 2017; Jakonen & Jauni, 2021; Licoppe & Morel, 2018; Luff et al., 2016). These studies presented rigorous,

evidence-based analyses of multimodal interaction in video-mediated settings, employing CA terminologies such as sequence, turn-taking, and repair. One challenge in analysing multimodal interaction in video-mediated settings (see Section 2.3.3) is that trajectories of actions in video-mediated settings may differ from those in in-person settings. This difference can be attributed to interacting in fractured ecologies (Luff et al., 2003, 2016) and the constraints the setting imposes on the production and recognition of conduct.

Following this brief introduction to CA and its main focus, the next section elaborates on its epistemological foundation: ethnomethodology (Garfinkel, 1964,1967) and its basic principles.

3.4 Ethnomethodology

This section briefly introduces ethnomethodology, together with the emic perspective and the basic principles that have influenced CA. As noted above, CA has its roots in Garfinkle's ethnomethodology. For a clear understanding of CA's theoretical roots, one must take a closer look at ethnomethodology, which appeared as a radical departure from the dominant Parsonian theoretical paradigm at that time. According to (Seedhouse, 2004), ethnomethodology is concerned with the principles that form the basis of people's social actions, while the focus of CA is narrowed down to the study of how people use language and other conduct to accomplish interaction with each other.

It is also worth mentioning that both ethnomethodology and CA follow an emic perspective of analysis. This means that analysts examine human social interaction from inside the system, rather than from outside (etic perspective) (Pike, 1967). Analysts do not start from theories or hypotheses formed by social scholars regarding social interactions and cast those on their analysis (Lett, 1990). Instead, they analyse interaction following the participants' orientations and interpretations of social actions. In the emic perspective, analysts refrain from any prior assumptions, as the social world is only examined through the participants' perspectives and what occurs in a certain setting.

Ethnomethodology has a number of core principles: indexicality, documentary methods of interpretations, reciprocity of perspectives, and normative accountability. Indexicality refers to the notion of the context-boundedness of utterances. To elaborate, interactants do not explicitly state everything they want to convey. Rather, they orient to a shared and mutually understood background context to convey the full meaning. This shared background, or indexical knowledge, is something that interactants talk into being (Seedhouse, 2004).

Interactants demonstrate a reflexive relationship between talk and the social context in which it is produced through their talk (ibid). Hence, CA does not consider any contextual features that the interactants do not orient to in their talk. The only contextual features to be considered are the ones that the interactants mutually orient to in their social interaction.

The documentary method of interpretation is a core principle of ethnomethodology. It is a term that Garfinkle (1984c) used following Mannheim (1952), who observed that an action is treated as a 'document' that points to a presupposed pattern underlying it. The interactants treat actions as a document of an already known pattern to them. For example, an invitation is identified as such and is treated according to the already known invitation patterns. Once the interactants are introduced to a new way for inviting, this new document is added to the existing patterns that fit it. Seedhouse (2004) points out that it is fundamental that CA analysts also use this method of interpreting social interactions as maintaining an emic perspective requires the examination of the data from the participants' perspectives.

Another important principle of ethnomethodology is the reciprocity of perspectives, which reveals the interactants' tendency to follow the same norms and to display their alignments with their counteractants' perspectives in order to achieve mutual understanding (Seedhouse, 2004). This is not to say that the interactants do not deviate from these norms. However, a difficulty in reaching intersubjectivity becomes noticeable when these norms are breached by the participants. In his breaching experiments, Garfinkle (1967) showed that challenging the achievement of the reciprocity of perspectives resulted in rage and the interactants "vigorously sought to make the strange actions intelligible and to restore the situation to normal appearances" (p.47). The reciprocity of perspective links to the preference organisation in CA, and a manifestation of it can be seen in the structure of adjacency pairs. For example, Seedhouse (2004) points out that some of the first pair parts in adjacency pairs can be met with more than one possible second pair part. Some of these pair parts are preferred, while others are dispreferred (ibid). The production of the preferred response is seen but unnoticed, as it displays reciprocity of perspective. In contrast, the production of the dispreferred response becomes noticeable, as it breaches this display of affiliation (Seedhouse, 2004).

The last ethnomethodological principle to be discussed here is normative accountability, which forms the basis for both ethnomethodology and CA. According to Heritage (1984), Garfinkle's notion of normative accountability indicates that the interactants have normative expectations and that these are not regulative of actions, but rather constitutive of actions. The

interactants can "design their social actions and interpret those of others by reference to these norms" (Seedhouse, 2004, p.10). These normative expectations are not to be looked at as social rules; rather, they are used by the interactants as 'action templates' to interpret social actions (ibid). For example, the production of a normatively expected action (acceptance to an invitation) is seen but unnoticed, while the failure to produce an acceptance or a declination to that invitation can be noticeable, accountable and perhaps sanctionable (Heritage, 1984).

This section has presented an overview of ethnomethodology, which is considered to be the theoretical foundation of CA. The next section presents the interactional machineries that organise speech and other conduct production in interaction.

3.5 Interactional Machineries

In interaction, there are a number of interactional organisation types that people use to produce social actions and analyse those of others (Hoey & Kendrick, 2017; Seedhouse, 2005a). As Sidnell (2011) puts it, the term organisations, or machineries, as metaphorically referred to by Sacks and colleagues, simply means that there is an organised number of practices that people use to gain and put together their turns or to repair problems in conversation. It is by using these types of machineries, analysts can reveal the orderliness of conversation and of interaction. The following section briefly discusses some of these machineries, as they guide the data analysis for this study.

3.5.1 Sequence Organisation

Sequence organisation is a general term that refers to how turns are relatively positioned in a successive nature to form coherent and orderly courses of social actions (Schegloff, 2007). A clear manifestation of sequence organisation is the concept of adjacency pair, which consists of two actions linked together in its basic form. There are many examples of adjacency pairs, such as invitation-acceptance/decline, question-answer, or greeting-greeting. They are considered the building blocks for sequences and "the basic building-blocks for intersubjectivity" (Heritage, 1984b, p.256). An adjacency pair is a couple of turns produced by different speakers in which the first pair part (FPP) is followed by a type-matched second pair part SPP which is conditionally relevant to the (FPP) (Seedhouse, 2004). The conditional relevance of an adjacency pair means that the SPP production depends on the production of the FPP. Moreover, when an FPP is produced, an SPP of the same type is normatively expected, and its absence is noticeable by the fellow participant(s), accountable and may be

sanctionable. For instance, when a person greets another, the other person is normatively expected to return a greeting. However, when returning the greeting does not occur, this may be noticeable and accountable.

Moreover, Schegloff (2007) points out that sequences can also be expanded beyond the basic form of the adjacency pair by inserting a sequence in different positions in relation to the base adjacency pair. To elaborate, these inserted sequences can occur in a pre-sequence position (i.e. before the base adjacency pair), between the two parts of the adjacency pair, or in some cases following the production of the SPP (post-expansion) (ibid).

An important aspect to consider alongside adjacency pairs is preference organisation. According to Hutchby and Wooffitt (2008), the way conversation analysts use the notion of preference "is not intended to refer to the psychological motives of individuals, but rather to structural features of the design of turns associated with particular activities" (p.46). The interactants can then utilise these features to interpret the type of actions in these turns (ibid). Regarding adjacency pairs, the available SPPs to an FPP are not of the same value, which can be seen in the different designs of each. To elaborate, Sacks (1987/2020) demonstrated that interactants can design their turns (invitation) to receive a particular response (acceptance). As noted earlier, the absence of this particular response can be accountable. Sacks also draws attention to the preference for contiguity, which refers to the notion that both parts of an adjacency pair follow each other (i.e. no talk inserted in between) (ibid). Additionally, there is a relationship between the design of these responses and their alignment or affiliation with the previous turn. In the case of an invitation, for instance, an acceptance is produced immediately while a rejection might be delayed, which are both markers of preference or dispreference (Pomerantz, 1984).

3.5.2 Turn-Taking

Turn-taking is a term used to describe how people normally take turns to speak in normal conversations. The study of turn-taking in CA delves deep into the organisation of the practices by which people take turns in talk-in-interaction and adhere to the rule of "one party talking at a time" (Schegloff, 2000a, p.1). Sacks et al. (1974) illustrated that two components influence the process of turn-taking in talk. The first is what they termed 'Turn Constructional Unit' (TCU) and the latter component is 'Turn Allocation'. Each turn in a conversation is built from TCUs, and each turn can consist of one or multiple TCUs, which can be any linguistic unit, such as a word, a clause, a sentence or even small units like 'hmm' (Hoey & Kendrick, 2017; Sidnell, 2010). TCUs are not exclusively spoken; they can be a gesture, body

movements, or a click on a button. The possible completion of each of these units creates a Transition Relevance Place (TRP), where a transition space to the next speaker is open (Hoey & Kendrick, 2017). The opening of a transition space is not only marked by the completion of the TCU, but is also "projected in advance through various practices that are understood in context as foreshadowing that the turn-in-progress may be winding down" (Clayman, 2013, p.151). This feature offers the recipient (s) the ability to systematically identify the current turn's completion in advance, and thus, anticipate the TRP at which he/she can speak (Sacks et al., 1974; Schegloff, 1984; Sacks, 1992).

A distinction has been drawn between the two levels of macro- and micro-projection by Schegloff (2013). Macro-projection concerns the projection of the overall structure of the whole TCU or turn. Contrastingly, micro-projection concerns the linguistic organisation of a TCU (smaller components within a TCU). Even though Schegloff (2013) is more concerned with the syntactic level of TCU(s), the literature shows that there is a number of indicators or cues that can project the possible completion of a TCU, such as syntactic (Duncan, 1972; Hayashi, 2004; Lerner, 1991), prosodic (Auer, 1996; Bögels & Torreira, 2015; Rühlemann & Gries, 2020), pragmatic (Levinson, 2013; Levinson & Torreira, 2015), and non-verbal (Mondada, 2006; Rossano, 2013; Streeck, 1995).

The second component of turn-taking is the turn allocation component. In ordinary conversations, Sacks et al. (1974) note that a general 'rule' in conversation is one party speaks at a time and there is a tendency towards no gaps between speakers. Moreover, overlaps are kept short using resolution devices (Schegloff, 2000). Also, Sacks et al. (1974) presented a set of locally managed practices by which the interactants take turns in a conversation. The transition from one speaker to another can occur in different scenarios; the current speaker, during an ongoing turn, can select the next speaker. Also, the current speaker may not specify who is going to speak next, which means that any of the participants can self-select for the next turn (usually the first to do so gains the interactional floor). In the cases where the current speaker does not select the next speaker and none of the participants self-selects, the current speaker may or may not keep the turn by continuing to speak.

3.5.3 Repair

The notion of repair is an important CA principle, which can be defined as a set of practices that people use to treat speaking, hearing or understanding problems that can occur in a conversation (Schegloff et al., 1977). These problems in talk are vital, as they can hinder the progressivity of talk and, subsequently, threaten intersubjectivity. Repair is a solution to

secure the flow of interaction and achieve or re-establish intersubjectivity, as participants treat trouble sources during their interaction (Schegloff et al., 1977). As a concept, repair is far broader than correcting and replacing incorrect linguistic forms with correct ones (ibid). Evidently, errors in linguistic forms in many cases do not require repair by participants, who may choose to let it pass as this does not threaten intersubjectivity (Firth, 1996). Problems that need to be fixed in talk are referred to in CA literature as "repairable" or "trouble sources" (Schegloff et al., 1977). Sacks (1987) argued that repair is motivated by the participants' desire to get things right, and, therefore, this triggers the initiation and execution of repairing trouble sources in their talk.

Three components constitute the complete repair sequence: trouble source, repair-initiation, and repair execution. A trouble source in a conversation can be of any form, such as an ambiguous word, sentence, or turn that the participants detect. According to Schegloff et al. (1977), there is nothing that cannot be considered a trouble source. The current study presents types of trouble sources that are not spoken and may not necessarily be produced by the interactants, but are oriented to as such by fellow participants.

Upon the occurrence of a trouble source, there are four possible sequences of repair that can occur in relation to who initiates and performs repair (Liddicoat, 2021). In terms of initiating repair, it can be done either by the current speaker (self) or by other interlocutors (other). The speaker of the trouble source can point out the trouble and carry out the repair. Also, the speaker can indicate the trouble, but the repair can be carried out by another participant. Other participants can indicate the trouble and the trouble source producer carries out the repair. The same goes for who repairs the trouble; the other participants can indicate the trouble and carry out the repair as well. In summary, the current speaker can do self-repair. Alternatively, other interlocutors can repair the trouble source.

In terms of repair sequences initiated by other than the trouble source producer, Kendrick (2015) presented two types of other-initiated repair structures: minimal other-initiated repair sequences and non-minimal other-initiated repair sequences (p.166). The former type is constituted by three turns, where the trouble source is produced in the first. The second turn is the repair initiation by other than the trouble source producer. Finally, following the repair initiation, the repair solution is provided by either the trouble source's producer or any of the other participants. However, non-minimal other-repair initiated sequences ensue when a repair initiation fails to resolve the trouble, which can lead to multiple other-repair initiations, especially in a multi-party interaction. The repair initiation can be ambiguous or insufficient,

which may then be a trouble source itself. This, in turn, results in another repair initiation (Kendrick, 2015).

3.6 Reliability and Validity of CA Research

In both quantitative and qualitative approaches to research, researchers need to explicitly explain their procedures to ensure the validity and reliability of their work. Despite being a requirement in both approaches, Creswell (2018) states that these two terms have different meanings in each approach. In this section, the measures taken to ensure the reliability and validity of this study are showcased. CA has developed its own procedures to maintain validity and reliability due to its development of an emic perspective, and it maintains reliability throughout the stages of collecting data, transcribing them, and building collections of instances (Seedhouse, 2005). Below, the procedures taken throughout this study to maintain a high level of reliability are explained.

In conducting CA, the researcher needs to ensure a high-quality recording of what is being studied. In the current study, all the recordings are in full high-definition resolution (1080p) and made from an angle that allows the capture of different activities on the screen (the process of recording the data will be further illustrated in Section 4.4.1). Prior to the data collection, the broad interest was in examining the teachers' managing interaction in online, video-mediated, English speaking classes on Zoom. Thus, the researcher offered training on using the screen recording software to the teachers and ensured that technical support was always provided when required. This was undertaken in order to ensure that the recordings were of high quality.

Another factor that raises the level of reliability is producing clear and adequately detailed transcripts of the data, which are always provided, along with the results in CA publications as a general practice (Seedhouse, 2005). To be able to create adequate transcripts, the researcher attended a number of transcription workshops and CA training courses at Newcastle University, Loughborough University, York University, and the University of Oulu. Also, the transcripts were developed at several stages to improve their quality and enhance their readability (the data transcription process is further elaborated in Section 4.4.2). CA studies enable the reader(s) to view the data while reading the researcher's analysis, which makes it possible for them to carry out their own analysis and agree or challenge the analyst's claims. Conversation analysts nowadays offer readers not only still images of the data; they also present the video recordings, which makes it possible for a "a strict empirical discipline" (Arminen, 2017, p.68).

Additionally, CA practitioners often present their data and transcripts in public venues such as conferences and data sessions to gain more insights into the accuracy of the transcripts and the interactional phenomena presented. Regarding the current study, multiple segments of the data were presented in a number of data sessions, including the multimodal analysis research group (MARG) at Newcastle University, (CADSS) at Southampton University, a as well as various other venues. These presentations allowed the possibility for the data and transcripts to be shown to experts in the field in order to receive and consider highly constructive feedback and insights regarding both the transcripts and the analysis.

In regards to validity, it is concerned "with the integrity of the conclusions that are generated from a piece of research" (Bryman, 2012, p.47). Seedhouse (2005) presented four types of validity in CA studies: internal validity, external validity, ecological validity, and construct validity (pp.255–259). In what follows, the first three types of validity will be briefly presented.

Internal validity questions the credibility of the findings and whether the dataset does back the researcher's claims. Due to maintaining an emic perspective in CA research, all claims made by the researcher are entirely based on the data that is presented for the reader to test. Researchers are not allowed to invoke contextual details on the analysis, and they are also unable to cast their ethnographic knowledge over the analysis. Therefore, the analysis in this study is entirely based on the participant's orientations to the use of the platform features to organise turn-taking and repair. Such transparency in CA research makes it possible to test internal validity, whilst simultaneously increasing it.

External validity concerns the extent to which the research findings are generalisable and can be applied in other settings. Generalisability is often linked to quantitative research, which has large sample sizes, and there is an impression that qualitative research is difficult to generalise due to its small sample sizes (Bryman, 2016). In CA literature, there have been attempts to add quantification to generalise results, such as in the works of Stivers and Rossano (2010), as well as Zimmerman and West (1975). However, as Schegloff (2010) pointed out, this kind of quantification ignores the details of each occurrence. Additionally, CA research aims to explain how something happens in social interaction, not how many times it happens. Nonetheless, Seedhouse (2005) noted that CA can produce descriptions of a certain settings' interactional organisation, such as classrooms, where interaction is organised around a social goal, which can be generalisable.

Ecological validity concerns the applicability of the findings to the daily natural social settings of people (Bryman, 2016). CA researchers, as a standard practice and to maintain an emic perspective, tend to record naturally occurring interactions that would have happened regardless of the researcher's presence. This practice directly addresses the concern of ecological validity and renders CA research highly ecologically valid (Seedhouse, 2005).

3.7 Conversation Analysis for this Study

Making the decision to choose a particular research method stems from the research interests and objectives. Within the field that examined interaction in L2 video-mediated classrooms, previous researchers focused on the way interaction in L2 video-mediated classrooms is constructed. Within this research interest, a number of approaches exist, such as Multimodal (inter)action Analysis MIA (for a full introduction, see Norris, 2004, 2020) and CA. In this regard, Antaki (2008) states that there are multiple methods for researchers interested in analysing what people say or write to choose from; however, the researchers' choice must be made based on what they seek to examine. In what follows, a justification for choosing CA over MIA in the current study is provide.

In the case of the current study, the broad interest was in examining how interaction is managed in L2 video-mediated classrooms. On the one hand, MIA takes as its primary aim the presentation of a holistic analysis of interaction in the multiple levels of micro, meso, and macro levels from a broad sociocultural viewpoint (Norris, 2004, 2020). It affords the researchers the ability to move between these levels of analysis. To this end, it provides the researchers with novel analytical tools such as modal density, modal configurations, and foreground-background continuum of attention/awareness, among others (Norris, 2020). However, as Antaki (2008) noted, the choice of an approach to examining discourse is tied to the research interest. The interest in the current study was not in carrying out a macro analysis or examining the "organisation of the hierarchical ordering of modes" (Norris, 2020, p.17) or, "cognitive/psychological aspects such as attention/awareness" (Pirini, 2015, p.12).

Rather, the interest was in conducting a microanalysis to reveal the underlying structural organisations of naturally occurring interaction in L2 video-mediated classrooms. As noted, CA follows an emic perspective which enables the generation of empirical evidence that is entirely based on the data and the participants' orientations, which increases the level of validity. Additionally, using the CA's next-turn proof procedure adds to the robustness of the analysis and makes it available for the readers to test and challenge.

As will be shown in the analysis chapters, CA's analytical tools enabled the provision of detailed empirical evidence of the sequential moment-by-moment unfolding of the participants' use of the audio activation/deactivation features of Zoom in the management of turn-taking and repair. Further, using CA's analytical tools combined with the notion of affordances (Hutchby, 2001a, 2014) enabled the explication of how the participants' use of the technology features afforded or constrained their interaction. Also, it enabled the demonstration of the participants' use of the audio activation/deactivation features beyond their intended design. CA's analytical tools also enabled the unpacking of the participants' coordination of actions in L2 video-mediated interaction and how fellow participants adjusted their conduct accordingly. Moreover, CA's analytical tools enabled the presentation of empirical evidence relating to aspects of preference organisation, such as preference for progressivity and for self-repair.

CA has proved its usefulness in examining video-mediated settings by providing valuable insights into how interaction is organised in multiple settings, as previously shown in Section 2.3. However, there is no perfect methodology that does not receive criticism or suffer limitations. The following section presents some of the main criticisms of CA as a research methodology and demonstrates how CA practitioners have treated such issues.

3.8 Limitations and Criticisms of CA

As with any other methodology used to study social interaction, CA has faced a number of criticisms. CA has been criticised for that readers are only able to see data transcription to represent the data in published works. However, this drawback is not an issue anymore, as researchers have begun to provide their recordings along with the articles, as seen in the study by Mondada (2019). A reader can easily view video recordings and gain a deeper understanding of the phenomenon being analysed, and in so doing, they can validate or challenge the analyst's work. Also, transcripts are also accompanied by still images to illustrate the interactants' embodied practices where relevant. As aforementioned in Section 3.6, CA practitioners present parts of their data at data sessions and conferences, along with their analysis and results to be scrutinised by experts in the field.

Moreover, CA has been criticised for a focus on or preference for spoken interaction over other elements of interaction, such as gaze, body movement and other embodied practices. Looking back at the inception and early history of CA, it is possible to note that the originators benefitted from the technology available at that time, namely audio recording devices. This was a great step forward compared to note-taking or having to depend on what

the researcher or the participants could remember and report. As technology advanced and the recording of videos became possible, CA practitioners (see, for example, Goodwin, 1986, 2000; Heath, 2002; Mondada, 2006, 2007) began to make use of this new technology, and could thus take into account bodily conduct and use of objects in the environment.

In regards to using CA to examine video-mediated interaction, researchers may face a number of challenges, which result from the physical dispersion of the participants (Due & Licoppe, 2021). This can contribute to creating challenges related to data collection, transcription and analysis (ibid). However, as Section 2.3 demonstrated, previous research that employed CA to examine L2 video-mediated interaction revealed significant findings. In the current study, a number of challenges regarding the aforementioned areas were faced. The next chapter presents the processes of data collection and analysis and highlights such challenges.

3.9 Summary

This chapter has introduced the methodology that guides this research project. Section 3.2 revisited the purpose of the study and introduced the research aims. Section 3.3 introduced CA, covering its basic principles and how the advancing technologies have helped widen its focus and placed more weight on analysing multimodal interaction. Next, Section 3.4 provided an overview of the ethnomethodological foundations that have influenced CA's approach to the study of social interaction. The chapter then turned to present the interactional machineries that organise the production of social interaction in Section 3.5. Section 3.6 discussed the reliability and validity of CA research and how this study seeks to meet rigorous standards. Finally, longstanding and common criticisms of CA and ways of addressing them were presented in Section 3.7. Following the CA methodology, the next chapter presents the design of the current study.

Chapter 4. Research Design

4.1 Introduction

Chapter 3 presented CA, including its methodological principles and theoretical stance. The current chapter moves on to present the research setting, the process of data collection and analysis using CA. According to Bryman (2015), the term 'research design' refers to the "structure that guides the execution of a research method and the analysis of the subsequent data" (p. 40). The chapter is organised as follows: the first section (4.2) presents the research setting that covers the Zoom platform and the nature of the classes recorded. Section 4.3 introduces the participants; Section 4.4 describes the process of data collection; and Section 4.5 presents the ethical considerations this study considered. Next, Section 4.6 presents the process of data transcription and presentation. Finally, Section 4.7 presents the procedures for data analysis.

4.2 Research Setting

The context for the current study is small-group, online, L2 video-mediated speaking classes, hosted on the Zoom platform. This section is further divided into two subsections: the first presents the platform and elaborates on how it works; the latter provides a view of the nature of the classes recorded in this study.

4.2.1 The Platform

When analysing online video-mediated data, a good understanding of how the technology works can be highly beneficial to both the analyst and the reader. Hence, this subsection offers an overview of the Zoom platform and how it works. Zoom (www.zoom.us), founded in 2011, is a platform that offers an opportunity for participants to communicate online using video. It also enables the participants to text each other using the chat box provided. In addition, the participants can share their screens with the other participants to allow easy collective access to the materials that may be central to the activity, such as presentation slides, and videos.

Zoom equips the participants with a number of features that enable them to manage their interaction. The host (usually the teacher in the current study) has access to different managerial features, such as disallowing the learners to immediately access the meeting room by placing them in the waiting room, inviting new participants, assigning co-hosts, placing the

participants into breakout rooms, and - of special importance to this study - the ability to activate/deactivate the audio of self and others. It is also of importance to this study to note that the learners can also have access to some of these features, such as the ability to enable/disable video, activate/deactivate their audio, and share their screen with the rest of the class.

Figure 4.1 below is a screenshot from the data of the teacher's (host) screen showing the different features available to manage the interaction. Before elaborating on the different elements in Figure 4.1, it is worth noting that the teacher, in this instance, was on gallery view without sharing his screen. The different users' preferences can yield different screen layouts, as will be shown later. In Figure 4.1, the numbers 1-7 are used in the description to ease the reference to the different features.

Number 1 refers to the control bar located below the participants' videos in both the computer and mobile applications. This bar includes buttons through which the teacher can invite new participants, create polls, share their screen, open the chat box, place learners in breakout rooms, or send reactions to the learners' contributions. Of special relevance to the current study, number 2 highlights the window where the teachers can see a list of the participants' names and the status of their audio and videos. Number 3 shows the teacher's self-audio/video activation/deactivation buttons. These self-audio/video activation/deactivation buttons are similar to what the learners have on their screens, whether they are connecting from a computer, tablet, or mobile phone. Number 4 demonstrates the chat box, where the participants can communicate with the group or privately with another participant by typing a message (number 6) and sending it. The chat box is used in some classes to share links to Google Slides, discussion questions, share files, or, on some occasions, spell new words. Number 5 displays the view mode that enables the participants to choose how to view others on their screen. The different layouts the platform can take will be explained below in Figure 4. 2. Number 7 shows the green box that appears around the current speaker's image on the screen.

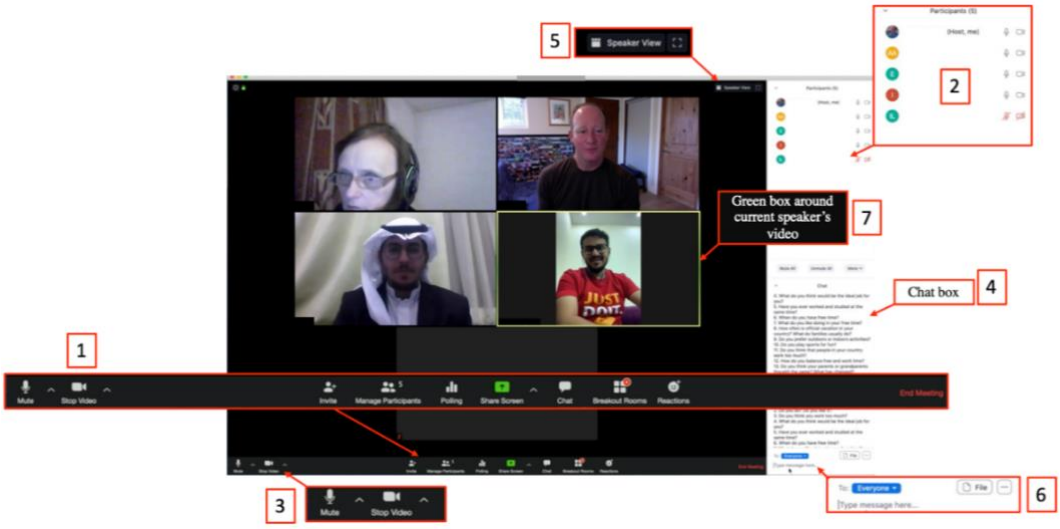


Figure 4.1: The host's (Teacher's) Options to manage Video Calls on Zoom.

Zoom can take different layouts on a screen depending on each user's preference. Figure 4.2 shows screenshots taken from the data collected for this study, which illustrate these different layouts. Users can choose to use either the gallery view or the speaker view to determine how to see the other participants' videos on their screens. In the gallery view, the participants are placed in a grid pattern which displays the thumbnails of other participants. When a participant has an active audio (current speaker or having audible noises, as will be seen later in the analysis), they are highlighted and made recognisable to others by a green box around their thumbnail. Also, in some cases when there is an overlap, one of the participants will have the green box around his/her videos, while the other participants' videos will be underlined by a green line. In the speaker view, the current speaker is shown in a large video window, while the other participants are shown in smaller windows at the top.

When a participant (usually the teacher) shares the screen, the layout can take different shapes on the screen as well. In the Speaker mode, the shared content takes the larger portion of the platform interface and only the current speaker's video or image, or the one with audible noises, appears on the screen. The users are afforded the choice of gallery side-by-side view to split the interface between the participants' thumbnails and the shared content. In the full-screen standard view, the participants' thumbnails are moveable and can be placed at different spots on the screen depending on the user's preference. The same movability feature applies to the control bar, as will be shown later in the analysis. Also, the platform window, when not in the full screen view, can be placed anywhere on the screen for computer users, and the control panel for the hosts can be dragged to different places when screens are shared.

Furthermore, to enhance the representation of the data in the analysis chapters, the readers will be informed of whether or not there is screen sharing by any of the participants, as this may have an impact on the interaction occurring at that particular moment.

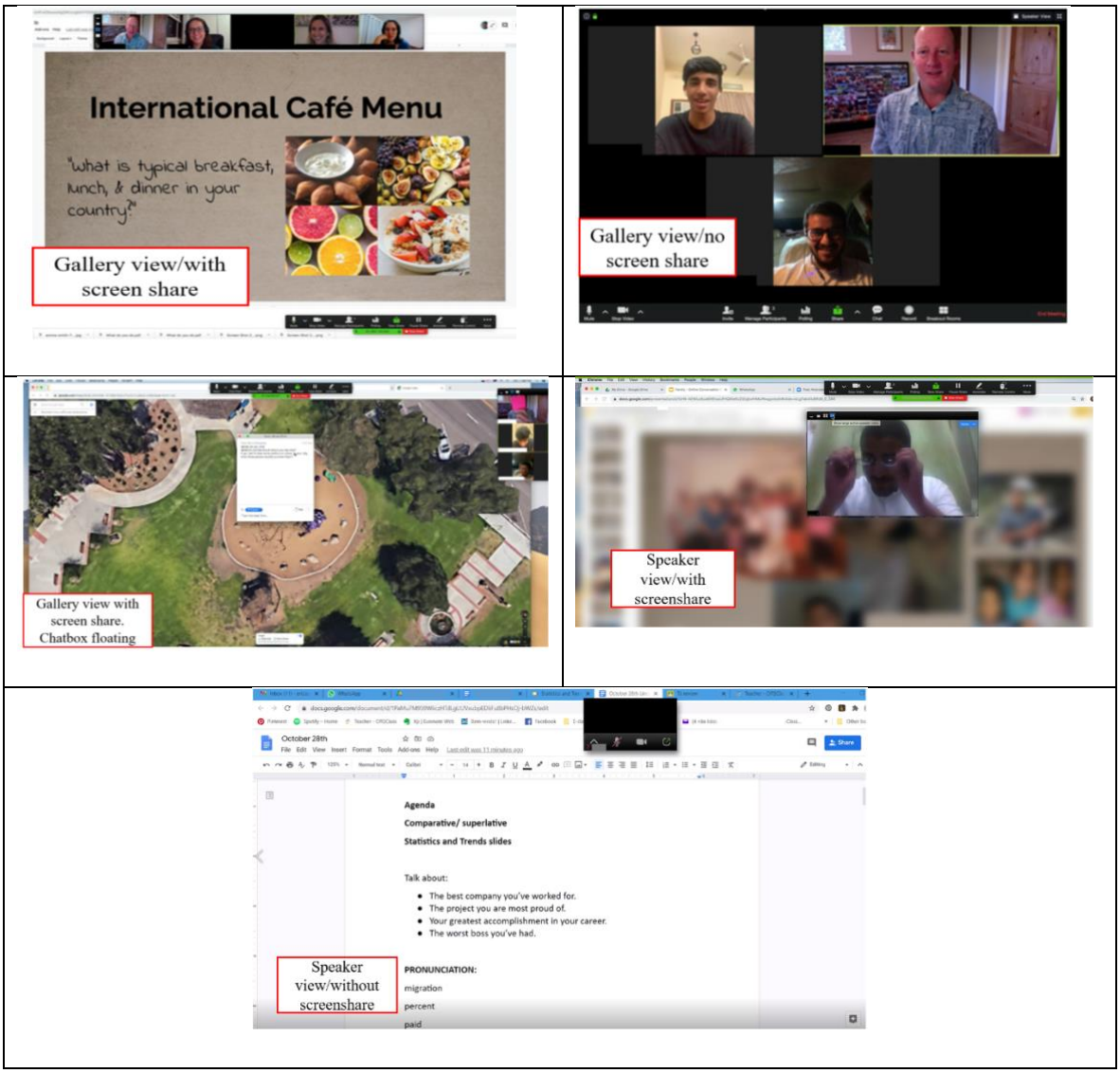


Figure 4. 2: Different Zoom Layouts of the Data, Depending on Users' Preferences

The platform's features that are of special interest to the current study are the audio activation/deactivation. These features are available to both the host of the meeting and the participants. The difference lies in the fact that there is an ability to deactivate the audio of others exclusively available to the host. A participant who self-deactivates their audio will have a red microphone crossed by a red slash next to their names, making his/her status available to the other participants to see (see Figure 4. 3).

Figure 4. 3: The Audio Status Next to The Participant's Name

When the host deactivates a participant's audio, the participant receives an on-screen notification informing him/her of the host's decision (see Figure 4. 4); in addition to the appearance of the red microphone icon. The host can access these features by either clicking on the mute icon on the top right corner on the participant's thumbnail or by clicking on the microphone icon next to his/her name on the participants' list. The participants can activate/deactivate their audio by clicking the microphone icon in the bottom left of their platform interface. They can temporarily activate the audio by clicking and holding the space key.

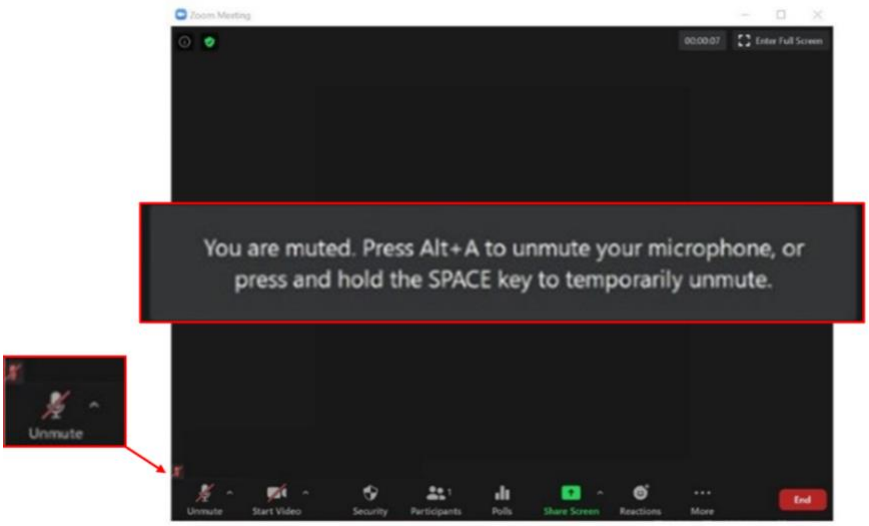


Figure 4. 4: On-screen Notification Informing a Participant that They are Muted

Following the introduction of the different features of Zoom and how it works, it is important to elaborate on the nature of the classes recorded in the current study.

4.2.2 The Classes Recorded in the Study

The classes recorded for the current study are speaking classes, with or without a focus on grammar. The classes revolve around speaking about topics that address cultural and daily life activities, such as work and play, home, justice, and other current world events. In some of these classes, teachers used other platforms in conjunction with Zoom, such as shared Google Slides and Google Maps to aid participation in class activities and provide materials for discussion, such as family photos, favourite places, or national costumes. The number of learners attending each class ranged from 1-12. This variation in the number of attendees mainly depended on the time of the class. As mentioned, the learners come from different countries, and so time zone differences played a significant role in preventing all the students from connecting altogether at one time. For example, all classes were between 8:00-15:00 PST, so classes between 8:00-11:00 were suitable for learners from the Middle East because their time zone is +8 hours. In comparison, it was not a suitable time for Honduran learners because the time zone difference is +3 hours, and they would be attending classes at their university. Another issue that played a role in this variation was connectivity for some of the learners. In addition, the lack of the technical knowledge necessary to join and participate in these classes led to a fluctuating number of attendees in some cases. The classes target small group discussions to allow more speaking time for the learners. To this end, when the number of learners exceeded 4, the teacher placed them into breakout rooms after the task's pedagogical and technical requirements had been explained.

The teacher(s) began each class by explaining what the class would be about and introducing the assisting tools (e.g. Google Slides, Google Maps, or Mentimeter) that would be used. Following this, the class then began the discussion following the points illustrated, either on the shared document or the chat box (see Figure 4.5).

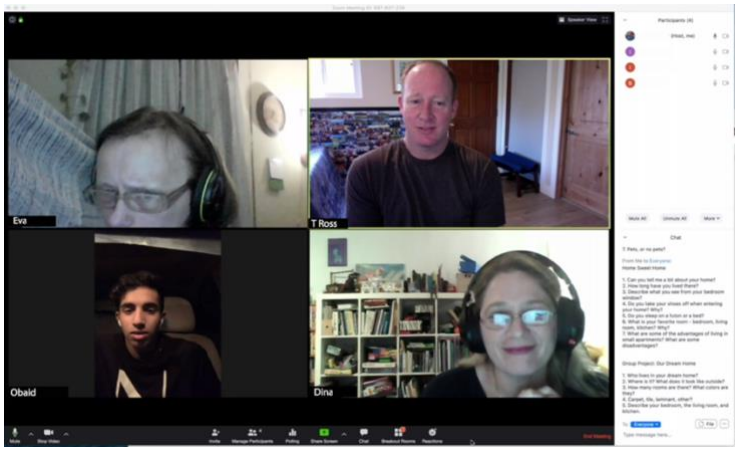


Figure 4.5: A Screenshot of the Discussion Topics in the Chatbox at the Beginning of Each Class

4.3 The Participants

A total of 30 L2 learners from different countries and ages (all aged 18+), as well as five teachers, participated in this study. The learners originated from different countries, including Brazil, Honduras, Saudi Arabia, and Japan. These demographic details were collected from the participants' conversations. Their language proficiency levels ranged between A2 to B2 according to the Common European Framework. The learners' language levels were not obtained via a placement test, but were estimated by the researcher based on the classes they were attending. Some of the classes were labelled, A1-A2 while others had B1-B2 labels on them on the website. On some occasions, learners from lower levels attended those of the higher level and vice versa. This classification is for the purpose of providing the reader with an indication of the learners' language levels. The majority of the learners joined these classes in conjunction with their morning classes at university. Some other learners joined these classes to improve their English for work-related purposes, while others sought an opportunity to practice speaking the target language. Again, these details regarding the learners were learned from their conversations with each other, which made them available to the researcher.

The lead teacher in these classes was from the USA, while the other teachers were from Brazil; all the teachers have had the experience of teaching online before these classes. Nevertheless, their expertise was mainly in teaching one-to-one classes, and leading language small group classes was a new experience for them. Following the introductions of the platform, the nature of the classes and the participants, the next section presents the process of data collection.

4.4 The Process of Data Collection

Obtaining or recording naturally occurring interaction is a basic starting point in CA (Liddicoat, 2022). Naturally occurring data refers to any interaction that is 'non-experimental' or provoked by the researcher (Ten Have, 2007). The main purpose for collecting data in EMCA research is to capture the interaction of all the participants involved as it unfolds in real time. A central source of multimodal CA data is video recordings, as found in the key works of Goodwin (1994) and Luff, Hindmarsh, and Heath (2000). Mondada (2012) indicates that video recordings enable the researcher to capture the whole participation framework. In the context of the current study, where the interaction is video-mediated, the screens are the sites where the interaction occurs. To elaborate, the participants can see each other's videos on the screens and, thus, screen recordings become essential to the analysis of video-mediated

interaction. According to Arminen et al. (2016), it is important to "take advantage of the embedded recording potentialities of the medium used for the interaction during data collection" (p.305).

The next section considers the methodological frameworks used to inform the data collection for the study. The data collection tools are also presented with the data collected elaborated upon.

4.4.1 Data Collection Tools

Recording screens is practically the closest option possible to observe the participants' perspective in interaction. There were two options to obtain recordings of the interaction in these L2 video-mediated classes. Firstly, Zoom has its own local screen recording feature, which is available to the hosts of the meeting (teachers in this case) and any participant that the host assigns as a co-host. This recording feature enables users to record meetings either in a local computer or to the cloud. All it requires the users to do is to click the 'record' button on the control bar and the platform will automatically record the meeting and upload it to the Cloud. Users are then able to download the meeting with multiple recording layouts, including speaker view, gallery view or screen-shared screen. However, these recordings can only capture the platform's window and there is no access to recording the full screen. This means that the recording will not capture the teachers' use of the control bar features or other elements, such as mouse cursor movements or other activities sites, such as Google Slides. The second option was to use an external screen recording software to capture all the activities on the teachers' screens. For this feature, the decision was made to use this tool for data collection. Two options of screen recording software were selected for testing (Camtasia and Snagit). Testing showed that both were stable and easy to use, but Snagit was lighter and required less hardware power.

Of course, the option of using an external software recorder comes with its own challenges as well; for instance, training the teachers to use them, uploading recordings to the internet and sending them afterwards to the researcher. Such a method means that certain technical requirements needed to be met, such as sufficient hardware to handle the multiple tasks and a good internet connection to upload the large data files to the Cloud afterwards. To address these requirements, the teachers were trained on 'Snagit' screen recorder software to capture their screen activities. As noted, Snagit is lighter and does not require high hardware capabilities. The training was conducted remotely on Zoom, as the teachers and the researcher were located in different geographical locations. After the training, a running test was

conducted a week prior to the beginning of the data collection to ensure that everything was running smoothly, and technical support was provided by the researcher as and when required.

Recording learners' screens would have given access to the data from multiple angles. However, this was a challenging task to accomplish, as the learners were located in different countries. This brought its own challenges, as some of the learners could not have access to a stable internet connection all the time, which imposed a tremendous burden for them to upload large files of high-definition video recordings to the internet. Further, some of the learners did not have sufficient hardware, which would make running Zoom and Snagit simultaneously a difficult task and might hinder their attendance and participation. Nevertheless, the teachers' screen recordings were "good enough records of what happened" (Sacks, 1984, p.26), despite the lack of access to all the other participants' screen activities.

4.4.2 Data

As stated above, the data recorded for this study were screen video recordings of the teachers' screens in addition to 2 breakout rooms recorded by the researcher. There are two data sets in this study: the first data set was collected between October 2019 and February 2020, with each class lasting for about 50 minutes; the second set of recordings was collected between October 2020 and November 2020, resulting in a total of 32 hours of video recordings collected.

4.5 Data Transcription

In this section, the process of transcribing the data will be explained. In CA, data transcription is seen as a core procedure of analysis and is considered an early step to enable the analysis of the data following CA methods, and it is also viewed as part of the analysis itself (Hutchby & Wooffitt, 2008). To ease the organisation of the data, the video files were imported into 'Transana' (Woods, 2021), a software used for managing, coding, and transcribing audio/video data (see Figure 4.6). Transana enables the users to organise videos into libraries to ease the location of different datasets. The user is then able to watch the videos repeatedly and control the playback (i.e. move forward/backward, stop, play) from the transcription window using keyboard shortcuts. Also, Transana displays the audio in waveform, which makes it easier to track rising/falling intonations and measure pauses and gaps. Regarding the transcription on Transana, the user is able to control the video while transcribing using the keyboard shortcuts, take and include screenshots in the transcripts and insert timestamps. The

transcription window is also equipped with the feature of adding some of the transcription conventions, such as rising/falling intonations, aspirations, and in-breaths, among others.

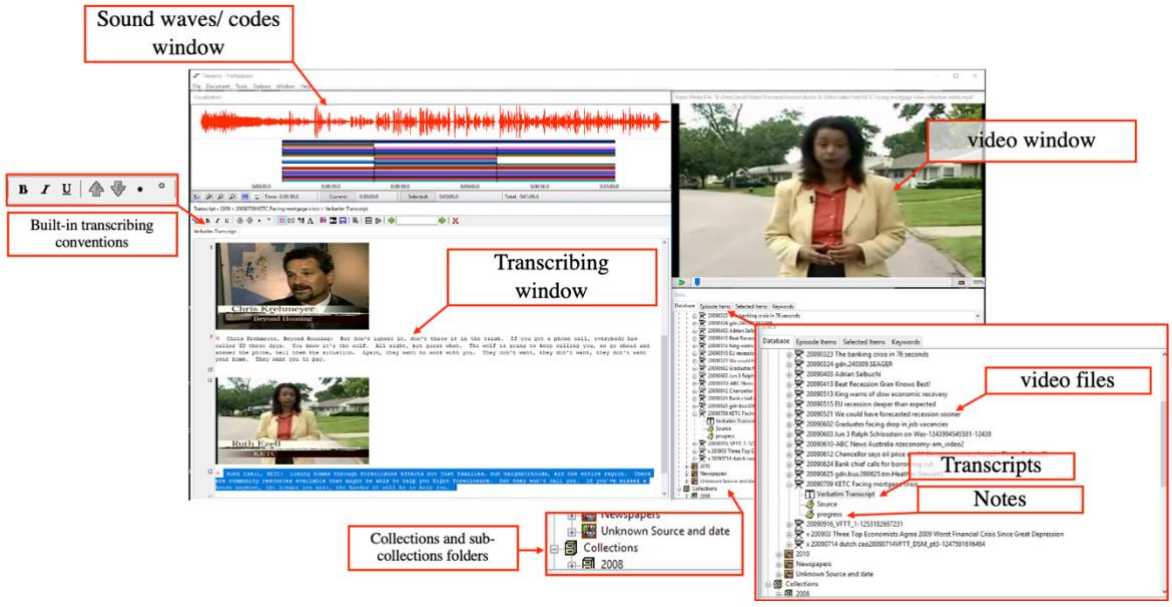


Figure 4.6: A Screenshot of the Transana Software Interface (Woods, 2021)

A core activity of doing CA is careful repetitive watching of the data collected in order to produce detailed accurate transcripts that capture, not only "what has been said, but also how it has been said" (ten Have, 2007, p.94). To this end, the video files were watched repeatedly and the initial transcripts were produced using Transana. The initial version of the transcripts was then exported to a Word document and placed into tables to preserve its format and representation after adding screenshots and other illustrations. It is worth noting that multiple refined versions of the transcripts were produced following repeated viewings of the data or feedback from supervisors, data sessions, and presentations.

Following the current study's focus on studying the participants' use of audio activation/deactivation features, the analysis of such subtle practices makes necessary the production of transcripts that present the finest details of the sequential unfolding of these practices. To present the systematic and orderly character of the participants' use of these features to organise their interaction online, the Jefferson transcription system (Jefferson, 2004) was used to transcribe the verbal conduct, combined by Mondada (2018) transcription conventions for non-verbal conduct and to display their temporalities. These conventions were also accompanied by other additions by Balaman (2016) to capture the participants' screen activities, such as clicks and mouse cursor movements. Additionally, new additions were created for the clear illustrations of the participants' use of audio activation/deactivation features (See all transcription conventions in Appendix A: Transcription Conventions).

Another level of refining the transcripts is to add some of the non-verbal conduct, such as gaze shifts and mouse cursor movements, and more importantly audio activation/ deactivation use. This was done using VLC media player, as it offers the feature of slowing down playback, which enables the analyst to track precisely the trajectory of different elements of the phenomena under examination. This is of special importance to the current study, as will be observed in the analysis chapters, where mouse clicks on audio activation/deactivation become relevant.

Furthermore, anonymising the data in both the data sessions the analysis chapters was carried out using Adobe Premier pro-video editing software and other programmes, while Adobe Illustrator and PowerPoint were used to capture some of the details, such as mouse cursor movements and clicks in the written analysis. Moreover, they were used to anonymise and zoom in on parts of the screenshots in the transcripts to capture the change in audio activation/deactivation status. Having presented the process of transcribing the data, the chapter now introduces the procedures used to analyse the data.

4.6 Data Analysis

The repeated and careful viewings of the data recordings and note-taking following 'unmotivated looking', which continued for several months, paved the way for the identification of the unique interactional phenomena. The initial focus was on how the teachers repaired speaking/hearing background noises caused by background noises and other reasons using audio activation/deactivation features. The repeated viewings of the data revealed another recurrent use of these features by the learners(i.e. to organise turn-taking). Once the use of these features was established as playing a role in organising repair and turn-taking, it was necessary to examine who uses them and where these practices are positioned in the sequences.

In the repair sequences, it was important to look at the timing of the trouble source's occurrence (background noise, for example) and to examine when the participants oriented to them as problematic. Similarly, it was also important to track the trajectory of the repair sequence, i.e. who initiated repair and how they designed their repair initiation. This stage included the use of the audio activation/deactivation features accompanied by verbal means in some cases to identify the source from where the disruptive noises were emanating from. The next step was to investigate how the trouble was resolved, which in most of the cases entailed deactivating the audio feed of the participant with the disruptive noise. Finally, the analysis demonstrates how the participants resumed the suspended class activity.

Regarding the analysis of the participants' use of the audio activation/deactivation features to manage turn-taking, it began by tracking down the sequential positions at which these features were used. Two sequential positions emerged in the data, whereby the participants activated their audio to self-select for the next turn at talk. Another two positions for when the participants deactivated their audio were discovered. The analysis of these positions was conducted with reference to the ongoing participation frameworks to illustrate the systematic character of the participants' use of these features to access interaction. This quest also led to the observation of the participants' accountability of the delay or the absence of audio activation by their fellow participants. The next step was to set the procedures in relation to how to present these analyses to the readers in a readable manner. How the data is presented to the readers is illustrated in more detail in the following section.

4.7 Data Presentation

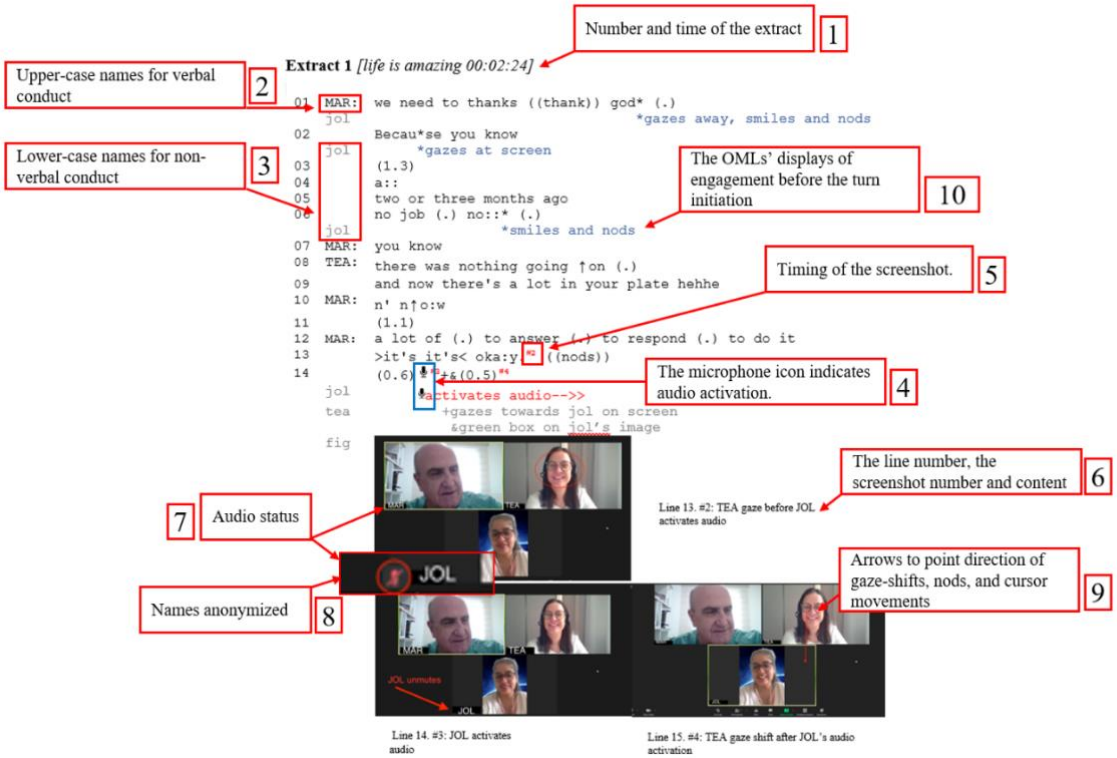


Figure 4.7: Illustrations of the Transcripts in the Analysis Chapters.

Figure 4.7 is a screenshot of one of the transcripts presented in the analytical chapters. The numbers on the figure ease the navigation of the description of how the data is presented to the readers. Number 1 shows the extract number and its timing in the class, which is useful for displaying to the reader the duration of a certain sequence or parts of a sequence. This is beneficial in the encounters analysed in Chapter 6, where longer extracts are divided into smaller ones to ease their presentation. Numbers 2 and 3 illustrate the names of the

participants and their verbal and non-verbal conduct, with the upper-case names preceding the verbal and the lower-case names preceding the non-verbal. It is also worth noting that non-verbal conduct was placed according to its occurrence in order to present it as close as possible to the data. Number 4 shows a microphone icon, which represents the audio activation action by the participants and shows the position at which it occurs. The audio deactivation is presented by another icon that shows a microphone with a slash on it to enact the icon shown on the platform. The audio activation/deactivation use is typed in red to highlight them in the transcript, as they are the main practices examined in the current study.

Screenshots are presented to provide the readers with as much access to the data as possible and to provide them with a holistic view of that data. They are also used to illustrate certain non-verbal conduct in the transcripts (i.e. changes in audio status, displays of engagement and attention such as smiles, fixed gazes, etc.). Number 5 shows the exact position at which the screenshot was taken during the talk. The screenshots in the transcripts are accompanied by illustrations next to or below them to indicate the line number on which it was taken and a note of their content (number 6). Numbers 7 and 8 show the audio status displayed next to the participant's pseudonym. Number 9 displays some illustrations that are used to mark different non-verbal elements, such as arrows for eye gaze movement, cursor movement, double yellow circles for clicks, or red zooming in circles for changes in audio status. Finally, number 10 shows the OMLs' displays of engagement during the turn in progress before activating their audio and launching the next turn.

Having presented the procedures for the data collection and analysis in this study, the following section sheds light on the steps taken to ensure the full adherence to the research ethics and regulations.

4.8 Ethical Considerations

The current research adheres to rigorous research ethics and regulations. Permission to conduct the study was granted by Newcastle University. All the teachers and learners who participated in this work were sent consent forms and information sheets prior to the classes commencing (see Appendix B: Information sheet (Teachers)Appendix C: Information Sheet (Students) and Appendix D: Informed Consent.Appendix A: Transcription Conventions The participants all voluntarily agreed to participate in the study. The consent form contained options regarding the appearances of the participants' faces/audio in the data presentation in the study, data sessions, and future conferences or publications. The participants had the freedom to choose whether to have their faces shown clearly or blurrily in the study, data

sessions, conferences and future publications; the same options were provided for the audio recordings. In addition, at the beginning of each of the classes, permission to record was orally obtained from the participants.

The researcher's presence in these classes was very limited. The researcher attended some classes when there was a need to record the breakout rooms. During these classes, the researcher's audio and video were deactivated and there was no contact with the participants, in order to minimise the 'observer's paradox' effect (i.e. the notion that participants act differently than they normally do because they are being observed) (Labov, 1972). Limiting the researcher's presence and also the use of screen recording software instead of Zoom's built-in recorder (which shows a flashing recording icon) were additional measures taken to obtain interaction that was as 'naturally occurring' as possible. Goodwin (1981) points out that participants "never behave as if they were unobserved; it is clear that they organize their behaviour in terms of the observation it will receive from their coparticipants" (p.44). Thus, being recorded or observed should not affect the way the participants go about organising their social interaction.

In the classes recorded for the current study, the participants' 'real names' appeared next to their images/videos. To maintain the privacy of their identities, all the participants' names in this study were replaced by pseudonyms. If any of the participants decided that they wanted their faces covered or blurred in still and moving images, this was respected. Also, their voices were altered or muted upon their requests. The researcher ensured that the participants' identities and all of their personal information would remain confidential and that they were used only for research purposes. All collected data were kept confidential and stored on a password-protected computer on the university's server in line with EU GDPR requirements. The findings will be presented using pseudonyms instead of the participants' real names, and no information will be included that could identify participants' identities.

4.9 Summary

This chapter has provided an overview of the study design. It began by offering a background of the context of the study in Section 4.2, which included an introduction to Zoom, its different features and how it works (4.2.1). As noted, sufficient knowledge of how the platform works can aid the analysis and understanding of the social phenomena observed. Rintel (2015) indicates that a good knowledge of the technology features is required to be able to build an analytical argument regarding technology-mediated interaction. The focus then

shifted to the elaboration on the classes recorded in this project (4.2.2). This included the nature of the classes and how they were carried out by the participants.

Section 4.3 Presented the participants in this study and offered an overview about their locations, language proficiency and other characteristics. Section 4.4 introduced the data collection process. It also presented the procedures taken to adopt a data collection tool and offered the reason behind choosing a screen recording software over Zoom-generated recordings (4.4.1). Additionally, the nature of the data recorded for this study and when it was collected are discussed in 4.4.2. Section 4.5 then demonstrated the process of data transcription, which includes introducing how the data was organised in Transana, and first version of transcripts was produced. Moreover, the process of refining the multiple versions of the transcripts and the procedures taken to ensure their accurate capturing of the examined social phenomena were demonstrated. Section 4.5 demonstrated how the data are presented in the analysis chapters. It showed the different illustrations accompanying the transcripts to ease the readability of them and to help the reader to gain as accurate as possible representation of the data. Following this, Section 4.8 introduced the ethical considerations that the researcher adhered to throughout the project. This includes gaining permission from the university and the participants, during data collection, presentation in the study and other venues, and maintaining the confidentiality of the data and participants' personal information following GDPR.

Having introduced the methods that were used for data collection, transcription, and analysis, the next three chapters present the analysis of observed social phenomena.

Chapter 5. Analysis: Using Audio Activation/Deactivation Features to Organise Turn-Taking

5.1 Introduction

The Zoom platform affords the participants the ability to deny or grant access to their audio to other participants. This can be done simply by using the audio activation/deactivation features of the platform. Some of the learners in the data choose to stay on-mute during the whole or parts of class time for a variety of reasons, such as engaging in non-class activities, having, or expecting to experience background noises. In the data in the current study, the on-mute learners (OMLs) oriented to having family members, pets, or TVs in their physical spaces on a number of occasions during the classes. Therefore, a potential reason for being on-mute is to adhere to the social norms of the online settings and stated class guidelines as noises caused by background activities can be consequential for the ongoing interaction (Jenks, 2014). Also, the guidelines of the classes recorded for this study clearly requested the participants to connect from a quiet place and that they were expected to deactivate their audio should they (expect to) have background noise(s). Thus, the OMLs needed to self-activate/deactivate their audio in order to gain access to the ongoing interaction. It should be noted at this stage that the OMLs are labelled as such only to ease the readability of the analysis and not to impose any category on them.

In this chapter, the OMLs' methodical use of the features of the online L2 video-mediated ecologies is analysed. In particular, the analysis focuses on their use of self- audio activation/deactivation features to organise their turn-taking, mainly for projecting self-selection for the next turn in an ongoing interaction by activating their audio and to mark their turn(s) as complete by deactivating their audio. Moreover, an analysis of some extracts to assess the consequentiality of the delay or absence of audio activation in these classes is also presented. It will be argued that there is a reflexive relationship between the participants' use of these features and the management of turn-taking in L2 video-mediated classroom interaction. Moreover, it will be argued that the delay or absence of audio activation can be consequential to the OMLs' ability to access an ongoing class activity.

According to Hutchby (2001, p.129), there is a difference between the developers' intended use of these features and the participants' actual use of them. The aim here is to explore the audio activation/deactivation *features-in-use* (ibid) by examining the orderliness and systematic character of their use in the organisation of turn-taking in online, synchronous video-mediated L2 classes as they unfold in the interaction. In order to do this, the analysis in

this chapter tracks the trajectory of using these features and their methodical use by the OMLs. This entails the location of the sequential positions of using these features and observing what actions they are used to achieve.

As previously mentioned in Section 4.2.1, Zoom can have different layouts on users' screens. When a participant (the teacher in this context) shares their screen, it is possible that the other participants are on speaker view, and thus, they may no longer have access to each other's videos on screen. The speaker view constrains the participants' access to the videos of fellow participants, and only the participants with active audio (either as the current speaker or any other participant with audible background activities) at a particular moment will appear on the screen next to the shared content. This may render the visual indicator of the audio status change, represented by the appearance of a microphone with a line crossing it next to the participant's name, to become invisible to the others in the class. Hence, the audio activation /deactivation actions can become invisible to the other participants (see Figure 4. 4, Chapter 4). This, however, remains a speculation, as the data does not contain recordings of the learners' screens. The participants could enable the side-by-side gallery mode to view all the participants, but this is unknown as the other participants, as well as the researcher, do not have access to this. To treat this uncertainty, only cases where no participant is sharing his or her screen are included in the analysis.

The chapter is divided into two main analytic sections. The first section (5.2) presents cases where the OMLs activate their audio to project their incipient speakership. It is further divided into two subsections based on the position of the practice: where the OMLs activate their audio and immediately take the floor (Section 5.2.1); and where they activate audio and delay the turn-initiation until the next possible transition space (5.2.2). The second section (5.3) presents cases where the OMLs return to their on-mute status after activating their audio to participate. This section is also divided further into two sub-sections:5.3.1 presents extracts showing the OMLs deactivating their audio immediately upon the completion of their turn, while 5.3.2 demonstrates a delayed position for audio deactivation to accommodate for further participation. Section 5.4 then illustrates the consequentiality of the delay of audio activation on the ability to access the ongoing interaction, while Section 5.5 presents extracts showing accountability for the absence of or delay in audio activation.

5.2 Audio Activation to Self-Select for the Next Turn

In this section, the focus is on how the OMLs use the Zoom feature of audio activation to project the possible completion of the current speaker's turn and their self-selection for the

next turn. The analysis will also show how the other participants recognise such use and adjust their conduct accordingly. The Sequences presented in this chapter are the ones where the non-primary speakers use the feature of audio activation to project their self-select for the next turn and how they display (or not) recipiency during the current speaker's turn. It is also important to look at the turn(s) preceding the audio activation to demonstrate the participants' close monitoring of the ongoing talk and, thus, projection of its completion.

The analysis in this section shows the two sequential positions of the OMLs' audio activation. In both positions, the audio activation is a pre-beginning activity (Schegloff, 1996). A pre-beginning activity is the conduct (verbal or non-verbal) that the incipient speaker produces to project the onset of their next turn as the current speaker's turn is reaching possible completion (ibid). The difference between the two positions lies in how long the audio activation is done prior to launching the next turn by the OMLs. The following subsection offers a number of extracts illustrating the first position, where the OMLs activate their audio immediately at the TRP and launch their turn. To ease following the exact positions of audio activation/deactivation in the extracts, they will be represented by the two symbols: (🎙) represents audio activation and (🎙) for audio deactivation. Moreover, the symbol (#+number) will be used to refer to still images in the extracts, as well as in the analysis.

5.2.1 Audio Activation at the Transition Relevance Place

Figure 5. 1 below is a post-analysis representation of the pattern this subsection presents. It shows the position whereby the OMLs use the audio activation feature to project self-selection for the next turn. As the figure shows, the OML is not selected by the teacher or the current speaker for the next turn. As will be shown in the following analysis, the participants orient to the ongoing participation framework by activating their audio upon the possible completion of the participation of the current speaker.

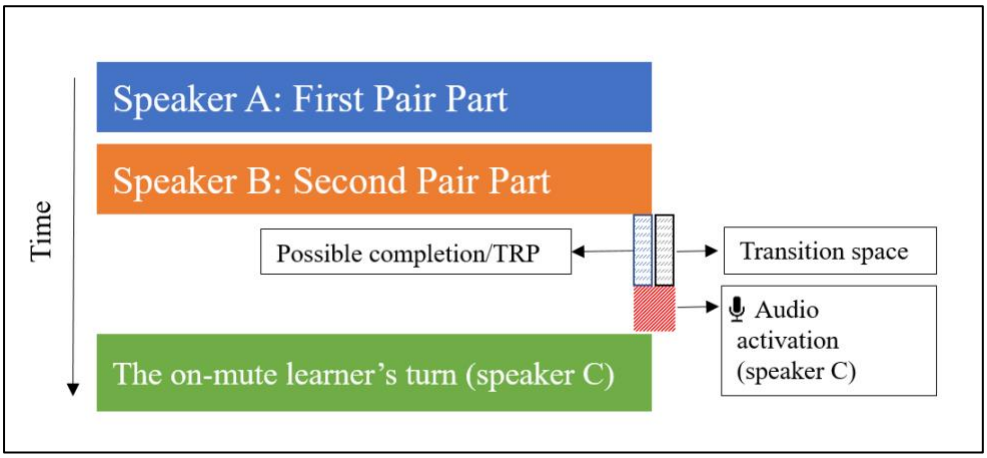


Figure 5. 1: Schematic Representation of the Audio Activation at the Transition Space

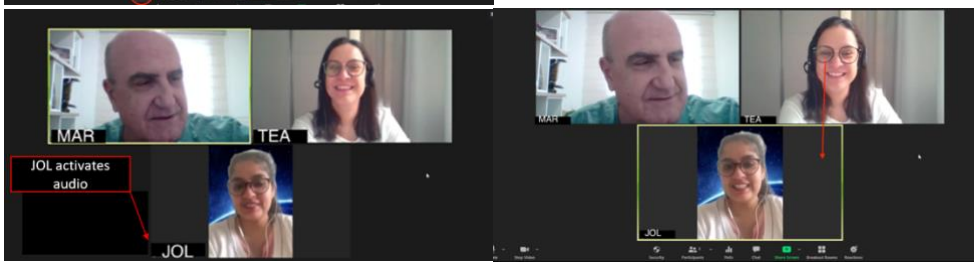
Extract 5.01 below is taken from a screen recording of a class that involves two learners and their teacher. It takes place in the opening phase of the class when the teacher usually asks learners about their week. Leading up to the extract, the teacher produced a first pair part asking MAR about his week and, thus, selected him as the next speaker. MAR answered that he was busy doing different jobs for his company. The interaction is entered as MAR extends his answer to the teacher's question.

Extract 5.01 *[life is amazing 00:02:24]*

```
01  MAR:  we need to thanks ((thank)) god* (.)
    jol                                  *gazes away, smiles and nods
02        Becau*se you know
    jol         *gazes at screen
03        (1.3)
04        a::
05        two or three months ago
06        no job (.) no::* (.)
    jol                 *smiles and nods
07  MAR:  you know
08  TEA:  there was nothing going ↑on (.)
09        and now there's a lot in your plate hehhe
10  MAR:  n' n↑o:w
11        (1.1)
12  MAR:  a lot of (.) to answer (.) to respond (.) to do it
13        >it's it's< oka:y.#1 ((nods))
14        (0.6) ^#2+&(0.5)#3
    jol         activates audio-->>
                 ^low background noise occurs-->>
    tea            +gazes towards jol on screen
                     &green box on jol's image
```

fig

Line 13. #1: TEA gaze before JOL activates audio



Line 14. #2: JOL activates audio

Line 15. #3: TEA gaze shift after JOL's audio activation

15 JOL: ↑that's life (.) ah ↑marco
16 (0.4)
17 TEA: ehh[ahha
18 MAR: [ye:ah ((smiles))
19 JOL: life's amazing ehheheh
20 MAR: yeah ((nods with a smile))

In line 1, as MAR explains the situation in a multi-TCUs turn, JOL appears to gaze away, smiles and nods to someone in her physical space. Due to the asymmetrical nature of these fractured ecologies, it is not uncommon in the data to see participants engage in side or parallel short activities with people or pets in their physical spaces (Ruhleder & Jordan, 2001b). JOL gazes back at the screen in line 2 and nods vertically with a smile indicating active listenership. A 1.3-second pause occurs in line 3 but as MAR's ("Because you know", line 2) projects more talk to come, as the reason is yet to be stated for why they need to thank God, the other participants do not take the floor. Indeed, MAR continues with his elaboration in lines 4-6. JOL further displays active listenership and engagement by smiling and nodding in line 6 as MAR is reaching a possible turn completion. MAR's discourse marker ("you know", line 7) marks an end to the description of how the situation looked two months before but also projects more talk. The teacher self-selects and comments on MAR's answer in lines 8-9 by providing a reformulated description of the situation and also a candidate completion of the answer. MAR uses ("n' n↑o:w", line 10) with a high stretched intonation projecting the pursuit of a comparison between the previous couple of months and the time being, which indicates more talk is coming next, and he, thus, holds the floor regardless of the 1.1-second pause in line 11.

After a 1.1-second pause, MAR continues with the comparison by listing the things that need to be done ("`a lot of (.) to answer (.) to respond (.) to do it`", line 11). MAR follows this up with a turn indicating that he is not complaining about the current situation of his job ("`>it's it's< oka:y`", line 13). MAR's turn at this point is recognisably complete (syntactically, prosodically, and pragmatically), and it is followed by a 1.1-second silence, which projects no further talk is coming. This creates a relevant transition place for the next speaker.

To sum up what has happened so far, the teacher produced a first pair part (a question), and MAR provided the second pair part (an answer). If the participation framework is considered here, it can be observed that the teacher's question prior to the extract has set it in terms of questioner-answerer(s) roles. Thus, JOL activates her audio in the middle of the 1.1-second pause right at the TRP indicating self-selection for the next turn in line 14 (See #2). The teacher's gaze shift[1] from the middle of her screen (#1) to gazing down towards JOL's video (#3) following JOL's audio activation may be indicative of the accountability of the audio activation action. The teacher's gaze shift, combined with a silence following JOL's audio activation, indicates that JOL is the possible next speaker. The trajectory of the audio activation use is well-coordinated with the current turn in progress and adjusted to the ongoing set participation framework. JOL begins by referring to MAR's complaint of the unexpected turn of events regarding the amount of workload. She designs her turn in a humorous manner in line 15 ("`↑that's life (.) ah ↑marco`"), which is met with laughter from the teacher and a smile from MAR.

Extract 5.01 demonstrates how JOL uses the Zoom feature of audio activation to project self-selection for the next turn. The timing of the audio activation is finely coordinated with the possible completion of the current speaker's turn. It is also accompanied by other visual displays of engagement, such as smiling, nodding and maintaining a fixed gaze at the screen. Even though the audio activation trajectory involves the arm extending to touch the screen, moving the mouse cursor towards the unmute icon, clicking the icon or even pressing and holding the space bar in the keyboard may not be visible to the other participants; this is still a pre-requisite that the OML needs to fulfil before activating his or her audio. Extract 5.02

[1] I am aware of the challenging nature of following eye gaze in video-mediated settings. However, this extract comes from the teacher's screen recording and the teacher's gaze direction can be roughly followed given that only three participants are in the room with the gallery view activated for the teacher.

below shows a similar example in which an OML activates his audio to self-select for the next turn.

Extract 5.02 *[removing shoes 00:37:27]*

```
01  MOH:  Yeah+ (.) cultura- a cultural thing
    mar       +nods-->
02        that (0.5) >saudi arabians< most of the time: (.)
03        a:m (.) remove their sho:es (.)#4 ↑out
    fig
```

Line 3. #4: MAR nods slowly

Line 5. #5: MAR nods faster

```
04        so they: they would be+ >you know< (.)
    mar                         +nods faster-->
05        free to move#5
          like mr ↑marco
06  TEA:  Yeah
07  MOH:  ^ahum::^
          ^a slow nod^
    mar           activates audio-->>
08        (0.6)
09  TEA:  &+#6a[hm:
          &gazes to the right side of the screen-->
    mar   -->+
10  MAR:      [her&e in #7brazil (.)
    tea       -->&shifts gaze to the left side and nods
11        there are ((is)) a culture
12        that use a small carpet (.) out of the house (.)
13        in order to clean the (.) the shoes
    fig
```

Line 9. #6: teacher gazes to the right.

Line 10. #7: teacher's gaze back to the left.

In this class, the teacher is leading the discussion by asking the questions placed in the chat window. Prior to the extract, the teacher had asked MOH if he considers the act of taking his shoes off before entering the house to be a cultural trait in his country; thus, selecting him for the next turn. The extract marks the beginning of MOH producing a second pair part to the teacher's question. He answers positively and takes a multi-unit turn (lines 1-5). MAR displays attention and active listenership by a slow continuous nodding during MOH's turn,

as notated in the transcript (lines 1-8) and in #4 and #5. MOH's turn reaches a possible completion in line 5, while MAR's nodding becomes more intense (faster), as shown in #5. The teacher follows MOH's answer by providing an assessment ('yeah', line 6). MOH marks his answer as complete using a stretched minimal response token a ('ahum::', line 7) accompanied by one slow nod stretched throughout his response token. At this possible TRP at the end of MOH's turn in line 7, MAR activates his audio. He then launches his turn and overlaps with the teacher's turn initiation in line 9. It is worth noting the teacher's gaze direction at the time MAR activates his audio; the teacher appears to look at the right side of her screen to where the chat box containing the questions is placed in the gallery view (#6 and #7). This gaze to the left side of the screen by the teacher occurred every time before she introduced a new question from the list in the chat window. Due to MAR and TEA overlapped turn initiation, the teacher gazes to the centre side of the screen and nods in acknowledgement of MAR's turn initiation, and thus, drops back to enable MAR's contribution.

To summarise the above analysis, MAR's activation of his audio channel and self-selection for the next turn comes after MOH's confirmation of the completion of his participation. The audio activation is conducted immediately at the TRP, which, in addition to displays of attention and active listenership, reveals MAR's close examination of the turn in-progress and orientation to the participation framework at play. The adjacency pair is complete and followed by the asker's assessment, which can project an open floor at that particular moment for any participant to take the next turn. This is exactly when MAR activates his audio and self-selects. Note that prior to beginning his turn he was displaying engagement and, possibly, preparing to self-select for the next turn by shifting from slow to fast nodding and activating his audio.

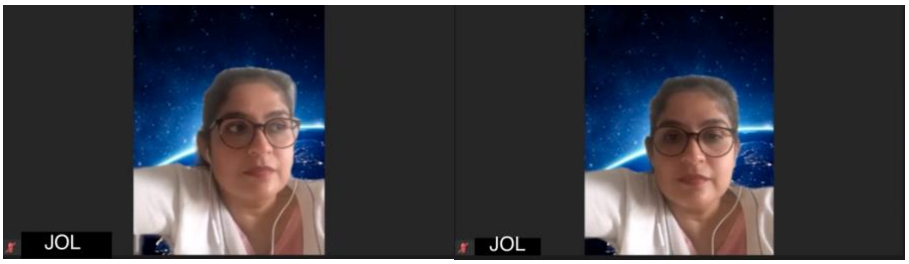
The OMLs may also activate their audio and self-select for the next turn without displaying much willingness to be the next speaker (Mortensen, 2009), unlike extracts 5.01 and 5.02. A close examination of Extract 5.03 illustrates this more.

Extract 5.03 *[re-election 00:17:07]*

```
01  TEA:  because our president is (.) friends with trump (.) nar
02  MAR:  Yeah
03        (0.6)
04  TEA:  >n he's gonna< try the re-election as well
05  MAR:  yeah(.)
06        to the fact (.)our election*#8 is on two thousand second(.)
    Jol                              *gazes away-->
07        Probably*#9
```

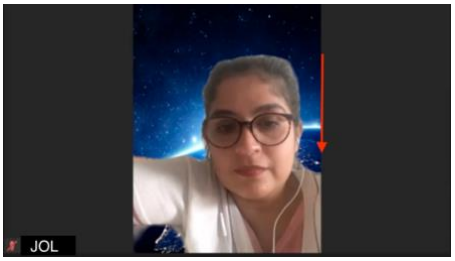
Jol -->*gazes back to the screen
08 VER: °a::h°*#10
*extends arm towards screen-->>

Fig



Line 6. #8: JOL gazes away

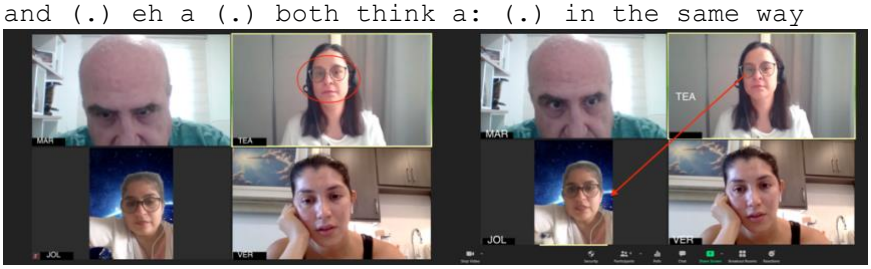
Line 7. #9: JOL gazes back to screen



Line 8. #10: JOL extends left arm towards the screen

09 TEA: yea:h
10 i agree&#11 (.)i agree
Jol activates audio
&tv noise appears-->
11 #12(0.9)
12 JOL: a:h↑ (.) me too
13 and (.) eh a (.) both think a: (.) in the same way

Fig



Line 10. #11: teacher's gaze before JOL audio activation

Line 11. #12: TEA gazes towards JOL

Prior to the extract, MAR and the teacher inform VER that they hope Biden wins the elections in the USA because they hope this will positively affect the elections in their own country. VER did not seem to understand the relationship between the outcomes of the two elections. This confusion makes relevant an explanation by MAR and the teacher. The extract marks the beginning of this information in lines 1-6. JOL gazes away in the middle of MAR's turn in line 6 (#8). So far, the conversation is between the triadic (TEA, MAR and VER), with the first two as information providers and the latter as the intended recipient. JOL is more of an overhearer in this instance (Goffman, 1981), as she is from the same country as the teacher and MAR. In line 7, MAR extends his turn to indicate uncertainty regarding the time of the elections in his country. At the end of MAR's turn, in line 7, JOL's gaze is back and fixed on the screen (#9). The explanation provided by MAR and the teacher makes a response by VER

relevant next, who is the recipient of that explanation. Indeed, VER self-selects next and provides an assessment (Goodwin, 1986), using a minimal response token to indicate a shift in epistemic status (Heritage, 1984a) ('°a::h°', line 8).

As VER's turn reaches a possible completion, JOL appears to extend her arm to the screen (#10). At this point of interaction, the adjacency pair is complete, and a mutual understanding is reached between TEA, MAR and VER, making it a possible transition place for JOL to self-select and access the interaction. However, the teacher also produces an extra assessment to MAR's explanation ('yea:h(.)i agree (.) i agree', line 9-10). In the middle of the teacher's assessment (at the possible TRP of 'i agree' TCU), JOL activates her audio. A TV background noise comes from JOL's background after her audio activation. The teacher shifts her gaze from the centre of the screen to JOL's image (#11 and #12). This gaze shift can indicate that audio activation to self-select is recognisable to the teacher. JOL immediately takes the floor in line 11 and accesses the interaction by expressing agreement of what had been discussed earlier ('a:h↑ (.) me too', line 12). As she is from the same country of the teacher and MAR, she has a higher epistemic position than VER. JOL further contributes to the explanation of the relationship between the two elections indicating that both Trump and the Brazilian president ('think a: (.) in the same way', line 13).

The analysis so far shows that the participants strategically use the feature of audio activation to project self-selection for the next turn. The participants' audio activation is finely timed with the possible completion of the current turn(s) in progress. The audio channel activation in this case can be seen as a pre-speech activity (Jefferson, 1983, p.14) or an activity that can function as a pre-beginning action (Schegloff, 1996) that can "project the onset of talk, or the beginning of a (next) [TCU] or a turn, but are not yet proper recognizable beginnings" (p.92). In a similar vein, Mondada (2007) points out that pointing gestures can be used by participants as resources to establish themselves as the next speakers. However, interacting in such fractured ecologies (Luff et al., 2003, 2016) in addition to the on-mute status of the learners in the data examined here, may render some of the pre-speech activities, such as throat clearing or turning one's head towards a potential recipient, insignificant. To handle this constraint, the OMLs in the current study use the platform features, such as audio activation accompanied (or not) by nodding, constant gaze shifts for a similar purpose, to indicate self-selection for the next turn at talk.

The analysis in this subsection also shows that the OMLs' use of the audio activation feature to project their incipient speakership is positioned immediately at the TRP of the current

speaker's turn. However, the data also shows another interesting pattern of audio activation concerning where it begins. The OMLs activate their audio way before the ongoing turn reaches a possible completion, but initiate their turn at the next possible TRP. In what follows, a number of extracts will be used to illustrate more on this feature-in-use.

5.2.2 Delaying Turn-initiation After Audio Activation

Using the audio activation feature to project incipient speakership can be positioned well before the current speaker reaches a turn completion. The OMLs activate their audio and initiate their turn at the next possible TRP. This entails monitoring the current speaker's turn to take the floor next. Figure 5.2 below is a post-analysis visualisation of this position for audio activation, which shows the audio activation occurs at some point during the current speaker's turn. However, the OML who is preparing to take the turn does not launch the turn until the possible completion of the current speaker's participation. Extract 5.04 provides evidence of how the participants use this feature to establish themselves as the next speakers at the next possible TRP.

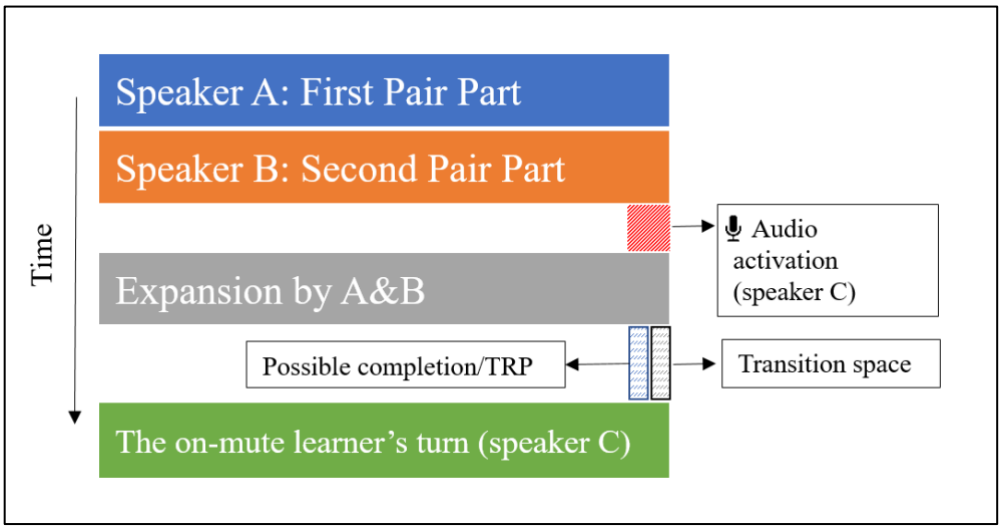


Figure 5.2: A Schematic Representation of Delaying Turn Initiation After Audio Activation.

Extract 5.04 *[sleep in a bed 00:40:35]*

```
01  ZAI:  of course
02        (0.5)
03        .hh living without a bed (.)
04        its look like that you are living without water
05        (1.0)
06  TEA:  °yeah°
07        >but some people< like to sleep on hammocks (.) you know
08        ^(0.7)
    pau   ^smiles and nods
09  ZAI:  ahmum:
10        +(1.1)+
```

```
    tea   +--nods--+
11  TEA:  I don't feel they're comfortable (.)
12        [and i'm too old for that
13  ZAI:  [ahm:: (.) they're not
14  TEA:  Ahhhh
15  ZAI:  &↑ouch [(.) ↑ouch
          &Mimecks having back pain-->>
    pau         activates audio-->>
16  TEA:          [ahhh
17  ZAI:  it ↑hurts^ (.) that's really#13 painful&
    pau           ^nods and smiles-->>
    zai                                    -->&
18        ↑ouch
19  TEA:  $that's true$
20  PAU:  ^somebody lives in trailer (.)
          -->^stops nodding
21        ↑you know
22  TEA:  ↑yeah
    fig
```

Line 17. #13: MAR nods and smiles while gazing at the corner of his

Extract 5.04 comes from a conversation class while working on a task about 'home'. Prior to the beginning of this extract, the teacher had addressed a question to ZAI about whether he sleeps in a bed, on the floor or in a hammock. The teacher's question makes relevant next an answer by the already established next speaker (ZAI in this case). Line 1 marks the onset of ZAI's second pair part (lines 1-5). This is then followed by an acknowledgement token by the teacher ('°yeah°', line 6). ZAI's answer suggests that it is impossible to live without a bed. This suggestion is then challenged by the teacher in line 7 which invites a response. PAU displays attention and engagement by continuous slow nodding. The form of the teacher's comment does not specify the next speaker, but it seems that ZAI and PAU take it as a follow-up question addressed to ZAI, who indeed self-selects in line 9 and produces a minimal response continuer (Schegloff, 1982) ('ahmum:'), followed by a 1.1-second silence in line 10, during which the teacher is nodding before she continues the turn ('I don't feel they're comfortable (.)', line 11). ZAI self-selects and overlaps with the teacher's turn initiation in line 13 while PAU continues nodding.

Both the teacher and ZAI produce overlapped complete utterances (lines 12-13). ZAI self-selects again and produces a verbal and visual mimicking of pain that can result from sleeping in hammocks in line 15. PAU activates his audio channel in line 15 at ZAI's turn initiation,

but does not initiate the turn just yet. If we examine ZAI's turn in line 15, it can be noted that it is not complete and it is also overlapped by the teacher's laughter. ZAI also extends his pain mimicking turn, as seen in lines 17-18, while PAU is nodding and smiling (#13). The teacher produces an evaluation to ZAI's answer in line 19, registering ZAI's answer as complete and thus creating a possible transition place. Indeed, this is where PAU stops nodding, self-selects, and initiates a turn ('*somebody lives in trailer (.) ↑you know', lines 20-21). The teacher acknowledges PAU's contribution immediately ('↑yeah', line 21).

Extract 5.04 shows an early position of the audio activation in the ongoing sequence, accompanied by visual displays of engagement by PAU, who closely monitors the turns in-progress. PAU also orients to the established participation framework in that he only initiates his turn after MOH's contribution is recognisably complete by the teacher's evaluation, which is the next possible transition place in this case. Extract 5.05 displays another occurrence where a OML activates her audio at an early position and initiates her turn at the next possible transition space.

Extract 5.05 *[he's won 00:07:10]*

```
01  MAR:  -about& (.) >you know<(.)the& election in the united states
    tea          &gazes at mar    &abandons minimising, retracts cursor
02  TEA:  oh my ↑go[:d
03  MAR:           [((inaudible))
04        (0.7)
05  TEA:  that's fierce* (.) ↑right
    jol                  *smiles
06  MAR:  he's- he's*#14 won
    jol              *extends arm to device screen
07        (1.0) +
    jol           activates audio-->>
                  +tv background noise appears-->
08  MAR:  he's#15 [gonna wine ((win)) it
    fig
```

Line 6. #14: JOL's arm in home position

Line 8. #15: JOL extends arm to the screen

Line 12. #16: TEA holds cursor on JOL's audio deactivation button.

```
09  TEA:        [>yeh<((opens eyes widely and mouth indicating
            surprize))
10  TEA:    he's going to win (.) ↑right&^
                                &places the cursor on jol's muting icon-->
    jol                                 ^opens mouth slightly-->
11  MAR:    °ye:ah°((nods))
12          (0.4)^#16
    jol       -->^
13  JOL:    bi↑den
14  TEA:    bi↑den
15  JOL:    a::h (.)&
    tea           -->& moves the cursor away
16          did you &guys watch the movie i sent to you
```

Extract 5.05 involves two learners (JOL and MAR), along with the teacher, while JOL's audio is deactivated. We can see MAR self-selects in line 1 and initiates a new a topic ('`-about& (.) >you know< (.)the& election in the united states`', line 1). The teacher gazes back to the centre of the screen after gazing up just prior to MAR's turn initiation. Even though the gaze movement can be insignificant in such fractured ecologies (Luff et al., 2003), gaze shifts coordinated with verbal conduct ('`oh my ↑go[:d`', line 2), as in this case, can be viewed as displays of attention. The teacher recognises MAR's initiative and engages further with it ('`that's fierce* (.) ↑right`', line 5). We can observe in line 5 that JOL is also displaying attention by smiling while gazing at the screen. MAR makes an announcement that one of the candidates has won the elections without mentioning the candidate's name ('`he's- he's won`' line 6). At the end of MAR's announcement in line 6, JOL extends her arm to the screen (#14 and #15). MAR's announcement is followed by a 1.0-second pause in line 7, after which JOL activates her audio. A TV noise from JOL's background is immediately heard in class after she has unmuted. Despite the activation of her audio channel, JOL does not take the turn just yet. MAR pursues his announcement ('`he's [gonna wine ((win)) it`', line 8), which overlaps with the teacher's turn in line 9. The teacher produces a statement with a tag question ('`he's going to win (.) ↑right`', line 10) using the pronoun (he), which displays an understanding of who is winning.

As the teacher's turn is reaching a possible completion, she moves the mouse cursor, places it on JOL's mute icon, and holds it there (Figure 16). At the teacher's possible completion in line 10, JOL opens her mouth slightly and maintains the position, possibly indicating she is 'gearing up' (Schegloff, 1984, p. 93) to take the next turn. Indeed, after MAR produces an answer to the teacher's tag question, JOL self-selects and launches the turn ('`bi↑den`', line 13). JOL's turn seeks confirmation of the identity of the winning candidate in the elections, whom the teacher and MAR are referring to. The confirmation comes from the teacher by

repeating the candidate's name (line 14). JOL acknowledges the confirmation ('a::h' line 15) and self-selects for further talk. At this point, the teacher drags the cursor away from JOL's audio deactivation button after she took the turn and became the current speaker. One may ask why the teacher held the mouse cursor instead of immediately deactivating JOL's audio. It seems that the teacher recognises that JOL's audio activation at that moment as projecting self-selection to take a turn in the ongoing talk. Also, JOL's slightly opened mouth can indicate that she is preparing to be the next speaker.

Extract 5.6 below presents another example of a participant's use of the audio channel activation to access interaction in the next possible transition place.

Extract 5.06 *[community 15:24:01]*

```
01   TEA:  seri(.)you ahm(.)you brought up a really interesting
           topic (.)
02         or an interesting ↑tangent
03         and >that's about<
04         what's that word
05         what's that word when
06         when you have a group of house:s (.)
07         a group of neighbours (.)
08         and everybody kind of gets alo:ng and shares ca:kes
09         >what's that< called* when you have that
     ter                        *smiles-->>
10   SER:  ↑we we are gathering (.)
11         #17very:*#18 often
     Ter           activates audio-->
     Ter         -->*
     fig
```

Line 11. #17 and #18: TER activates audio and audio status changes.

```
12   TEA:  ↓it's a gathering.
13   SER:  yes [gathering
14   TEA:       [a::nd >when you< (.)
15         when you have your* neighbourhood (.)
     ter                      *smiles-->>
16         a:nd the neighbourhood has lika:: (.)
17         What do ya you call that (.) when it's a neighbourhood
18         *&what's that English word am looking for&
     tea    &gazes down left and right at ter and ali&
     ter   -->*
19   SER:  $mmh$
20   TEA:  ^↓ali.
     ter   ^opens mouth slightly
21   TER:  community↓
22   TEA:  $↑community$
23   SER:  $community$
```

24 TER: hhh $ya::y$

The extract begins with the teacher producing a multi-unit question in pursuit of eliciting a candidate vocabulary item from SER. When a final version of the question is produced in line 9, TER begins to form a smile. The teacher's question selects SER as a recipient and, thus, the possible next speaker. Indeed, SER self-selects and provides an answer in lines 10-11. It seems that SER's answer does not provide the desired vocabulary item wanted by the teacher. This may explain why TER activates her audio at this particular point of interaction, as seen in line 11 (#17 and #18). It might be that TER is preparing to provide another candidate answer. The teacher's lower intonation while acknowledging the receipt of SER's answer can indicate that it is not the pursued one (line 12). The teacher then reformulates the question in lines 16-18. However, his gaze shifts from centre left (SER) to down left and right (TER and ALI) as the question is produced ('*&what's that English word am looking for&', line 18).

The question reformulation does not explicitly specify the next speaker. In addition, the teacher's gaze movement direction can project that the other participants can now try to answer.[2] TER is smiling again, which may indicate a willingness to self-select for the next turn, while ALI does not display much willingness to answer. SER minimally responds to the reformulated question ('mmh', line 19), indicating an inability to answer. TER opens her mouth slightly as if gearing up for turn initiation in an overlap with the teacher offering ALI the answering slot (line 20) in a falling question-like intonation. This is when TER initiates the turn by providing a candidate answer in a low questioning tone. The teacher repeats the answer in a high intonation close to a celebratory manner with a smile on his face in line 21.

The analysis so far shows that audio activation to access the interaction is preceded by visual displays of engagement and this action is well timed with the ongoing interaction. Moreover, the OML participant orients to the participation framework at that particular part of the activity by delaying their turn initiation. The technology features of audio activation/deactivation enable the participants to actively participate in class activities, while keeping their background noises or expected noises off the class shared space. However, simply activating one's audio does not guarantee successful access to the interaction. On the contrary, OMLs may miss the transition space, which might hinder their participation and,

[2] It is worth noting the asymmetrical nature of the setting that bars the participants from accessing each other's gaze direction. However, gaze movement to a different part of the screen in gallery view, accompanied by a question reformulation without specifying a recipient, can be recognisable by other participants as an open floor.

possibly, learning. This is because they typically have the extra task of audio activation before initiating the turn. This extra work can entail preceding steps that can include extending the arm to reach the screen, dragging the cursor to the audio activation button, or pressing the space bar on the keyboard. Other participants with already activated audio may be one step ahead of the OMLs and they may be able to self-select and initiate the next turn faster. The next section presents multiple occurrences to demonstrate this case.

5.2.3 Audio Activation Past the Transition Space

The OMLs strategically use the features of audio activation/deactivation to self-select for the next or subsequent possible turn(s), as shown in the analysis in the previous sections. The OMLs also use visual cues - nodding and smiling, etc. - to display engagement with the ongoing interaction, and they closely monitor the turn in progress before activating their audio. This, however, may not be enough because the *timing* of the audio activation is equally important in securing the interactional slot. Additionally, activating the audio can be seen as an extra task that the OML needs to accomplish before self-selecting for the next turn. Extract 5.07 below is an instance that can illustrate this idea. Prior to the extract, the teacher tells the learners about an old advertisement she saw on TV depicting how people used to have a single television in their home around which they gathered, but now they have multiple televisions that keep them apart. Line 1 is a continuation of the teacher's storytelling. PAU displays engagement by continuously nodding as the teacher describes the advertisement to the class.

Extract 5.07 *[bringing family together 46:43:00]*

```
01  TEA:  *^but then< (.)people started having their own t- tvs#19 (.)
    pau   >>* nods-->
    pau   >>^ gazes to left and right constantly-->
    fig
```

Lines 1 -9. #19 and #20: PAU gazes constantly to left and right sides of the screen

```
02        and#20 the idea of bringing family together was: (.)
03        a:α:
    zai     α moves head to left and right as saying((no))-->
04        (0.3)
05        [off
06  ZAI:  [αfinished (.)
          -->α
07        [no]thing+¤
```

```
    pau    --activates audio--
    pau                ¤raises eyebrows
08  TEA:   [yes
09  ZAI:   ^actually a am going to say something (.)
    pau    -->^
10         a are you finding (.)
11         do you find anything interesting at the tv this ↑time
```

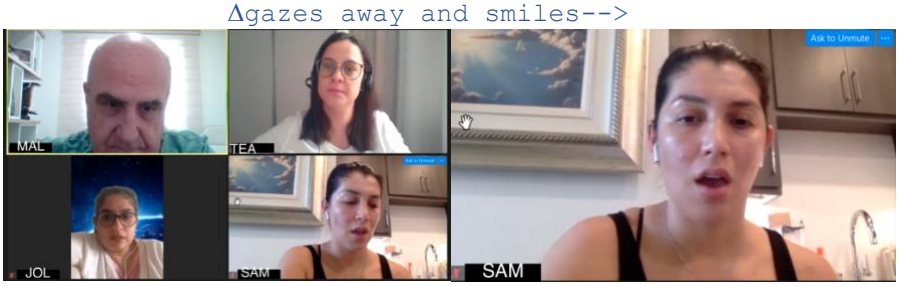
PAU also displays engagement in the ongoing activity and the monitoring of fellow participants by constantly nodding and gazing left and right (line 1) (#19 and #20). The teacher carries on with the storytelling in line 2, but as she reaches a possible storytelling closure, she initiates a word search to describe the end of the idea of TVs bringing the family together. This search is indicated by the pause in line 2 and the non-lexical speech perturbation, i.e. hesitation marker (‘a::’, line 3) followed by a 0.3-second pause in line 4. Such markers project the possibility for the other participants to initiate repair (Duran, Kurhila, & Sert, 2022), i.e. either the participant or other participants offer the candidate words (Hayashi, 2003). This is what occurs in the subsequent line as MOH’s offer of a candidate word overlaps with the teacher’s self-selection (lines, 5-6). Prior to ZAI’s contribution, we can see that he is displaying engagement and understanding as well. In fact, he is using his facial expressions and head movement to left and right, as if he is saying (‘no’) before providing a candidate word to resolve the teacher’s word search, which overlaps with the teacher in line 6.

Up to this moment, both learners are displaying engagement in the ongoing activity using various visual cues, including nodding, gaze movement, and head movements. They are both recipients of the storytelling and have the right to self-select to offer the candidate word the teacher is searching for. PAU activates his audio to self-select, but ZAI self-selects and offers a candidate word first and indeed he receives the teacher’s approval, while PAU drops back and deactivates his audio. It can be said that the delayed use of the affordance of audio activation constrained PAU’s access to the ongoing activity.

Slightly different from Extract 5.07, the following extract presents an instance where all the three participants are on-mute at the same time. In line 1, the teacher produces a polar question that does not specify the next speaker. Such questions addressed to the whole group are referred to in the literature as G-questions (Petitjean, 2014) or general solicit (van Lier, 1988, p.110). The question makes relevant next an answer, and in order to answer, the participants first need to activate their audio.

Extract 5.08 *[what video 11:41:07]*

```
01  TEA:  &>did you< see(.)
    sam   &>> does not gaze at the screen-->>
          any of you (.)
          did you watch the ↑video
02        (1.8)
03  TEA:  did you understand what it meant
04        (1.9)🎙#21&#22(2.0)Δ
    Mal        🎙activates audio-->>
    Sam        -->&turns, gazes at screen and opens mouth slightly-->
    Jol                  Δgazes away and smiles-->
    fig
```



Line 4. #21: MAL activates his audio.

Line4. #22: SAM turns and gazes at the screen and opens mouth slightly.

```
05  MAL:  what (.)
          the video 🎙&from: Δ↑sam
    Sam             🎙activates audio-->>
    Sam           -->&gazes away
    Jol                  -->Δ
06        (0.7)
07  TEA:  ahmu:m
08  MAL:  no (.) i see=
09  TEA:  =the video i sent during the week abou:t (.)
10        it was a ted video (.) if trum-
11        about conceding speech
```

The teacher's question makes the answering slot available to all three learners ('`any of you (.)`', line 2). Such questions may result in an answer delay, which is what happens in line 3, when a 1.8-second silence occurs. The teacher assumes that the learners have seen the video, as she further asks about the content of the video ('`did you understand what it meant`', line 3). This is followed by a relatively long verbal gap (3.9 seconds) in line 4. After 1.9 seconds of this long gap, MAL activates his audio following a fixed gaze towards the screen (#21). Despite his audio activation, MAL does not yet take the turn. It may be that SAM's upper posture turn, gaze shift from an activity that she is doing off camera towards the screen, and slight opening of her mouth (#22) are the cause behind MAL's 2.0- second wait. Nonetheless, SAM's gaze and mouth opening are not accompanied by the activation of her audio, which is a pre-requisite for speaking. It might be that MAL's audio activation and the green box around his image indicate that he is the next speaker. Thus, we see MAL self-selects and takes the floor in line 5. SAM activates her audio, as seen in line 5, but the

answering slot is already taken, and she then gazes away from the screen back to the off-camera activity.

In a manner comparable to Extract 5.07, Extract 5.08 demonstrates that delaying the activation of the audio results in the opportunity to participate being missed. Additionally, it is possible to draw the conclusion that when all the learners' audios are deactivated, the learner who is the first to activate it will likely secure the next turn. The first here refers to the one who carries the extra work of audio activation before the other fellow participants. At this stage, the analysis so far has displayed the positions at which the OML 'non-primary speakers' activate their audio to project self-selection for the next turn. The two positions revealed are both considered as pre-beginning practices but vary in terms of how long before the turn initiation they are produced. The affordance of activating audio shapes the turn-taking of the ongoing class activity, as it is essential for successful self-selection. The use of this affordance can be consequential for the participant's ability to self-select for the next turn in the sense that without its use, the participant's participation can be hindered, as seen in Section 5.2.3.

5.3 Deactivating Audio After Participation

One of the questions that this chapter seeks to answer is related to the position at which the OMLs who activated their audio to access the ongoing interaction return to their on-mute status. The data shows two recurring patterns in terms of the positions at which the OMLs deactivate their audio after participating in a class activity. In the first – and more common - position, the participants deactivate their audio immediately upon completion of their turn, i.e. the completion of the action they activated their audio for. In the second, the OMLs who activate their audio to take a turn in the ongoing class activity delay the deactivation action until an extended participation reaches a possible closure. The following sub-sections use a number of extracts to illustrate these two patterns and unpack the OMLs' use of this feature in situ.

5.3.1 Audio Deactivation at The Transition Space

The analysis in this sub-section will show a recurring pattern in the data, where the OMLs who activate their audio to access the ongoing interaction, return to their on-mute status by deactivating their audio right after their turn reaches completion, as illustrated in the post-analysis visual representation (see Figure 5.3) and exemplified by the following extracts.

Figure 5.3 illustrates the position of audio deactivation and the OML's turn followed by the audio deactivation exactly upon its completion.

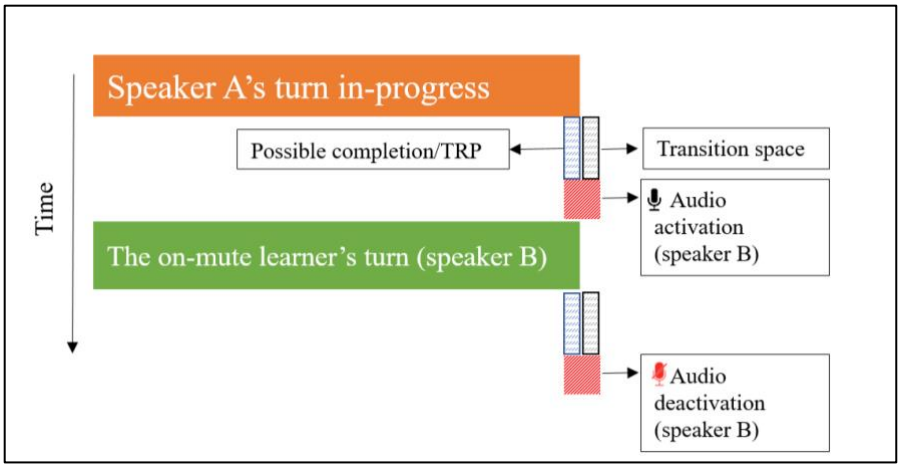


Figure 5.3: A Schematic Representation of Audio Deactivation upon Turn Completion

The analysis below demonstrates the participants' strategic use of these platform's features as it unfolds in the interaction. Extract 5.09 comes from a class that involves 3 learners along with their teacher. In this class, LOR keeps her audio deactivated and only activates it to initiate a turn at the ongoing class activity. Prior to the extract, LOR addressed a question to SUL regarding where they celebrate Eid (the Muslim religious celebration) in their village. He answered that they have places dedicated to this purpose, or sometimes they can stay outside (lines 1-5).

Extract 5.09 *[with all ages? 00:17:49]*

```
01   SUL:  we have a: (.)
02         like stay ↑outside^ (.)and
     lor                      ^nods-->>
03         (0.6)
04         ^we have prepared this from (0.9)
           -->^
05         like >three years< ago
06         (1.1)🎙(0.2)@(0.1)
     lor       🎙 activates audio-->>
     lor            @opens mouth slightly-->>
07   TEA:  a[hm:
08   LOR:  @[an- a- (.)
           -->@
09         you said you celebrate (.) like with all ages (.)
10         or only with your friends that are you're a:ge
11         (0.5)🔴(1.5)
             -->🔴 deactivates audio
12   SUL:  in the morning (.) with all ages&
     lor                                 &nods-->>
13         and a: at night with my friends&^
     lor                              -->&
     lor                                 ^smiles
```

LOR and the teacher display recipiency and engagement, as they both maintain a fixed gaze at the screen while LOR also nods, as seen in line 2. SUL's turn reaches a possible completion in line 5 and a 1.4-second silence occurs, marking SUL's turn as complete and creating a possible transition place. During the 1.4-second silence, LOR activates her audio and opens her mouth slightly in preparation to launch the next turn. This pre-speech activity by LOR is not recognised by the teacher, whose gaze just returns to the screen after briefly gazing down and, as a result, they both initiate the next turn at the same time (lines 7 and 8). The teacher drops back in resolution to the overlap. LOR produces a follow-up question addressed to SUL ('you said you celebrate (.) like with all ages (.)or only with your friends that are you're a:ge', lines 9-10).

The question's format requires SUL to choose from two candidate answers that have already been supplied by the asker. The complete production of these two options and the slower intonation of ('a:ge', line 10) projects the question's possible completion. What also marks the completion of LOR's turn is the audio deactivation during the 2.0-second silence in line 11. The deactivation of audio at this particular position can answer the question regarding "why this, in this way, right now?" (Seedhouse, 2004). It is done now as the action that triggered the audio activation (asking the follow-up question) is now complete. LOR activates the audio in the previous TRP, asks the question, and deactivates the audio right at the TRP. By doing this, LOR marks her participation as complete and hands over the interactional floor to the question's intended recipient. This momentary and finely-tuned use of technological features demonstrates the participants' awareness of participating in class activities without allowing their background noises to be heard in the class's shared space, as LOR, on a number of occasions during this class, appears to be talking to other family members.

In addition, Extract 5.10 below shows another example of this pattern when the teacher produces a G-question (Petitjean, 2014) that contains a multiple choice. The extract shows the interactional work the learner does before self-selecting to answer the teacher's question and also presents where the audio activation ends.

Extract 5.10 *[sleep in a bed 00:40:10]*

```
01  TEA:  ↑okay (.) a::
02        ↑where do you sleep boys^
    mar                      ^gazes constantly to left and right-->>
03        on a <be:d> (.) on a bean ↑ba:g^ (.)
    mar                               -->^
04        hhhh (.)
    mar       activates audio-->>
```

```
05          [↑hammocks (.)
06  MAR:    [i s-
07  TEA:    [↑where
08  MAR:    [i sleep in a bed (.)
09          a: (.)with my wife (.)
10          and a: (.) is (.) is a good bedroom
11          you know (.)
12          a: ↑nice (.) with a: television
13          and you know (.)
14          &some @↓facilities(.)↓as well&@
    mar     &------------nods-----------&
    tea           @---smiles and nods-----@
            (0.5)
15  TEA:    .hh (.)oka:y (.)
    mar                  -->deactivates audio
16          and you (.) ↑ZAI
```

In line 1, the teacher marks the previous discussion point as complete, using ('↑okay'), with a raising intonation functioning as a topic boundary marker (Svennevig, 2012), and projecting a new activity is coming up. This is indeed followed by asking the next question in the questions' list that is meant to guide the class discussion ('↑where do you sleep boys^', line 2). The teacher's question does not specify a learner to answer it; rather, it places the job of negotiating who is going to answer on the learners' side. As the transcript shows, in the middle of the teacher's question, MAR begins to gaze lift and right. It should be noted here that although it is difficult to precisely tell what a certain participant's gaze is directed towards, gaze constant movement can project an upcoming action. It can be assumed that MAR is monitoring his fellow learner to see whether he is going to self-select to answer the question. Indeed, his fellow learner is maintaining a smile and does not display a preparation to answer. MAR activates his audio after the teacher's laughter, before the teacher self-selects to add one more option. MAR overlaps with the teacher's question completion twice (lines 6 and 8), before the teacher drops back and hands the floor to MAR.

MAR provides a multi-TCU answer in lines 8-14 in which he describes his bedroom and who he shares it with. As the turn reaches a possible completion, MAR's intonation becomes lower and is accompanied by nodding, as seen in line 14. This is also accompanied by the teacher's smile and nodding, followed by a 0.5-second silence. The teacher then self-selects for the next turn and uses the same transition marker as earlier ('oka:y', line 15). This transition is recognised by MAR, who deactivates his audio right after the transition marker. Indeed, the teacher readdresses the same question to ZAI afterwards ('and you (.) ↑ZAI', line 16). The audio deactivation here is well timed with the end of MAR's participation and the teacher's disengagement and transition to the other learner in the class. MAR returns to the on-mute

status, but still displays active listenership and engagement using visual cues, such as nodding, smiling, and head movements.

The OMLs' strategic use of audio activation/deactivation to organise turn-taking can be further seen in the next extract. Extract 5.11 presents an example for a brief use of these features. This brief use is well coordinated with the other participants' actions and also finely tuned with the completion of own turn, even if it is as short as producing a one-word response.

Extract 5.11 *[sure 00:20:04]*

```
01  OBA:  it's really sounds great
02        &(2.1)&(1.0)
    lor   &smiles, nods and mouths 'yeah'&
03  TEA:  .hh=
04  OBA:  =↑so (.) can i ask you one question
05        &(1.5)&
    lor   &-nods-&
    lor           activates audio-->>
06  LOR:  sure
           -->deactivates audio
07  OBA:  >do you< know your: (.)
08        your neighbours (0.3)
09        ↑all your neighbours (.) or:
```

Prior to the extract, LOR was telling a story about an outdoor movie night she organised for her neighbours, to which OBA produces an assessment in line 1. LOR, who immediately deactivated her audio after completing the story, responds to OBA's assessment by smiling, nodding and mouthing 'yeah' during the 2.1-second gap in line 2. This is followed by a 1.0-second silence and neither LOR nor OBA claims the conversational floor. The teacher self-selects and prepares to launch the next turn using an in-breath (line 3). However, OBA initiates a turn, and the teacher abandons his turn initiation. OBA produces a polar question ('`=↑so (.) can i ask you one question`', line 4). Although OBA's question does not specify LOR as the question's recipient, the participants treat it as addressed to her as LOR was the last participant talking.

OBA's question is followed by a 1.5-second gap in line 5, which is only in the verbal conduct; the embodied interaction is ongoing as both the teacher and LOR nod positively. This is followed by LOR's audio activation to produce the verbal response to the question ('`sure`', line 6). This response works as a go-ahead to OBA's question. LOR deactivates her audio immediately upon the completion of the approval production. This audio deactivation does not only hand the interactional floor to OBA, but also marks the go-ahead action as complete. Indeed, OBA self-selects and produces a multi-TCU question in lines 07-09.

In all the three examples in this sub-section, the OMLs deactivate their audio upon turn completion. The audio deactivation action is positioned after the completion of the action that the participants activated their audio to accomplish. For example, in Extract 5.09 it was immediately following the completion of asking a question. In Extract 5.10, the audio deactivation is well timed with the completion of answering a question, whilst in Extract 5.11 it is positioned immediately after the completion of the go-ahead token. In all the previous extracts in this section, the participation is complete, and the participants return to their on-mute status immediately after turn completion. However, another recurring pattern discovered in the data reveals that the OMLs may delay the audio deactivation to accommodate an extension of their participation. The following sub-section provides examples to illustrate this pattern.

5.3.2 Delayed Audio Deactivation Beyond Turn Completion

This subsection aims to demonstrate another recurring pattern in the data in which the participants who activate their audio to participate in the class ongoing activity delay their return to the on-mute status. In order to accomplish this, the analysis seeks to track the trajectory of the OMLs' use of the features of audio activation until the time when they deactivate the audio again. As previously established (see Section 5.2), these participants prefer to remain on-mute when they have - or expect to have - disruptive noises from people, animals, or TVs in their physical spaces.

Figure 5.4 below shows a post-analysis visual representation of the pattern in this subsection, which illustrates the position whereby the participants who activated their audio to take a turn, return to the on-mute status. The participants delay the audio deactivation to accommodate an extension to their participation in the ongoing class activity. They deactivate their audio upon the completion of their extended participation. Extract 5.12 is an example of delayed audio deactivation that is used by a participant to achieve further extension to his or her participation.

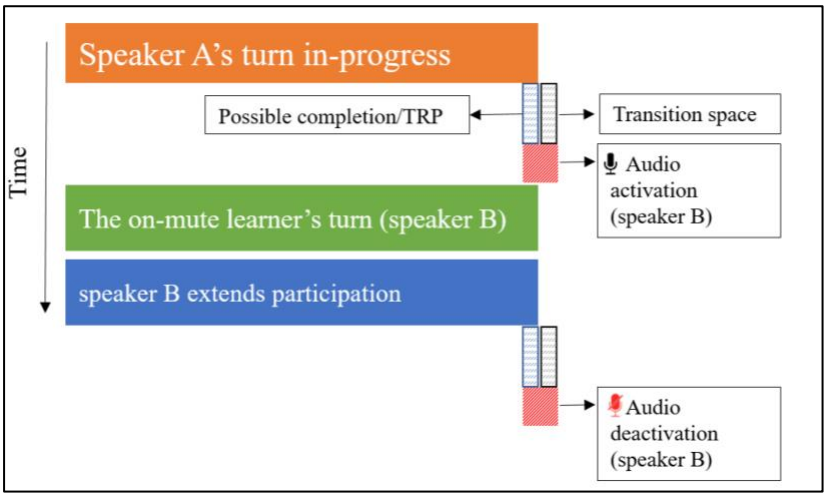


Figure 5.4: A Schematic Representation of Delaying Audio Deactivation Beyond Initial Turn Completion.

Extract 5.12 *[baseball 00:12:21]*

```
01   TEA:  $baseball$ (.) yea:h
02         i would say american football (.)
03         but american football is ^kind of easier^ than(.)baseball
     jol                          ^video shakes slightly^
     jol                                  activates audio-->>
04   JOL:  ↑yeah (.) you- you can get it (.) american football
05   TEA:  Yeah
06   JOL:  but baseball no
07   TEA:  baseball o:h yeah
08         it's true:
09         (0.4)
10   TEA:  .hh=
11   JOL:  =bu:t a:[after the
12   DIN:          [oh so boring
13         (0.3)((JOL laughs))
14   TEA:  yeah hhahh (.)
15         it is hh
16   JOL:  but after the explanation of dina husband(.)
17         i got it(.)
18         the elections(.)
19   TEA:  .h[h
20   JOL:    [↑bu:t (.)
21         i have questions about the concesh- conceding speech
22         (0.3)
23   TEA:  ↑ye:s
24   JOL:  it's not (.) a: it's a la:w
25         a- aft- if you could dina watch that video (.)
26         the ted video (.) that tea sent to us
27         it's [very interesting
28              [the one (.)
29         the last one
30   TEA:  ↑no: (.) was not the last one (.)&
                        &launches whatsapp page to resend video-->>
31         i'm gonna send it again
32         (1.7)
33   DIN:  °ok°
34         (0.8)
35   JOL:  a- ah
36         what i- i understood that a: (.)
```

```
37          trump is not oblige to: give up the:
38          (0.5)
39          his (.)
40    TEA:  Posit[ion&
                 -->&returns to zoom window
41    JOL:        [pres-
42          presid- (.) a:h (.) [his position
43    DIN:                      [a:h (.) okay
44          (2.3)^(0.2)#23^
      jol         ^video shakes slightly^
      jol            -->deactivates audio
      fig
```

Line 44. #23: JOL deactivates audio.

```
45    TEA:  .hh
46    MAR:  Yeah (.) but=
47    DIN:  =ah (.) okay (.) i didn't watch that
```

Prior to the extract, the teacher produced an evaluation by stating that American elections are very difficult to comprehend. JOL replied by telling the class that she compared it to baseball in her family group chat, and that only Americans can understand it. The first three lines of the excerpt show the teacher's agreement with this analogy. JOL's video slightly shakes as her arm stretches to activate the audio in preparation to self-select for the next turn at the possible TRP as the teacher's turn reaches a possible completion in line 3. The reason for the slight video movement is that JOL is using a smartphone to join the class and is holding it with her right hand while using her left hand to click the audio activation/deactivation buttons (see #23). JOL activates her audio, as shown in line 3, but delays the initiation of her turn until the teacher completes her current turn. Indeed, JOL self-selects and initiates the next turn to elaborate on the analogy mentioned earlier, to which the teacher displays agreement in lines 4-8.

A 0.4-second gap follows the teacher's assessment in line 8, which can be seen as a closure to the analogy sequence. At this point the sequence has reached a possible closure and JOL's turn is complete, which is a point at which the OML deactivates the audio again, as seen in Section 5.3.1. However, this is not what happens next; the teacher self-selects for the next turn in line 10 and begins with an in-breath, but this pre-speech activity is latched by JOL's turn initiation ('=bu:t a:[after the', line 11). DIN's assessment of baseball in line 12 overlaps with JOL's turn initiation, with both the teacher and JOL bursting into laughter in lines 13-15. The teacher acknowledges the assessment and displays agreement. This overlap and the

subsequent laughter and agreement delay the initiation of JOL's turn, and therefore, the delay of audio deactivation occurs as her turn is not yet fully produced.

JOL again pursues the initiation of the turn and repeats the previously overlapped turn, but succeeds this time in securing the floor ('but after the explanation of dina husband(.)' line 16). She stresses ('but') to allow her to maintain the interactional floor and secure the slot. JOL launches a multi-TCUs pre-question in lines 16-18 with reference to a past event when DIN's husband explained the process of the elections to the class. The teacher self-selects for the next turn using an in-breath, but this pre-speech activity is overlapped by another stressed with upward continuing intonation ('[↑bu:t (.)', line 20) to hold the floor further, and the teacher drops back as a resolution to this overlap. JOL further extends the pre-question by announcing that she has a question regarding Trump's conceding speech (lines-20-21). Up to this point in the extract, JOL has extended her participation on three occasions, in lines 11, 16, and 20. The delayed production of the question, due to the overlap in line 12 and the multi-TCUs pre-questions in both lines 16 and 20, all play a role in the extension of the participation, and thus, the delayed audio deactivation as a result. If the question is not produced yet, then the audio deactivation is likely to be delayed until its production is accomplished.

The announcement in lines 21-22 is followed by a 0.3-second pause, after which the teacher responds with a minimal response token ('↑ye:s', line 23) as a continuer that enables JOL to hold the floor and go ahead with asking the question that she said she had. Indeed, JOL continues the turn in line 24 and wonders if the conceding speech is part of the law, before inviting DIN to watch the video as well, referring to its location in the chat group. DIN enquires about the video order in the group ('the last one', line 29), as there is more than one. The teacher answers this and announces that she will resend the video. The teacher launches the chat group to resend the video during the 1.7-second silence in line 32. DIN appears to hold the phone, waiting for the video to be replied to by the teacher. At this point, JOL is heard initiating a new turn using restarts (Carroll, 2004; Goodwin, 1980) in line 35 and producing a candidate understanding of the video content (lines 36-37). In line 37, JOL seems to search for a word, as indicated by the stretching of ('the:', line 37) followed by the 0.5-second pause before the teacher initiates other-repair and provides a candidate vocabulary item in line 40. JOL accepts the teacher's vocabulary with a minimal response token, which indicates a shift in the epistemic status (Heritage, 1984a) and repetition of the suggested word ('position', line 42).

JOL activates her audio to elaborate on an analogy that she had made prior to the extract and received an assessment from the teacher. This analogy sequence is possibly complete, followed by a 0.4-second silence, which is a point where the participants can deactivate their audio. However, JOL does not deactivate the audio at this point. Instead, she self-selects for the next turn, latching on to the teacher's pre-speech activity and initiates a new sequence (lines 10-11). When JOL's turn initiation is overlapped in line 11, she momentarily drops back to self-select again to pursue that topic initiation in line 16. JOL holds the floor further and announces that she has a question in line 20, but ends up inviting the other participants to offer candidate understanding. As the transcript shows, the other participants are not familiar with the video; therefore, JOL offers a candidate understanding of the video content and further extends participation. This sequence comes to a possible closure after JOL offers the candidate answer to her own question, followed by a 2.5-second silence during which she maintains a fixed gaze at the screen. As seen preceding the audio activation, the video shakes slightly as a result of JOL stretching her arm out to click on the audio deactivation button on the screen. This is followed by audio deactivation (#23) and the teacher's preparation to initiate the next turn using an in-breath in line 45.

Another example of this pattern is illustrated by the next extract, in which a participant delays the audio deactivation to greet a fellow learner who has reconnected to the class after attending to an audio technical trouble. Prior to the extract, JOL has activated her audio and self-selected for the next turn.

Extract 5.13 *[the movie 00:07:13]*

```
01   JOL:  you guys watch the movie that i sent to you[ (.)
           >>activates audio
02   TEA:                                              [gasps
03   JOL:  in the: gro:up
04   TEA:  i ↑di:d (.)
05         that's ↑unbelievable
06   JOL:  ↑a:ha:[ (.)
07   TEA:        [gasps
08         they- they had to have a lot of guts (.)
09         to interrupt (.) the president (.)
10         and ^say >no it's< not true
     din       ^rejoins the class
11   TEA:  for sur:e
12         (0.5)
13   TEA:  it's because really you know (.)
14         ↑outrageous (.)
15         for them to interrupt like ↑that (.)
16         like ^what he was ↑saying was like (.)
                ^DIN's audio status changes
17         @#24$ridicules (.) ↑right$
```

		@gazes at DIN-->>
	fig	

Line 17. #24: The teacher shifts gaze to DIN as the audio status changes.

18	JOL:	hhehh (.)
19		↑$yeah$
20		(0.3)
21	TEA:	hi ↑din (.)
22		i think you c- i can hear you[now (.)
23	DIN:	[↑hi everyone
24	TEA:	↑ye:s (.)
25		we can hear you@
		-->@
26	JOL:	↑hi dina
27		(0.5)
28	DIN:	hi [jol and mar
29	JOL:	[we c-
30		we can hear you ↑now
31		(1.0)
32	DIN:	$perfect$
33	JOL:	((smiles))
34	TEA:	how was your week& (.) dina& jol &video shakes slightly& jol -->deactivates audio
35	DIN:	was good (.) was good

Extract 5.13 begins with JOL self-selecting for the next turn and producing a polar question regarding a video she sent in the chat group (line 1). The teacher's reaction in line 2 overlaps the production of the question as she gasps to indicate surprise at the video content. Indeed, the teacher answers the question in line 4 ('i ↑di:d (.)') and follows this by producing an assessment of the video content ('that's ↑unbelievable', line 5). JOL follows with a display of agreement with the teacher's assessment in line 6, overlapping with another gasp from the teacher.

The adjacency pair is complete at this point in terms of question-answer production. It appears, however, that JOL's question does not only seek a yes/no answer, but also seeks an evaluation of the content, including her own. As such, the participation is not complete, and therefore, the delay occurs to the audio deactivation. JOL self-selects to provide her own evaluation of the video content in lines 8-10. In line 10, DIN reconnects to the class after leaving for moments to handle an audio problem. The teacher displays agreement with JOL in

line 13 and then self-selects after a 0.5-second pause in line 12 to elaborate more on the video content in lines 13-17. In line 16, the status of DIN's audio changes from the wavy dots (…), indicating an ongoing audio connection attempt to a tick sign, which represents connection success. The teacher then shifts gaze to DIN's image on her screen in line 17 simultaneously with the production of the tag question addressed to JOL ('↑right$') (Figure 24). JOL replies with laughter followed by agreement ('↑$yeah$', line 19), followed by a 0.3-second pause. It seems at this point that the sequence of seeking evaluation has come to a closure. This is marked by the teacher's disengagement from the dyadic interaction with JOL and moving to greeting DIN and highlighting the shift in her audio connection status (lines 21-23). The teacher confirms that DIN's trouble has been solved now after her verbal greeting to the fellow participants is heard (lines 23-25). JOL self-selects and greets DIN. and also confirms that her audio is now working. DIN expresses happiness that the fellow participants are finally able to hear her.

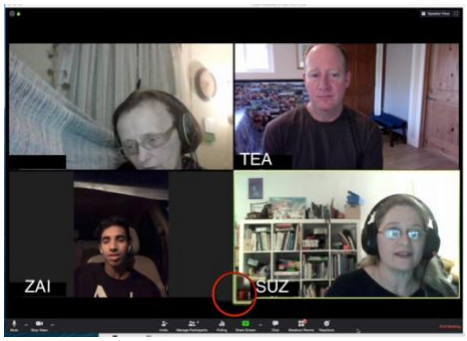
In line 34, the teacher addresses a question to DIN ('how was your week& (.) dina*&'). This question specifies DIN as the recipient and the next speaker. As the question is reaching a possible completion, JOL's video seems to shake slightly before she deactivates the audio. JOL keeps the audio activated beyond the question-answer adjacency pair completion, sequence closing and delays it until after greeting DIN. The delay to the audio activation can be extended beyond the completion of the initial action, which the participant activates their audio to do. The participants can also delay their audio activation to pursue eliciting further information from the recipient of their initial question. Extract 5.14 shows another example for this pattern, when SUZ delays the deactivation of the audio to self-select and produces a follow-up question to elicit a sufficient answer to her question addressed to ZAI. It may be worth noting that during the class, SUZ is very precise in terms of deactivating her audio immediately after the end of her turns. Therefore, delaying the deactivation of her audio is for the purpose of producing another question, as will be seen in the analysis of the next extract.

Extract 5.14 *[in the neighbourhood 00:17:46]*

```
01   ZAI:  ^we have to: (.) celebrate two times in the year(.) yes
     suz   ^smiles-->>
02         (1.0)🎙(0.7)
     suz         🎙activates audio-->>
03   SUZ:  ^zai (.) where do you celebrate (.)
           -->^
04         when you meet your community (.)
05         where do you do ↑that (.) all the:↓
06   ZAI:  in the neighbourhood
07         ^(1.1)^
     zai   ^fixed gaze at the screen^
```

```
08   SUZ:  do you have a [pla
09   ZAI:                [yeah we hav-
10   SUZ:  you [like to s-     to set     ]<together>#25 @
                                          -->deactivates audio
```

fig



Line 10. #25: SUZ deactivates audio after the completion of the question.

```
                                                    @nods-->>
11   ZAI:       [yeah (.) we have a places]
12              (1.0)
13   ZAI:  @yeah we have a places (.)
     suz   -->@
14              not only place
```

In line 1, ZAI's answer to a question that he was asked prior to the beginning of the extract is reaching a possible completion. SUZ appears to display engagement by smiling, as seen in line 1. This is followed by a 1.7-second pause in line 2, during which SUZ activates her audio and self-selects for the next turn. SUZ produces a question and specifies ZAI as an intended recipient ('`^zai (.) where do you celebrate (.)when you meet your community`', line 3-4). Both this version of the question and the reformulated one in line 5 begin with (`where`), which places some type of restriction on what the next speaker says (i.e. it requires the name of a place for an answer). Indeed, ZAI produces the answer and provides a place ('`in the neighbourhood`, line 6). This is followed by a 1.1-second silence in which ZAI maintains a fixed gaze at the screen and does not seem to self-select to elaborate more on the answer. Moreover, during this silence, SUZ appears to move her arm slightly, which can be seen as grabbing the mouse again, to deactivate the audio. However, SUZ self-selects to ask a follow-up question to pursue an answer that is more than just naming a large space, such as a neighbourhood ('`do you have a [pla=`', line 08).

The question's initiation is overlapped by ZAI's self-selecting to answer the question, even before it reaches its possible completion. Both SUZ and ZAI drop back to resolve the overlap and SUZ pursues the production of the question's full form ('`you [like to s- to set ]<together>`', line 10). Nevertheless, this attempt is also overlapped by ZAI's turn-initiation to answer the question and provide more information about the place they celebrate at (line 11). Line 10 shows the position where SUZ deactivates her audio (#25). This is located right at the TRP, following the completion of the follow-up question. The audio deactivation is also accompanied by continuous nodding. Audio deactivation and the nodding at the final turn

position both mark SUZ's participation as complete and hands the interactional floor to ZAI to produce the preferred answer to the follow-up question. Indeed, ZAI takes the floor and produces a multi-unit answer providing more details on the celebration place.

The previous sections in the chapter demonstrate that the participants use the features of audio activation/deactivation strategically to gain access to the ongoing interaction and to mark their participation as complete and partially disengage from the class activity. The analysis shows that the use of these two features is essential to manage turn-taking for the OMLs to participate in the class activities. A valid question at this point would be related to the absence of audio activation/deactivation: *how is this absence treated by fellow participants?* The following section presents two extracts to show that the delay or absence of activating audio is noticeable and accountable by the other participants.

5.4 Absence or Delay of Audio Activation

Crucial to the success of the interaction in a video-mediated L2 classroom is the existence of a clear audio channel. The absence of a video or audio feed can be noticeable and accountable by fellow participants. A participant is able to choose to deactivate his or her audio during part or whole class activities and activate it prior to taking turns in that activity, whether they self-select or are selected by current speakers for the next turn. However, the failure to activate the audio on time can be noticeable and accountable by other participants.

A close examination of the data in the current study shows that the delay of audio activation is noticeable and accountable by reference to the norm (Seedhouse, 2004). The norm in this case is that the OMLs need to activate their audio before taking the turn for successful communication. Hence, the absence or delay of audio activation is consequential to the progression of the ongoing activity and fellow participants may treat it as problematic. Extract 5.15 below illustrates more on this observation.

Extract 5.15 *[unmute yourself please 00:35:34]*

```
01  TEA:  back to you recardo (.)
02        .hh (.) do you take your shoes off
03        (0.3)
04        when entering (.) your home:
05        (4.5)
    rec   Sits up
    rec   opens his mouth slightly
06  TEA:  Unmote#26((unmute)) yoursel- (.) your (.) mic please
```

```
    fig
```

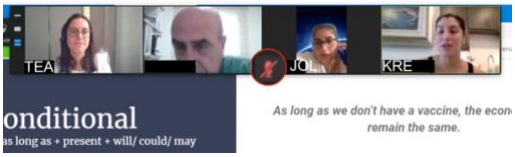
Line 6. #26: MAR's audio status and the teacher's gesture requesting him to activate his audio.

```
07          (2.0)
    mar           activates audio-->>
08  REC:    a::h (.) no (.)
09          because my: (.) my (.) my apartment
```

The teacher addresses a yes/no question to REC in lines 1-4. Her question is syntactically, semantically, and prosodically complete in line 4. The teacher's question assigns REC as the next speaker and its format makes relevant next a yes/no answer. A 4.5-second gap occurs in line 6 during which REC straightens up after leaning towards the screen and opens his mouth slightly. As #26 shows, REC has his audio deactivated and in order to produce the answer, he needs to activate it first. The teacher treats this as problematic, and initiates repair by explicitly asking REC to activate his audio ('`Unmote+((unmute)) yoursel- (.) your (.) mic please`', line 6). The teacher's verbal repair initiation is accompanied by a hand gesture that mimicks the click on the unmute button on the screen (#26). REC activates his audio in line 7 after a 2.0-second pause and immediately produces the answer to the teacher's question. This consequentiality of the delay in audio activation is clearer when the OML is selected to speak next by the current speaker. Not only the absence of audio activation is noticeable, but its delay is also noticeable. In what follows is an example from the data which illustrates this.

Extract 5.16 *[your sound is off 00:20:14]*

```
01  KRE:    a::* (.) the #27*economy the united states will be [the same
    Jol        *--says something but mic off--*
02  TEA:                                                     [↑ye:s (.)
    fig
```



Line 01. #27: JOL speaks but her audio is deactivated.

```
03          will remain the same (.) yes
04          (0.9)
05  TEA:    ↑jol (.) &i can't#28 hear you&
                     &-points to ears-&
06          your sound is off&
    fig
```

Line 6. #28: TEA uses hand gesture indicating inability to hear.

```
07          *(2.1)*
```

```
    Jol   *--mouths a::h but mic off--*
08  JOL:  ↑sorry:+
          activates audio-->>
    Tea           +smiles
09        cause my tv is turn on (.) at cnn (.)
10        i was watching the elections
```

Prior to the extract, the teacher had asked KRE to complete the sentence appearing on the shared screen using the 1st conditional 'unless'. When KRE struggles to answer the question following multiple tries, the teacher asks JOL if she can help with the answer. The extract marks the beginning of a new attempt by KRE to answer. JOL visually overlaps with KRE's answer as her audio is off, so no verbal overlaps occur (#27). The teacher does not orient to JOL's attempt to access interaction at this point. Instead, the teacher offers a positive assessment of KRE's answer ('`[↑ye:s (.) will remain the same (.) yes`', line 2). After a 0.9-second pause in line 3, the teacher addresses JOL by using her name and telling her that she is not being heard in class ('`your sound is off`', line 5). The teacher uses a hand gesture (#28), indicating the inability to hear what JOL is saying. The use of visual cues, such as pointing to ears to orient to the absence of audio, is common in the data. The teacher's turn is designed as a noticing, which functions as a request to JOL to activate her audio. In fact, JOL recognises this request when she mouths ('`a::h`', line 6) during the 2.0-second pause. JOL activates her audio and starts the turn with an apology and accounts for staying on-mute ('`cause my tv is turn on (.) at cnn (.)i was watching the elections`', lines 8-9).

From these two extracts, it can be observed that the delay or absence of activating the audio is noticeable and accountable by the other participants in video-mediated interaction. Thus, the well-coordinated audio activation/deactivation with the ongoing turns is seen but unnoticed. The two extracts also provide evidence regarding the participants' orientation to the reason behind keeping their audio deactivated, which is to block their background noises from being heard in the class. The occurrence of these background noises could result in hearing or speaking troubles that impact the progression of class activities. Such occurrences and how the participants handled them will be illustrated in the following two chapters.

5.5 Summary

This chapter presented an analysis of the participants' methodical use of the audio activation/deactivation features to organise their turn-taking. Section 5.2 showed how the participants systematically use these features as resources to self-select for the next turn by tracing the sequential positions of these practices. Section 5.2.1 presented extracts showing that the participants activate their audio at the TRP and immediately take the floor, while Section 5.2.2 presented an early position of audio activation before taking the turn by the participant. In addition, Section 5.3 demonstrated how the participants mark their participation as complete by deactivating their audio; Section 5.3 also showed the positions of these practices. Section 5.3.1 presented extracts demonstrating that the participants deactivate their audio immediately upon the completion of their turns, while Section 5.3.2 showed a delayed position of audio deactivation by the participants to accommodate for extended participation in class activities. Moreover, Section 5.4 presented extracts to demonstrate that the absence or delay of activating audio is noticeable and accountable by fellow participants, which paves the way to Chapter 6 that displays the participant's use of the audio activation/deactivation to as resources to repair hearing/speaking troubles in online, video-mediated, L2 speaking classes.

Chapter 6. Analysis: Using the Audio Activation/Deactivation Features to Search for the Trouble Source Producer

6.1 Introduction

The analysis in the previous chapter showed the participants' strategic use of the audio activation/deactivation features to manage their turn-taking in video-mediated L2 speaking classes. In addition, Section 5.4 showed that the delay or absence of self-audio activation to take the turn is noticeable and accountable by fellow participants. The current chapter examines the repair organisation of trouble caused by the interference of disruptive noises in online video-mediated L2 speaking classes using the audio activation/deactivation features of Zoom. According to Brandt and Jenks (2013), trouble can emerge in online settings due to various reasons, including the interference of participants' background noises. Such background noises are considered one of the variables that can shape interaction in online synchronous communication, which can be consequential to the ongoing talk in terms of the participants' ability to suspend or terminate the ongoing activity (Jenks, 2014). Therefore, resolving troubles caused by the interference of participants' background noises can be fundamental to the progression of the online video-mediated classes' activities, as will be presented in this chapter.

This chapter focuses on an examination of the participants' use of audio activation/deactivation features as resources to search for the source of disruptive noises in order to repair the trouble and restore the progression of a suspended class activity. The study of repair within CA entails a close description of the participants' methods to identify and resolve "troubles of speaking, hearing and understanding" as they unfold in interaction (Schegloff, Jefferson & Sacks, 1977, p.361). Following the CA methodology, this chapter provides detailed empirical evidence of the sequential organisation of trouble emergence and the teachers' use of audio activation/deactivation features to locate the trouble source producer, repair the trouble, and restore the progressivity of the interrupted classroom activities. The chapter will argue that there is a reflexive relationship between the participants use of the audio activation/deactivation features and the management of repair in L2 video-mediated classroom interaction.

The organisation of the current chapter is as follows: Section 6.2 will present the key terms used in Chapter 6 and Chapter 7 to ease the understanding of the analysis. Next, Section 6.3

will show the participants' use of audio activation/deactivation features to locate the trouble source producer; and finally, Section 6.4 will summarise the chapter.

6.2 Definitions of Key Terms used in the Analysis

The first term to define here is the **trouble source**. According to Schegloff (2000), nothing in talk cannot potentially be seen as trouble source. Following the emic perspective, a trouble source is what participants treat as such. In this chapter, a trouble source is the noises that the participants orient to as impeding hearing or speaking during online video-mediated class interaction. Examples of these are background noises, sounds resulting from connection problems, or in some cases noises of participants talking to or being talked to by people in their physical spaces.

In face-to-face conversations, the speaker of a trouble source is oriented to as the trouble source producer. However, it can be different in the context examined here (i.e. online video-mediated classrooms). Background noises such as a dog barking, a motorbike passing by, or a child playing are not spoken by the participants. Rather, these noises are picked up by the participant's microphone and fed into the class shared space. Therefore, the participants whose microphones pick up these noises and feed them into the Zoom room are considered **the trouble source producers**. It is important to note that they are only considered as such when the other participants deem them to be.

Another term used in this chapter is **class shared space**, referring to the virtual room where participants interact with each other, whether it be the class main room or a breakout room. Contrary to that, the participant's **private space**, or physical proximity (Brandt, 2011), refers to the participants' physical place from where they connect to the class shared space, such as a house, a café, or a school.

Having introduced the key terms used in this chapter, the next section unpacks the repair sequences and the multiple uses of audio activation/deactivation videoconferencing features to search for the trouble source producer.

6.3 Audio Activation/Deactivation Features to locate the Trouble Source Producer

This section demonstrates the participants' use of audio activation/deactivation features as resources to identify the trouble source producer and restore the progression of the suspended class activity. It also shows the role played by the teacher's use of these features in shaping the repair sequences. Figure 6. 1 is a post-analysis visualisation produced to ease the

readability of this chapter's analysis, which shows a self-repair space that follows the emergence of the trouble source. In the extracts analysed below, the trouble source producer does not exploit the first self-repair space, which results in a verbal other-repair initiation followed by another self-repair space. The absence of self-audio deactivation by the trouble source producer triggers the carrying out of the repair work by the other participants. However, to do this, the trouble source producer needs to be known to the other participants. In the following extracts, the trouble source producer is unknown to the participants. Thus, a search for the trouble source producer occurs using Zoom's audio activation/deactivation features. Once the participant feeding the background noise is located, the teacher deactivates their audio and the suspended classroom activity is resumed.

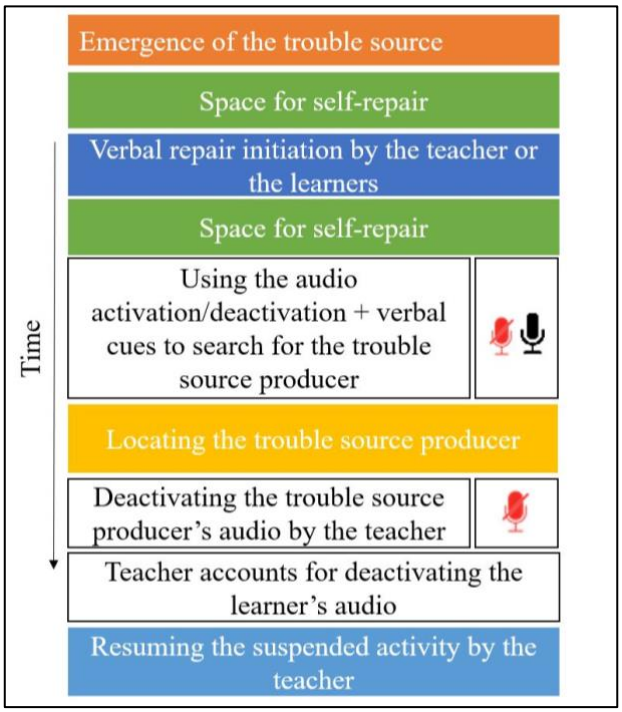


Figure 6. 1: A Schematic Representation of Searching for the Trouble Source's Producer Pattern

Due to the length of the extracts analysed in this section, each will be divided into three smaller extracts: the first shows the emergence of the disruptive noise, while the second and third show the participants' orientation to the trouble source, searching for its producer, repairing the trouble, and resuming the suspended class activity. At the beginning of each of the following extracts, a summary is provided before the detailed analysis.

Extract 6.01 is chosen to begin with, as it represents a clearer case that shows the teacher's use of the audio activation/deactivation features to locate the trouble source producer and restore the progression of the suspended class activity. The other two extracts in the section will also be compared against it.

Extract 6.01

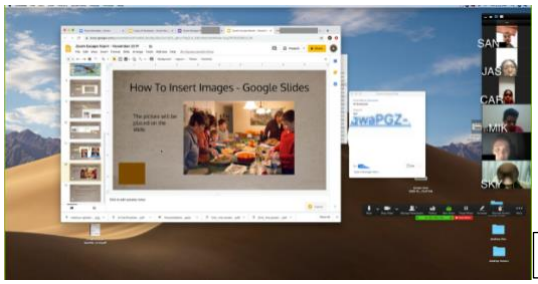
Participants: TEAcher, MIKki, JASmin, CARla, ZAId, GABriel, SANtiago, SKYla

Summary

Extract 6.01 comes from the beginning of a class that involves one teacher (TEA) and eight learners. Prior to the extract, the teacher had introduced the discussion topic of the session and the types of class activities that each group will collaboratively work on in their breakout rooms. The teacher then moved to the technical side of the class introduction and shared his screen to show the learners how to use Google Slides (see #1). During the instruction-giving, a background noise occurs in the class-shared space. Initially, the noise was not treated as problematic. However, another persistent noise occurs, which the teacher orients to as trouble and initiates repair. Following this, an insertion expansion ensues in which the participants collaboratively work to search for the trouble source producer. Next, the teacher deploys the audio activation/deactivation features to locate the trouble source producer. Once located, the audio of the participant who has the background noise is deactivated. The teacher then resumes the suspended class activity afterwards.

Extract 6.01.01: [*Disruptive noise emergence 00:17:44*]

```
01  TEA:  *#1when you select an image (.)
          >>*indefinite continuous noise-->
      fig
```



Line 01. #1: Teacher's screen layout.

```
02        you then click the insert button (.)
03        and it will: (0.4)
04        >↑pop< (.) pop into your slide
05        (2.4)&
                &loud bang noise
06  TEA:  >whatdya think of that< (.) JA- ↑JAS
07        (1.1)
08  JAS:  a: (.) i mean (.)that's great (.)
09        i'm just trying to figure out how am going to do that
10        on my pho:ne (.)*
                        -->*
11        bu:t (.)<when i get my computer> (.)
12        that is going to work just fi:*ne
                                        *parrot noise appears-->
```

Extract 6.01.01 begins with the teacher listing the steps that the learners need to follow to insert images in shared Google Slides (lines 1-4). The teacher is sharing his screen and decomposing his instructions to make them mimicable by the learners (Due, Lange, Nielsen, & Jarlskov, 2019). It is important to include lines 1-10 to show that background noises can occur in online video-mediated settings and are not uncommon. In fact, the participants hearing and dealing with such noises is at the core of audio/video-mediated interaction (Jenks, 2014). Nonetheless, the participants do not necessarily orient to these background noises as problematic with every occurrence. As the analysis will show, fleeting background noises are often treated as unproblematic by the participants.

Line 4 marks the end of the steps to insert an image and is followed by a 2.6-second pause in line 5. A loud banging noise in one of the learners' backgrounds occurs after this 2.6-second pause. Despite the loudness of the bang, the participants do not seem to treat this as trouble that needs to be repaired. After the 2.6-second pause, the teacher self-selects and addresses a question to JAS regarding her opinion of how easy it is to insert images in the dedicated slide ('`>whatdya think of that< (.) JA- ↑JAS`', line 6). After a 1.0-second pause in line 6, JAS responds in a multi-TCU turn in lines 8-12, indicating that she is trying this on her phone at the moment and that in the next class it should be easier when she joins from a computer.

At this point of the interaction, a disruptive background noise occurs in the class shared space, as JAS is reaching a possible turn completion in line 12. However, the background noise is not a fleeting one this time. It is a persistent noise as indicated by the arrows in the transcript above, in line 12. The noise is made by the purring of a participant's parrot (learned later in other lessons).

Below is Extract 6.01.02, which is a continuation of Extract 6.01.01 and shows the participants' orientation to the persistent background noise as being disruptive.

Extract 6.01.02: [*Orienting to the disruptive noise 00:18:11*]

```
13  TEA:  do you: aa (.) does everybody have a google account?
14  MIK:  yeah ((other learners nod positively))
15  JAS:  yeah i think so
16  TEA:  ↑holy moly*
                  -->*parrot noise quiet
17  MIK:  Yeah
18        (0.8)
19  TEA:  it sounds like somebody is wrestling with
20        a monster [over there
21  GAB:            [ha ha
22        ((learners laugh))
```

```
23  TEA:  $↑skyla what're you ↑watching$
24  ZAI:  ha ha*
                *continuos unknown loud noise-->>
25        (0.6)
26  SAN:  it's hungry
27        (0.9)*(3.1)
                *loud dog bark noise appears-->
28  JAS:  $oh go:d$
```

The teacher does not orient to the newly emerging noise as trouble yet in line 12, as he produces a polar information-seeking question to the class (‘`does everybody have a google account?`’). This is a requirement to the task accomplishment as the learners need such an account to work on Google Slides. MIK and JAS self-select to provide a positive answer to the teacher’s question, while the other learners nod positively in lines 13 and 14.

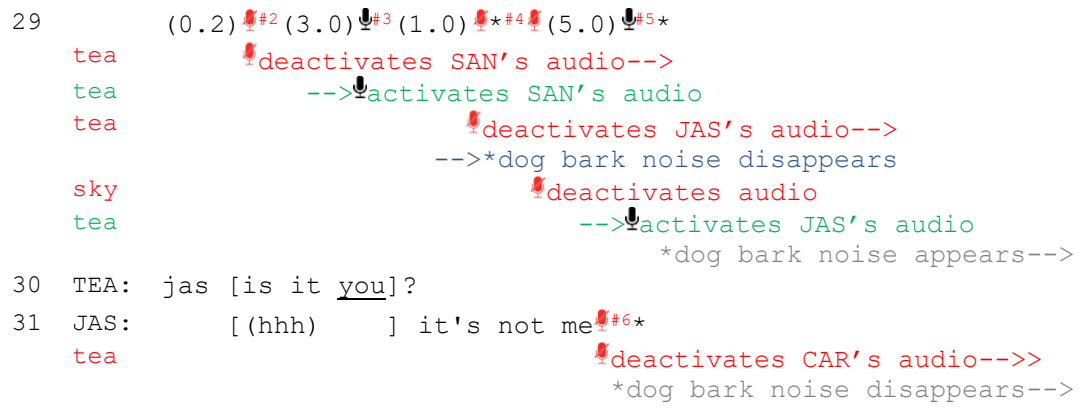
Up to this point, the noise that occurred in line 11 is still persistent but not yet treated as trouble. The initiation of other-repair is delayed and produced in line 15, which might be due to the expectancy that the noise will disappear, as these noises are not uncommon in online communication. The delay can also be seen as a repair space provided by the participants to the trouble source producer to carry out self-audio deactivation. It is worth reiterating that the teacher is sharing his screen, which can limit the vision of the learners of their fellow participants. As mentioned in Section 4.2.1, the Zoom platform allows teachers to share their screens with learners to ease instruction-giving, but this can also constrain their accessibility and that of others to see all the participants, especially those on speaker view. Speaker view only enables the participants to see fellow participants with active audio channels at a particular moment.

The teacher orients to the disruptive noise using a raising intonation (‘`↑holy moly`’, line 16), marking self-repair (self-audio deactivation) as ‘officially absent’ Schegloff (1990, p.213). This turn also marks an official departure from the class’s ongoing activity (instruction-giving). As the teacher’s exclamation turn reaches a possible completion, the purring noise becomes quiet. MIK’s (‘`yeah`’, line 17) seems to be a late response to the teacher’s enquiry about having a Google account. A 0.8-second pause occurs in line 18, followed by the teacher’s production of another repair initiation using an extreme case formulation (Pomerantz, 1986) (‘`it sounds like somebody is wrestling with a monster [over there`’, lines 19-20). The teacher’s complaint is met by laughter by the learners, as seen in lines 20-21. However, no self-repair is carried out up to this point of interaction.

As no self-repair is carried out by the participant with the noise in their background, the teacher produces another repair initiation in a more specific manner ('$↑skyla what're you ↑watching$', line 23). The information-seeking question addressed to SKY implicitly suggests that she is the trouble source producer. It also functions as an indirect request to SKY to carry out self-repair by deactivating her audio. This question-formatted request is met by laughter from the learners as well (line 24). At this point, a continuous unknown loud noise occurs, followed by SAN describing it as the monster being hungry (line 26). A 4-second pause occurs in line 26, during which a loud dog bark noise (barking dog) emerges. JAS orients to the new noise by expressing frustration ('$oh go:d$', line 28). SAN and JAS's orientation might indicate that by orienting to the noise, they disclaim being the trouble source producers. However, this is not recognised by the teacher at this moment. The suggested trouble source producer (SKY), for some unknown reason, does not respond to the teacher's question, which had been addressed to her. It is possible that SKY had connection trouble at that moment.

As the candidate trouble source producer is not responding to the teacher's repair initiation, the repair sequence is further extended to identify the participant that has the trouble source in their background. In order to locate the trouble source producer, more interactional work needs to be done by the participants. This extra interactional work is demonstrated in Extract 6.01.03 below, which is a continuation of the previous two extracts. It shows the use of audio activation/deactivation by the teacher to identify the trouble source producer. It is evident in the literature that in multiple repair initiations, participants move from weak to stronger repair devices (Schegloff et al., 1977), and also that extra interactional work is required when there is an unspecified trouble source producer (Brandt & Jenks, 2013). This is what can be seen in the following extract:

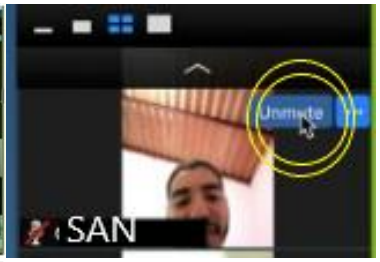
Extract 6.01.03: [*Searching for the trouble source producer and resuming suspended activity 00:18:34*]



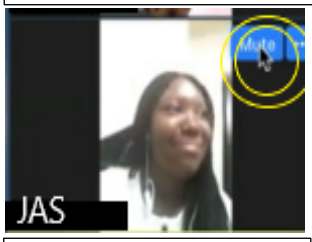
32 [(hhh) you're silencing me]
33 SAN: [(hhh) it's not me neither it's not me neither]
34 JAS: but it's not me (hhh)
35 SAN: i don't know whose dog is that
36 TEA: is it carla?&
37 (1.8)
38 SAN: maybe it's carla#7* i think
tea -->activates CAR's audio-->>
*dog bark noise appears-->
39 MIK: yeah it was carla
40 TEA: carla:::&*
tea *deactivates CAR's audio-->>
car &smiles
-->*dog bark noise disappears
41 ((learners laugh))
42 (0.8)
43 TEA: $you are mute:d$ (.)(hhh) (3.0) wow (.)
44 sorry carla (.) i muted you.
45 ((learners laugh))

fig

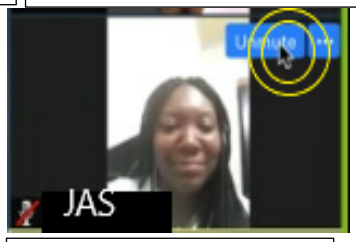
Line 29. #2: TEA deactivates SAN's audio.



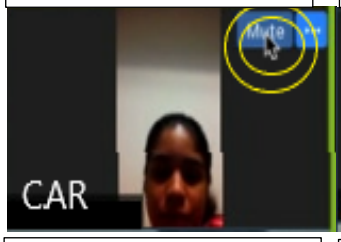
Line 29. #3: TEA activates SAN's audio.



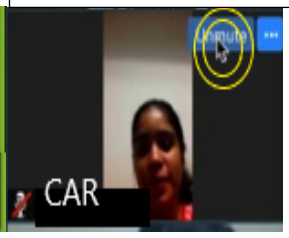
Line 29. #4: TEA deactivates JAS's audio.



Line 29. #5: TEA activates JAS's audio.



Line 31. #6: TEA deactivates CAR's audio.



Line 38. #7: TEA activates CAR's audio.

Following JAS's orientation to the noise, a 9-second pause occurs in line 28, during which the teacher extends the repair sequence and deploys the audio activation/deactivation features as resources to locate the trouble source producer. The teacher begins by deactivating SAN's audio (#2), who is at the top of the participants' list as they appear on his screen (See #1). The teacher holds for 3 seconds to test the effect of deactivating SAN's audio (i.e. does the noise

disappear?), as seen in #3. When the noise is still on, the teacher activates SAN's audio and deactivates that of JAS (#4), following the vertical order on the list. The noise disappears in the next 5-second pause, during which JAS's audio is off.

To confirm that the noise is coming from JAS's end, the teacher activates her audio (#5), and the noise re-enters the room. The teacher follows this up by producing a confirmation check question ('JAS [is it you]?', line 30), to which JAS bursts into laughter and informs the teacher ('[(hhh)] it's not me+* [(hhh) you're silencing me]', line 31). The teacher's question indicates that he is not sure yet and wants to accurately locate the participant who has the disruptive noise in their background. SAN, who has also had his audio deactivated earlier by the teacher, overlaps with JAS to point out that he is not the trouble source producer either ('[(hhh) it's not me neither', line 32). By doing so, JAS and SAN show that the teacher's action of audio activation/deactivation of the learners' audio is available to and recognisable by the other participants. Zoom shows an on-screen notification to the participants, which indicates that they are 'muted' when the teacher deactivates their audio (see Figure 4.3 in Section 4.2.1 of Chapter 4).

Before JAS and SAN complete their disclaimer turns, the teacher moves on with the search to identify the participant with the noise in their background. He then deactivates CAR's audio in line 31 (#6). The noise disappears again, but the teacher is also unsure and produces another confirmation check ('is it carla?', line 36). Note that the teacher this time addresses his question to the rest of the group, making his audio deactivation action more visible to the whole class, as CAR will not be able to respond. The teacher's question is also an invitation to the learners to partake in this search for the trouble source producer. There is an assessment of the effect of CAR's audio deactivation taking place in line 37 in the form of a 1.8-second pause to see if the noise has gone. SAN then answers the teacher's question with a degree of uncertainty ('maybe it's carla* i think', line 38).

As SAN nearly reaches a possible turn completion, the teacher reactivates CAR's audio (#7). As a result, the disruptive noise reappears in the class shared space. MIK answers the teacher's question with a confirmation ('yeah it was carla', line 39). MIK's confirmation and the teacher's calling ('carla:::', line 40) work as an announcement that the trouble source producer is finally located. This announcement is followed up by the repair solution to the trouble at the end of line 40 as the teacher deactivates CAR's audio. The teacher then announces his audio deactivation decision ('$you are mute:d$ (.)(hhh) (3.0) wow (.)', line 40) and follows this up by apologising and accounting for his decision ('sorry

carla (.) i muted you.', line 41) before moving to explain to CAR (not in the transcript) how microphones can pick up ambient noise and feed them into Zoom rooms before recommending that she needs to connect from a quiet place next time or keep the audio deactivated. The teacher then resumes the suspended class activity.

In Extract 6.01, the teacher deployed the audio activation/deactivation features, along with verbal means to identify the learner feeding background noise into the class shared space. This search was only launched after the trouble source producer did not exploit the self-repair space, which shows the teacher's orientation to a preference for self-audio deactivation (Schegloff et al., 1977). The audio activation/deactivation features are used as stronger repair devices to locate where the trouble source is coming from. Also, fellow participants held CAR accountable for noises that may have been beyond her control, in this case, a parrot purring and a dog barking. The repair sequence is extended until the trouble source producer is located, repair is completed, and the suspended class activity is resumed. It is worth noting that the platform's affordance of screen sharing may have contributed to the lack of accessibility of the learners to each other's videos and, therefore, constrained their ability to locate where the disruptive noise was coming from. However, the participants managed to use other platform affordances to repair the trouble that halted the progressivity of the class activity.

Using audio activation/deactivation features to locate the trouble source producer can also occur in classes with fewer participants. Also, its use does not guarantee success in locating the trouble source producer, as disruptive noises can disappear on their own. Extract 6.02 below demonstrates an example of this and shows the teacher's and the learners' collaborative work to identify the trouble source producer.

Extract 6.02

Participants: TEAcher, SUZzan, JASmin, MIKki, RON, SANtiago

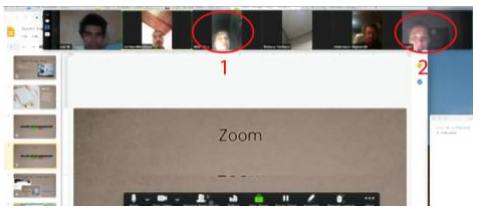
Summary

In this extract, the teacher is joined by 8 learners in an L2 speaking class. It occurs in the opening phase of the class while the teacher is explaining to the learners how to use Zoom, as most of them are relatively new to this mode of learning. The trouble source emerges when SUZ joins the class from two devices at the same time. Connecting from two devices in close proximity can produce a disruptive noise as the microphones on both devices pick up ambient noises and transmit them to each other, causing an echoey noise that gets louder when the participants speak. The participants orient to this disruptive noise and invite the unknown trouble source producer to carry out self-repair. The teacher, who has access to the audio activation/deactivation features, uses them to self-audio activate/deactivate his own audio to check whether the noise is emanating from his end. After multiple repair initiations, the noise disappears on its own when SUZ logs out from one of the devices. The class then resumes the suspended class activity. What follows is a detailed analysis of the extract.

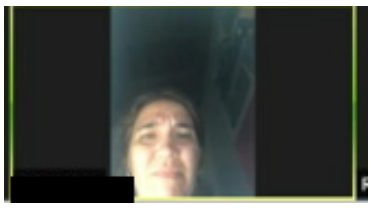
Extract 6.02.01: [*Where is that coming from 00:10:23*]

```
01   TEA:  ^*do you #8see (.)
           >>*the noise emerges (louder after pps speak)-->>
     Suz   ^appears in two videos-->>
02   TEA:  in this
03         #9(3.0)#10(2.0)
04   TEA:  ↑ho
```

fig



Line 1. #8: SUZ connecting from two devices



Line 3 #9: SUZ reaction to the noise loudness

Line 3. #10: RON orienting to the disruptive noise

```
05   JAS:  is anyone else getting that sound↑ (.)
06         where is that coming from
07   SUZ:  $ye:s$
08   MIK:  i ken hear it[ bu]t i don't know
09   RON:                [yeah]
10         (1.0)
11   RON:  yeah i ken hear it too*?
```

-->* noise quiet here-->

The echoey noise occurs at the beginning of the extract while the teacher is attempting to continue with instruction giving, which means that he does not yet treat it as trouble. As mentioned in Extract 6.01, it is not uncommon for such passing noises to disappear of themselves (Jenks, 2014). The teacher produces an incomplete question that is cut off by the echoey noise ('do you #see', line 1) and in line 2 after a micro pause. The teacher then orients to the trouble by halting the progressivity of his turn. A 5-second silence occurs in line 3, which is a relatively long silence and can be indicative of trouble. During this silence, as #9 and #10 show, RON and SUZ's orient to the disruptive noise by frowning. The teacher's turn in lines 1-2 projects more talk to come, as the question is syntactically incomplete. However, it is interrupted by the occurrence of the disruptive echoey noise. The silence offers a space for the trouble source producer to carry out self-audio deactivation, which does not occur.

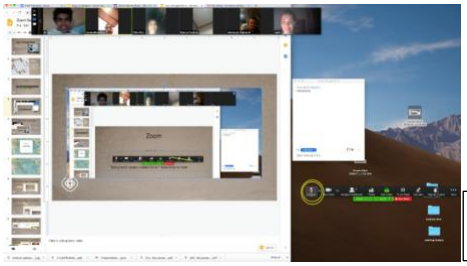
JAS initiates repair by interrogating the fellow participants' ability to hear the disruptive noise ('is anyone else getting that sound↑(.)', line 5) before she follows up with another question interrogating the source of the disruptive noise ('where is that coming from', line 6). Both the silence in line 3 and JAS's repair initiation mark an official departure from the class's ongoing activity. JAS's follow-up question makes relevant a candidate trouble source producer's name as a response. SUZ, MIK and RON, respectively, acknowledge the presence of the noise (lines 7-11). However, JAS's follow-up question is not answered yet. It can work as a request to the participant having the trouble source in their end to self-deactivate their audio. Note that SUZ responded to JAS's question, which might indicate that she is not aware of causing the echoey noise by connecting from two devices simultaneously.

As self-audio deactivation is not undertaken, the trouble source producer needs to be located for the repair work to be done. In this case, the trouble source is a disruptive noise, and it is putting the progression of the class activity at risk. A range of possible next actions can occur; either the noise producer self-audio deactivate (self-repair), or the teacher manages to identify where the trouble source is originating from and does the audio deactivation (other-repair). Following JAS's inquiry regarding the trouble source location in lines 5-6, MIK and RON co-orient to the trouble source and also negate knowledge of its whereabouts (lines 5-9). In addition, no self-repair is carried out by the trouble source producer. The continuity of the disruptive noise occasions the use of a stronger repair device - audio activation/deactivation features.

Extract 6.02.02 below is a continuation of extract 6.02.01, which demonstrates the search for the trouble source producer using audio activation/deactivation features accompanied by verbal means.

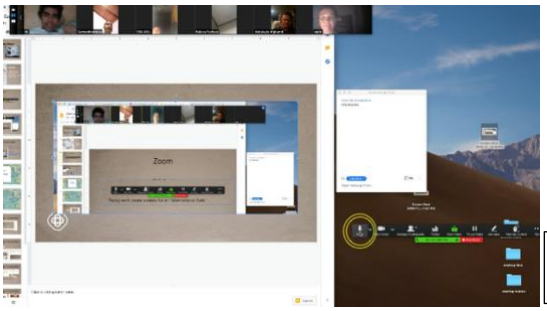
Extract 6.02.02 [*Searching for the trouble source producer 00:10:43*]

```
12  SAN:  i can hear not a single word+
    tea                              +moves cursor to own audio
          deactivation button
13  TEA:  we can hear ↑you
14  SAN:  oh (.) thank you#11
    tea                  deactivates his audio-->>
    fig
```



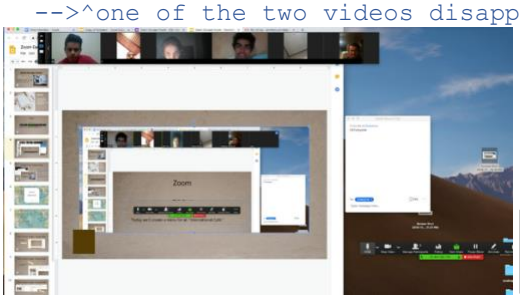
Line 14. #11: The teacher deactivates his own audio.

```
15  RON:  i can hear you*
                     *mild noise of people chatting in the
                           background of a participant-->
16        (4.0)
17  TEA:  i don't know#12
    tea                -->activates his audio
    fig
```



Line 17. #12: The teacher activates his audio.

```
18  JAS:  sounds like a tv:
19  MIK:  yeah that's coming from a tv (inaudible) he he
20  JAS:  ri:↑ght
21  MIK:  >yeah< he he
22  JAS:  can someone please turn off the tv:
23        ((laughter))
24  SAN:  i think that's mikki i think*
                                  -->* noise louder again-->
25        (2.0)(3.0)
    tea        deactivates his audio
    tea                activates audio
26  JAS:  *okay^#13
          -->* noise disappears
    suz     -->^one of the two videos disappears
    fig
```



Line 26. #13: one of SUZ's two videos disappear.

Extract 6.02.02 begins with a complaint from SAN of hearing difficulty using an extreme case formulation (‘i can hear not a single word’, line 12). As SAN’s turn reaches a possible completion, the teacher drags his mouse cursor from the instruction-giving area to his audio deactivation button and holds it there. The teacher does not click on it; instead, he replies to SAN’s complaint by assuring him that his sound is heard in class. SAN indicates a shift in the epistemic stance (Heritage, 1984a) (‘oh (.) thank you’, line 14). At the end of SAN’s turn, the teacher deactivates his audio (see #11). However, the noise persists during the 4.0-second silence that follows the teacher’s deactivation of his audio. Moreover, another mild noise - people chatting - occurs. The persistence of the disruptive noises indicates that they are not coming from the teacher’s end. The teacher then activates his audio (see #12) and explicitly expresses a negative epistemic status regarding the location of the trouble source producer (‘i don't know’, line 17).

Following this unsuccessful repair attempt to locate where the noise is coming from, another repair is initiated by JAS to narrow down the search for the trouble source producer (‘sounds like a tv:’, line 18). This is met by a confirmation from MIK (‘yeah that's coming from a tv’, line 19). Having received this confirmation, JAS redesigns the repair initiation as a request this time (‘can someone please turn off the tv:’, line 22); JAS’s repair initiation is met by laughter from the learners (line 23). Meanwhile, the noise volume is lower but still audible to the participants. SAN extends another opportunity for other-initiated self-repair, as he suggests that the noises are coming from MIK’s background (‘i think that's mikki i think’ line 24). However, his repetition of (‘i think’, line 24) indicates a degree of uncertainty. The noise gets louder again by the end of SAN’s turn, causing a disruptive echo in line 24.

The loud noise hijacks the floor for 5 seconds, during which the teacher deactivates his audio and holds for 3 seconds before reactivating it as the noise is still there, but in a lower volume. As the noise gets very low in volume, JAS orients to that in line 26 before a 2-second silence occurs. The noise disappears in line 26 in synchronisation with SUZ disconnecting one of her two devices (see #13), as it appears on the teacher’s screen. This disconnection may not have been seen by the participants as some of them might be on speaker view, which only displays the video of the current speaker. Indeed, this can explain the participants’ reference to MIK as the trouble source producer in the following extract, despite the fact that he was not.

Extract 6.02.03 is a continuation of the previous two extracts, which shows the resumption of the suspended class activity.

Extract 6.02.03 *[Resumption of the suspended class activity 00:11:20]*

```
27        (0.2)
28  SAN:  now he's gone he??
29  JAS:  yeah he heh
30  TEA:  Hmm (.)
31        was it mikki↑
32        (1.2)
33  TEA:  do you see the slide that
```

The extract begins by a 2.0-second pause following the disappearance of the disruptive noise. SAN suggests that ('`now he's gone he??`', line 28), in reference to MIK as the trouble source producer, which is met with an agreement and laughter by JAS in line 29. It is not clear why SAN and JAS assume that the disruptive noises are coming from MIK's end. The teacher does not seem to be sure in line 30 and wonders if it was actually MIK in line 31. After a 1.2-second pause, the noise is no longer audible, and the teacher resumes the suspended class activity ('`do you see the slide that`', line 33) by picking up from where he had left off.

In Extract 6.02, the trouble was caused by a noise resulting from a participant connecting from two devices simultaneously. There were multiple repair initiations by both the teacher and learners. The teacher used the audio activation/deactivation feature to test whether the echoey noise was coming from his end. The noise disappearing after his second use of the features in line 25 may explain his uncertainty when SAN and JAS assumed that MIK was the trouble source producer. It is also worth noting the collaborative work by the participants to identify the trouble source producer, despite the lack of mutual access. The participants worked to locate the trouble source while it was persistent, showing a preference for the progressivity of the suspended class activity, which manifests in the successive repair initiations.

In the previous two extracts, the self-repair was absent and extra work was carried out by the participants to repair the trouble. Extract 6.03 below shows a slightly different situation, whereby the teacher and learners search for the trouble source producer using the audio activation/deactivation features accompanied by verbal means, and the trouble source producer carries out self-repair after being identified. The difference lies in that the learner (i.e. candidate trouble source producer) is using the audio activation/deactivation features to check whether the noise emanates from her background.

Extract 6.03

Participants: Teachers: ROSsi, SALly, LIZ / Learners MOHammed, NADin.

Summary

This extract occurs in the class's opening phase, where the participants exchange greetings before the teacher introduces the discussion topic. This class involves three teachers and two learners. The reason this class has three teachers is that more learners are expected to attend and they will be placed into breakout rooms with teachers managing each room. However, only two learners showed up for this session, and the decision was made for all to stay in the main room. A TV noise from one of the participant's physical spaces is heard in the class shared space. The teacher orients to the noise and initiates repair by a direct information seeker to MOH. When MOH struggles to answer, an expansion insertion ensues to search for the participant who has the TV on. Another teacher intervenes and offers a candidate trouble source producer (NAD). NAD uses the audio activation/deactivation features to test whether her TV is causing the trouble. When confirmed that the noise was coming from her end, she carries out self-repair while the other participants resume the suspended class activity.

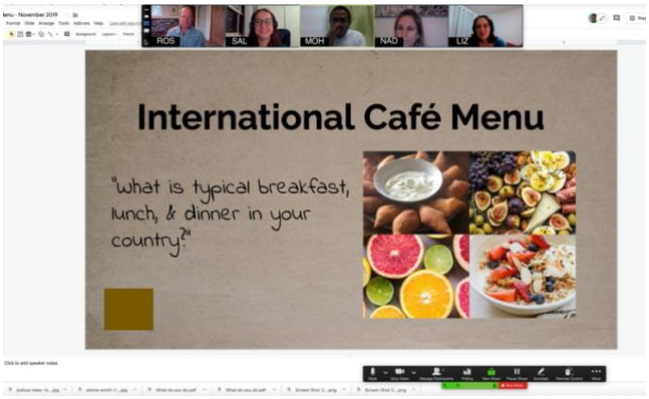


Figure 6.2: ROS's Screen Layout

Extract 6.03.01: [*Emergence of the background noise 00:01:48*]

```
01   SAL:  *ahm: (.) ↑great
           >>*tv noise in the background-->>
02   ROS:  thank you fr join=
03   MOH:  =so let's
04   ROS:  with computer today MOH (.)
05         hey do you hav=
06   MOH:  =$yeah i appreciate it$
07   ROS:  ahm
```

Prior to Extract 6.03.01, ROS had introduced MOH to the rest of the class, as they met for the first time while sharing his screen in preparation for the next phase of the class (see Figure 6.2

above). The background noise occurs before the beginning of the extract and continues, as indicated by the arrows in the extract above. Extract 6.03.01 begins with ROS thanking MOH for joining the class from a computer, as Google Slides work better on a computer than on a mobile phone. The first orientation to the continuous TV noise is in line 5 (‘hey do you hav=’,) by teacher ROS, before MOH’s expression of appreciation cuts it off (‘=$yeah i appreciate it$’, line 6). ROS acknowledges MOH’s appreciation using a minimal response token (‘ahm’, line 7). Note that the TV noise is not very loud throughout the encounter, but it is still audible in the class.

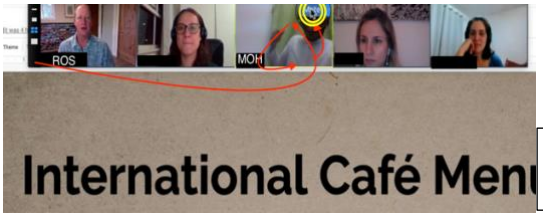
Extract 6.03.02: [*ROS initiating the repair sequence 00:01:56*]

```
08          (1.9)
09    ROS:  do you (.) ar- are there people in the background (.)
10          Ahm
11          (0.7)
12    ROS:  s- (0.4) is that is that your+#14 backg round MOH (.)&
                                              +points to MOH
```

Line 12. #14: ROS points to MOH’s

```
      ros                          &moves the mouse cursor to MOH’s video
      liz                                          deactivates audio-->>
13          +or is that(.)somebody else's& background
      Ros   +hovers cursor over MOH’s video-->

      moh                                  &turns and looks at his
            physical background
14          (1.1)
15          no it's mine&
      moh                 &points at his physical background
16          (0.9)#15(0.6)
      ros         deactivates MOH’s audio-->
      ros            -->activates MOH’s audio
```



Line 16. #15: ROS deactivates MOH’s audio.

```
17    MOH:  Oh
18    ROS:  ahm:
19          (1.7)%
      liz         %speaks with the mic off
20    MOH:  it's mine sir
      liz               -->activates audio
21    ROS:  amm (.) i was wondrin=
22    LIZ:  =no i think it's NAD:'s [(.)] tv:
```

```
23   NAD:                              [o:h]
24          is your tv on ↑NA:D*
     nad                        deactivates audio-->>
                              -->* noise disappears
25          (1.0)^(0.6)
     nad          ^mouths 'no?' with a head shake to right and lift
26   MOH:   it's not me it's not me (.) i have nothing in here*
     nad                                              -->  activates audio
                                                  *noise appears again-->
```

Following a 1.9-second pause in line 8, ROS reinitiates the repair by producing a polar question ('do you (.) ar- are there people in the background', line 9). As no candidate recipient to ROS's question is stated, a 0.7-second pause occurs in line 11, before ROS reformulates the question and addresses it to MOH ('is that is that your backg% round MOH', line 12). ROS points to MOH's video, as he assigns him as the question's recipient (see #14). This gesture adds ambiguity to the question and can explain MOH's misunderstanding in the subsequent lines. Following the pointing gesture, the teacher drags the mouse cursor towards MOH's video but does not deactivate MOH's audio yet, which can be described as an orientation to his preference for self-repair (see #15). After a micro-pause, the teacher adds more recipients to the question ('or is that(.)somebody else's& background', line 13). MOH turns to look at his physical background following ROS's enquiry, which means that he does not understand the teacher's question as a repair initiation. After a 1.1-second pause in line 14, MOH answers the teacher's question by claiming that the background is his while pointing to his physical background ('no it's mine', line 15). Despite MOH's apparent misunderstanding, ROS does not initiate repair to it. MOH does not understand ROS's question as an invitation to self-repair, and thus, it is not carried out.

As a result of self-repair absence, ROS moves on to a stronger repair device using the audio activation/deactivation features to identify the trouble source producer. He deactivates MOH's audio during a 1.1-second silence in line 16 (see #15). ROS then holds for 0.6 seconds to evaluate the effectiveness of the audio deactivation action before revoking the decision by activating MOH's audio, as the noise is still audible. Following this, MOH orients to deactivating/activating his audio by ROS using a minimal response token ('oh', line 17), which indicates the accountability of the teacher's decision. As shown in Figure 4.4 in Section 4.2.1, a notification appears on the learner's screen showing him that their audio is deactivated by the host (i.e. the teacher). ROS also produces a minimal response token ('ahm:', line 18), which may display uncertainty regarding who is feeding the TV noise into the class shared space. This is followed by a relatively long pause of 1.7 seconds in line 19.

During this pause, teacher LIZ appears to be saying something, but her audio is deactivated. MOH repeats that it is his background in line 20, while LIZ eventually manages to activate her audio. ROS also does not repair this misunderstanding trouble. Instead, he acknowledges MOH's response using a minimal response token ('amm', line 21) and begins to initiate another repair attempt ('(.) i was wondrin=', line 21). Before ROS reaches a possible turn completion, LIZ suggests a candidate trouble source producer ('i think it's NAD:'s [(.)] tv:', line 22). NAD produces a minimal response token ('[o:h]', line 24) in an overlap with LIZ's turn right after hearing her name, which displays a shift in epistemic status (Heritage, 1984). This may also indicate that NAD is not aware of the audibility of the TV noise coming from her end. Also, LIZ's turn design indicates a degree of uncertainty, which is then followed up by a confirmation check addressed to the new potential trouble source producer ('is your tv on NAD↑', line 24). In response to the question addressed to her, NAD deactivates her audio at the end of LIZ's turn (line 24). As LIZ's turn in line 24 is reaching a possible turn completion, NAD deactivates her audio. Consequently, the TV noise disappears from the class shared space. Following this, a 1.6-second pause occurs, during which NAD mouths 'no?' and shakes her head left and right as if to check whether the noise has gone. Her question is followed by MOH disclaiming the noise using an extreme case formulation (Pomerantz, 1986) ('i have nothing in here', line 26), which indicates that he now recognises the repair initiations. NAD activates her audio after a 0.6-second pause and the noise becomes audible again.

The following extract (6.03.03) displays the repair resolution and the resumption of the suspended class activity.

Extract 6.03.03: [*Repair resolution and resuming class activity 00:02:25*]

```
27  ROS:  Ahm
28  NAD:  ye:s= ((nods))
29  LIZ:  =it's nad's tv
30        (0.5)
31  NAD:  ahmm
32        (0.5)
33  NAD:  ^okay
    nad   ^goes to turn off the tv
34        (3.7)
35  ROS:  a:h [tv::
36  SAL:      [↑so tell us about your family moh↑
```

Following self-audio deactivation/activation, NAD acknowledges that the TV noise is coming from her end ('ye:s=', line 28). LIZ is now certain and produces a confirmation, which also works as an announcement that the trouble source is located ('it's nad's tv', line 29). Her

suggestion is proven right as she confirms that it is NAD's TV in line 28. A 0.5-second pause occurs in line 30 as NAD seems to be preparing to take off her headphones and stand up over lines 30-33. After uttering ('okay', line 33), she leaves the chair to handle the TV noise. In the subsequent turns, a 3.7-second silence occurs in line 34 before ROS produces a minimal response token indicating an epistemic shift followed by the actual trouble source ('a:h [tv::', line 35) as opposed to his initial repair initiation suggesting someone has people in their background. Teacher SAL resumes the suspended class activity, during which came the introduction of MOH with a request addressed to him ('so tell us about your family moh↑', line 36).

Extract 6.03, as analysed above, suggests that using the audio activation/deactivation features to identify the location of a trouble source and repair it is not exclusively carried out by the teachers; learners can also use these features. The difference lies in the fact that teachers are afforded the ability to deactivate the audio of others, while learners are only able to activate/deactivate their own audio. The analysis also shows the participants' commitment to the progression of ongoing class activities by working collaboratively to repair such troubles.

All three extracts examined here demonstrate a developing digital competence as these encounters occurred in the early recordings at the beginning of the small group conversation classes. Both teachers and learners were not experts in using Zoom for group classes. This can be seen in the participants' struggle to locate the trouble source producer, despite the existence of the affordance of the green box around the current speaker or the speaker with louder audio at a particular moment. This correlates with the notion that affordances do not determine the way interaction goes but that it is determined by the participants' use of these affordances (Hutchby, 2001).

The analysis above also shows the participants' preference for self-repair, as there is always a space provided for it before other-repair is carried out. The nature of the online setting in these encounters also plays a role in this preference as the trouble source producer is unknown due to the participant' lack of mutual access to each other's immediate environments.

6.4 Summary

This chapter has presented three encounters, which demonstrate the participants' use of the audio activation/deactivation features to locate the trouble source producer to repair hearing/speaking troubles. The chapter began by introducing the key terms that would be used in the analysis (Section 6.2). A post-analysis schematic representation of the pattern presented

in the chapter was then presented to improve the readability of the analysis. The chapter subsequently presented three encounters demonstrating the participants' use of the platform features of audio activation/deactivation to locate a trouble source producer and repair the trouble and restore the progression of the suspended class activity. The chapter showed the participants' extension of the repair sequence until the unknown trouble source producer is located and a successful resolution is reached. This extension was due to the participants' inability to identify the trouble source producer, sparking the need for collaborative work to resolve the trouble. The following chapter demonstrates shorter and smoother repair sequences, where the trouble source is known to the participants using the feature of audio deactivation.

Chapter 7. Analysis: Using Audio Activation/Deactivation Features to Manage Disruptive Noises from a Known Source

7.1 Introduction

The previous chapter presented an analysis of the use of audio activation/deactivation features by both teachers and the learners to search for the trouble source producers to repair hearing/speaking troubles caused by the interference of background noises in online video-mediated L2 classes. The current chapter demonstrates examples of shorter repair sequences discovered in the data. The reason these repair sequences are shorter than the ones shown in the previous chapter is that the trouble source producer is known to the participants, whereas in the longer sequences, it is unknown. The ability to quickly locate the trouble source producer results in shorter and smoother repair sequences and, thus, less hindrance to the progression of the ongoing class activity. The chapter will argue that there is a reflexive relationship between the participants use of the audio activation/deactivation features and the management of repair in L2 video-mediated classroom interaction. In addition, the chapter will argue that the development of the participants' CIC and e-CIC plays a role in advancing the progressivity of interaction and creating more opportunities for participation and learning in L2 video-mediated classrooms.

This chapter is organised as follows. Section 7.2 presents extracts where the participants use the audio activation/deactivation features to repair speaking or hearing troubles caused by the interference of a fellow participant's background noises. This is followed by Section 7.3, which presents extracts displaying the participants' use of the audio deactivation feature as a 'preemptive move' (Schegloff, 1986, p.133) to potential troubles that can hinder the ongoing class activity. Section 7.4 then offers a summary of the analytical findings presented in the chapter.

7.2 Audio Deactivation as a Repair Device to a Known Trouble Source Producer

This section presents a number of extracts representing a pattern found in the data where the participants use the audio deactivation feature to carry out self- and other-repair to handle trouble sources that impede hearing or speaking during online video-mediated class activities. The audio deactivation feature is used to block the audio channel of (self) or (other) participant(s) whose background noise or audio trouble momentarily interrupts the progressivity of the ongoing class activity.

As Jenks (2014) points out, background noises can be consequential in that they can halt the ongoing talk, and participants can make them interactionally relevant by topicalising them. However, participants may carry out repair without topicalising the trouble source, resulting in a minor hindrance to the progressivity of the ongoing activity (Brandt, 2013). In the present data, the participants repair troubles impeding speaking or hearing in a similar manner by simply deactivating the audio of the trouble source producer. As established in the previous chapter, the term 'trouble source producer', as already noted in the previous chapter, is used to label the participant from whose end the disruptive background noise emanates. This labelling is to ease the description of the data despite the fact that noises can be in some cases be caused by other people's or pets' activities, which the participants do not produce themselves and may not have control over.

Figure 7. 1 is a retrospective summary of the extracts that will follow in this section and which demonstrate the pattern whereby the participants use the affordance of audio deactivation to repair speaking/hearing troubles caused by interference of background noises. Compared to Figure 6. 1 in the previous chapter, the repair sequence is noticeably shorter, and the progress of the class activity is restored much more quickly.

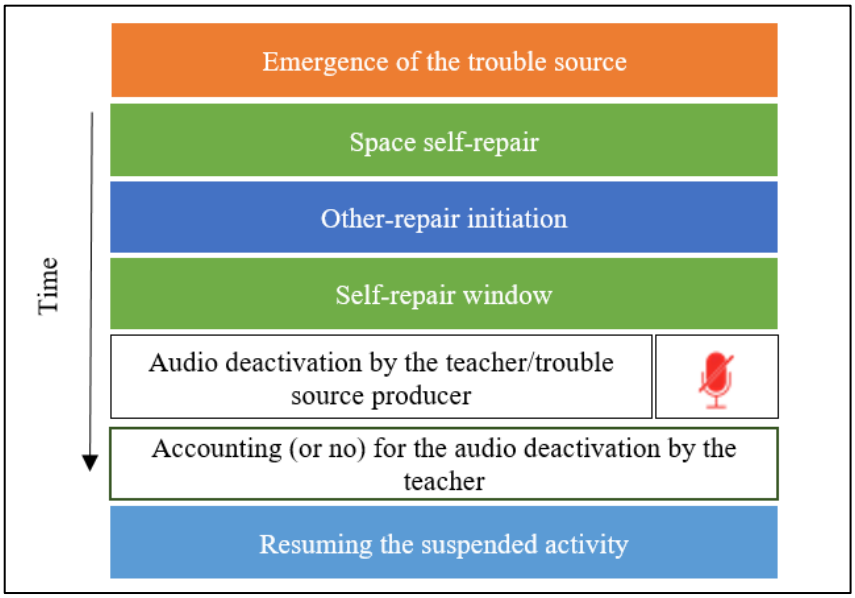


Figure 7. 1: A Schematic Representation of the Shorter Repair Sequences in this Section

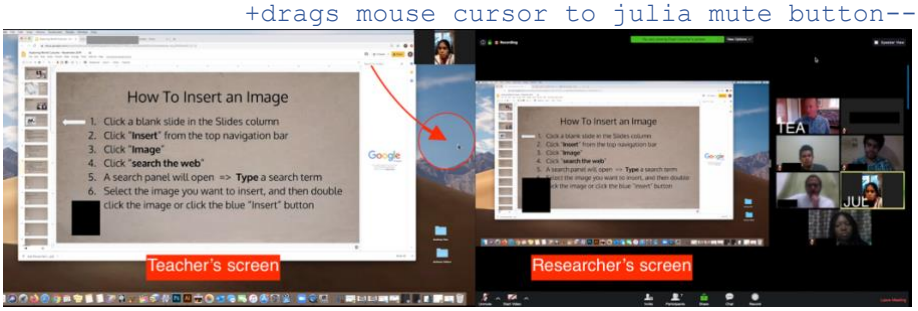
Extract 7.01 presents an example of these shorter repair sequences as disruptive noises which appear in the class while the teacher is giving instructions at the beginning of the class intending to ease the learners' work in the coming class activities. The teacher is sharing his screen at the beginning of the class time to explain to the learners how to search the web to insert images in their assigned Google Slides. These images will be used to drive the discussion later when they are put in breakout rooms. Since the class started, JUL had noises

of people talking (or TV), and passing by cars in her physical proximity. However, both the teacher and the learners did not orient to these noises as problematic, possibly because they were fleeting on some occasions. The teacher started listing the steps that learners would need to follow, accompanied by a live show on screen (lines 1-3).

Extract 7.01: [*it might be distracting 00:08:50*]

```
01  TEA:  so we >will go< insert image (.)
02        search the web (.) and when you insert (.) image (.)
03        >sea*rch< the web (.) opens up this side p-*
04            *background noise -->
                                        * indefinite explosive noise
05        (1.4)
06        ((people talking in JUL's background))+#1
                          +drags mouse cursor to julia mute button-->
```

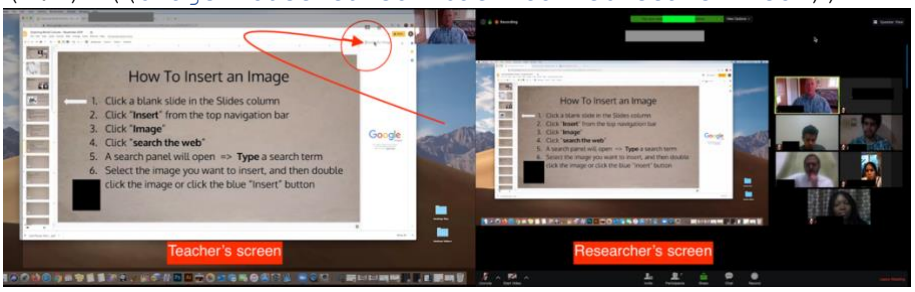
fig



Line 6. #1: The teacher drags the cursor from the search side panel to JUL's audio deactivation button.

```
07        (4.5)
08  TEA:  sorry julia
            -->deactivates jul's audio
09        (1.0)
10        julia i muted you
11        because: (.) ahm (.) i (.) a_
12        i feel like ahm: (.)
13        >it might be< distracting↓
14        (2.2)#2((drags mouse cursor back to web search icon))
```

fig



Line 14. #2: The teacher drags the cursor back to the search side panel after deactivating JUL's audio.

```
15        .hhh
16        °a:nd anyhow° (.) if you open up the side panel (.)
17        and then you can type for what you wanna type
```

As the teacher is giving the instructions, JUL's background noises appear again at the beginning of line 3. Up to this point, the teacher continues with the instruction-giving by decomposing his instructions to make them mimicable by the learners (Due et al., 2019)

(“>sea*rch< the web (.) opens up this side p-*”, line 2), and it is clear that the noise is not being treated as trouble yet, as there is no orientation to it. However, an indefinite explosive noise occurs in line 3, cutting off the teacher’s turn and hindering its completion. The teacher’s turn projects that more steps to add images into the slides were to follow but this is interrupted by the loud explosive noise.

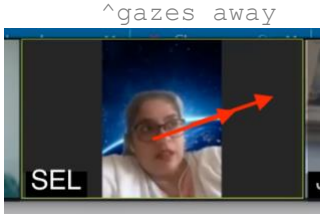
A 1.4-second silence occurs in line 5 before a more audible noise of people talking in JUL’s private space takes the floor in line 6. The teacher pauses at hearing the disruptive noise and holds the mouse cursor at the search side panel for 1.4 seconds (#1). The teacher does not take action during this pause, which may be seen as waiting for JULa to carry out self-repair (i.e. deactivate her audio). However, the self-deactivation of audio does not occur. The self-repair space is not exploited by the trouble source producer and another noise is heard in line 6. The teacher then drags the mouse cursor (line 6) from the search side panel, where it was being held towards JUL’s audio deactivation button. As he approaches JUL’s deactivation button, the teacher accounts for the audio deactivation action by apologising to Julia (‘sorry Julia’, line 8) while he clicks the button. This is followed by a 1.0-second pause in line 9, during which the teacher moves the cursor away from Julia’s audio deactivation button. The teacher then extends his accounting for his audio deactivation decision. He begins by announcing the decision (“julia i muted you”, line 9), and adding the logic behind the decision in an apologetic manner, as seen in the pauses, the low intonation, together with the hesitation both before and while he offers his reasoning (“because: (.) ahm (.) i (.) a”, line 11) and (‘i feel like ahm: (.)>it might be< distracting↓’ lines 12-13). The teacher then drags the cursor back to the search side panel (#2), which is the departure point before the initiation of the repair, and takes an in-breath. This projects a closing of the repair sequence and a resumption of the suspended class activity. Indeed, the teacher resumes the class activity in lines 16-17 and picks up from where the activity had been suspended.

In Extract 7.01, the repair sequence is notably shorter than the ones demonstrated in Chapter 6 since the teacher has epistemic access to the identity of the trouble source producer. Although the teacher has access to the feature of deactivating the others’ audio, the deactivation decision is delayed. This shows the teacher’s orientation to a preference for self-repair. Figure 7. 1 demonstrates the presence of space for self-repair after the disruptive noise emergence and after halting the progressivity of the ongoing class activity. When this window for self-audio deactivation is not exploited by the trouble source producer, other-repair is carried out by the teacher, followed by an account for the audio deactivation action.

Another example of this pattern is demonstrated by Extract 7.02, where three learners and their teacher are working on a fill-the-gap activity while discussing a written opinion shown on the teacher's shared screen. Prior to the extract, the teacher asks the learners about their thoughts of the stated opinion on the screen. SEL self-selects and shares her opinion with the class. The extract begins at the end of the dyadic exchange between the teacher and SEL. During this exchange, a TV noise is heard in SEL's background while she speaks, becoming clearer during pauses. This is, however, not treated as trouble when SEL is the current speaker or the main recipient of the teacher's explanation.

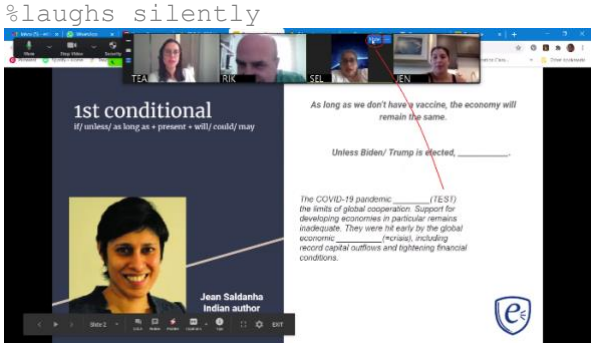
Extract 7.02 [*it's too loud 00:28:32*]

```
01  TEA:  *↑right+
          >>*tv noise appears-->
02        (2.0)
03  TEA:  @am[hm:
          @nods slowly-->>
04  SEL:     [Ok
05        (2.0)*^#3(3.0)
    sel          ^gazes away
    fig
```



Line 5. #3: SEL gazes away from the screen.

```
06  TEA:  @Oka:y↑
          -->@
07        %do you understand the meaning of+#4the:
                                          +moves the cursor to
          sel's mute button and holds there
    rec   %laughs silently
    fig
```



Line 7. #4: Moving the cursor to deactivates SEL's audio

```
08        (1.4)
09  TEA:  all the >words here<↑*
                              *deactivates SEL's audio
                            -->*tv noise disappears
10  RIC   ↑SEL
11  TEA:  >i i muted< you ↑SEL because it's too loud the:
12        (1.0)
13        your sound (.) oka^y↑^
    sel                      ^nods in agreement
14        >it's too loud yeah↑<
15        (3.8)
16  JEN:  Yeah
17        i i think ((clears throat))
18        i think it's ↑ha:rd
```

The extract begins after the teacher explains to SEL that what had previously been referred to as 'the third world' is now described as 'developing countries'. The teacher then produces an understanding check ('↑right', line 1), which is followed by a 2.0-second gap in line 2. The understanding confirmation is not yet produced by SEL. The teacher produces a minimal response token ('am[hm:', line 3), which overlaps with SEL's confirmation of understanding of the teacher's elaboration before the extract. As an overlap solution, both speakers drop back, resulting in a 5-second silence in line 5. During this long gap, the TV noise coming from SEL's background becomes louder, holding the floor during this period. In this silence, SEL appears to look away, possibly at the TV, while the teacher slowly nods (#3). The teacher's slow nodding during this relatively long silence can be seen as a pursuit of further participation by SEL. It can also be seen as giving space for self-audio deactivation by SEL. However, SEL is gazing away and does not produce either further talk or a self-audio deactivation (#4).

The teacher then self-selects and produces a transition marker ('Oka:y↑',) (Svennevig, 2012). This is followed by a polar question from the teacher ('%do you understand the meaning of+ the:', line 7). At the onset of the teacher's turn in line 7, RIC seems to be laughing silently. This can be seen as an orientation to the noise coming from SEL's background. As the teacher produces the question in line 7, she initiates a mouse cursor movement from the task area on the screen towards SEL's audio deactivation button, as seen in (#5). A 1.2-second pause follows this in line 8, during which the teacher holds the cursor over SEL's audio deactivation button before she completes the question in line 9. As the polar question reaches a possible completion, the teacher then deactivates SEL's audio right at the completion of the polar question, and the disruptive TV noise subsequently disappears. The audio deactivation is overlapped by RIC calling SEL's name (line 10) following his silent laughter, which can be described as a repair-initiation.

Following the audio deactivation, the teacher immediately self-selects for the next turn and accounts for the managerial action by explaining the reason that occasioned it ('>i i muted< you ↑SEL because it's too loud the:', line 11). A 1.0-second pause occurs before the teacher continues with the account ('your sound (.) oka^y↑', line 13). SEL responds by nodding in acceptance with the teacher's decision. As the teacher's turn is reaching a possible completion in line 14, RIC also nods in agreement and follows this up by agreeing verbally ('>it's too loud yeah↑<', line 14), which might explain his laughter in line 7. Following

the audio deactivation action and repairing the trouble caused by the interference from disruptive background noise, a 3.8-second gap occurs in line 15. JEN then self-selects for the next turn (lines 16-18) and resumes the progressivity of the suspended class activity by answering the teacher's question, asked in lines 7-9, before departing the activity to repair the trouble.

This extract also shows that the teacher is prioritising the progressivity of the class activity when the trouble source producer is the current speaker or the selected next speaker. However, it is a different scenario when the trouble source producer is not the current speaker. Extracts 7.01 and 7.02 demonstrate an immediate accounting for the decision to deactivate a learner's audio by the teacher. Nonetheless, this accounting is not always present, as the teachers in many other instances seem to have routinised the audio deactivation as a resource to repair hearing/speaking trouble caused by interference of background noises. Extract 7.03 presents an example of this.

In Extract 7.03, the noise of a non-participant talking to one of the learners disrupts the progressivity of the ongoing class activity, which causes the interaction to break down momentarily. The class involves three learners and their teacher, who enables the gallery view without a screen share. Extract 7.03 begins with MAR announcing that he has a question ('`i have a question`', line 1), which normatively invites a type-matched second pair part (i.e. either a go-ahead or a denial response by the teacher) (Schegloff & Sacks, 1973).

Extract 7.03 *[talking to a non-participant 00:15:28]*

```
01  MAR:  i i have a question [yes
02  TEA:                      [yes dear
03  MAR:  one person used this expression*#5 (.)
    sol                                *gazes up
    fig
```

Line 3. #5: SOL gazes up

Line 4. #6: SOL gazes back to screen

```
04        This#6 word^ (.) like ↑nevertheless
    sol          ^turns off webcam
05  TEA:  ah↑a:*
06        (0.5)
               *non-participant calls sol in the background
07  MAR:  a[::
08  SOL:   [ahm.((responding to the non-participant))
```

```
09          ((non-participant speaks in Portuguese))
10          (0.6)
11  SOL:    +#7[ta ((Portuguese. Responding to the non-participant))
    tea     +drags mouse cursor to sol's mute button
    fig
```

Line 11. #7: The teacher drags the mouse cursor to deactivate SOL's audio

```
12  MAR:      [what's the the the:
    tea                          deactivates sol's audio
13          (1.9)
14  MAR:    the best use of nevertheless (.)
15          in the beginning (.) or in the middle (.)
16          or in in the end
```

The teacher grants MAR the green light to ask his question in the next turn (line 2). MAR produces a storytelling formatted pre-question ('one person used this expression', line 3). Such prefatory work enables the current speaker to hold the floor longer and suspends the transition to the next speaker (Drew, 2005). During this pre-question turn in line 3, SOL gazes up before gazing back to the screen and deactivating her video but leaves her microphone on (line 4, #5 and #6). This may be explained as an anticipation of an oncoming engagement with a non-participant or expecting a kind of interruption, and thus, she blocks the visual link with the class. Indeed, a sound of a child is heard but is not treated as trouble. The teacher produces a minimal response token, functioning as a continuer to MAR's pre-question turn ("ah↑a:", line 5), which invites MAR to complete his question. A 0.5-second gap occurs in line 6 before MAR launches his turn, which is overlapped by the sound of someone in SOL's physical space calling her name; a sound which is clearly heard in the class shared space.

The interference of the non-participant's sound briefly occupies the conversational floor, which the teacher has allocated to MAR to complete his question. MAR does hold the floor briefly ('a::', line 07) before he drops out, as his turn initiation is overlapped by SOL's minimal response to the non-participant ('ahm.', line 8) in her physical space. By holding the floor, but then conceding, MAR orients to the interruption caused by the interference of sounds from SOL's physical space and treats it as problematic. Following this, a 0.8-second pause occurs in line 7, which marks a breakdown in the dyadic talk between MAR and the teacher. This pause is followed by a continuation of the non-class exchange in Portuguese between SOL and the person in her physical space in lines 8-9. A 0.6-second pause occurs

following this non-class-related exchange. The teacher drags the mouse cursor from the top of the screen to SOL's audio deactivation button at the end of this pause (see #7). The mouse cursor movement overlaps with both SOL talking to the non-participant ('[ta', line 11) and with MAR self-selecting to resume the momentarily suspended activity ('[what's the the+ the:', line 12). MAR's self-selection after the 0.6-second pause indicates that he judges the non-class related exchange to be over and the suspended activity is to be resumed. However, he overlaps with SOL's closing of the exchange with the non-participant, resulting in MAR dropping out again. The teacher then clicks on SOL's audio deactivation button to end this interference of SOL's private conversation with the non-participant.

Following this audio deactivation action by the teacher, a 1.9-second pause occurs in line 14, which marks the ending of the interference of the non-related exchange. MAR self-selects again and picks up asking the incomplete question from precisely the point he had stopped at ('the best use of nevertheless', line 14). MAR's completion of the interrupted question marks the repair success, and thus, a resumption of the momentarily suspended class activity.

In Extract 3, SOL may have anticipated that someone is coming into her physical space in which she is engaged in an online class, as she gazes away in line 4 and blocks access to her video. Indeed, this anticipation can be seen in the sounds from her physical space afterwards. SOL also uses minimal response tokens only when talking to a non-participant, possibly to reduce engagement in the side activity. Yet the audio channel remains active, causing the conversation breakdown, and she is, thus, held accountable when her audio is deactivated by the teacher. The teacher does not account for the audio deactivation action. Instead, the floor is opened to MAR to resume the completion of his question, which results in less hindrance to the ongoing activity. This extract differs from the previous ones in the sense that the trouble source producer is, in fact, verbally talking to a non-participant. Conversely, in all the other extracts, the participants are held accountable for noises in their physical spaces produced by other people, pets, or objects, such as TVs, which they may have no control over.

Contrary to the extracts in Chapter 6, the repair sequence in Extracts 7.01and 7.02 is smoother. For example, the participants do not topicalise the trouble source, but simply use the audio deactivation feature to block the audio channel of the trouble source producer to allow the suspended class activity to be resumed. Extract 7.04 further illustrates this point of smoother and shorter repair sequences, where there is no accounting for the audio deactivation feature after the participants have experienced interference from a noise produced by a child of one of the participants.

Extract 7.04 [*drug trafficking 00:26:20*]

```
01  TEA:  the penalty is worse you mean↑
02  ROD:  ye:s (.) the [penalty is worse
03  MAR:               [◦yeah◦
04  TEA:  for drug* trafficking than for killing people↑
    jol           *activates audio-->
05  ROD:  *yes (.) [yes
    jol   -->*deactivates audio
06  TEA:           [((gasps
07  ROD:  and i[even
08  MAR:       [the*((inaudible)) they have own^ (.)
    jol            *activates audio-->
                                          ^child noise
09        their own ↑law
10        (2.0)
11  TEA:  what↑
12  MAR:  ahm (.)
13        the ((inaudible)) group% of traffic
    tea                         %moves mouse cursor to deactivate
          jol's audio
14        *#8they have own law
    tea   *deactivates jol's audio
    fig
```

Line 14. #8: The teacher deactivates JOL's audio

```
15        (0.9)
16  MAR:  they killed
17        (0.9)
18  TEA:  Oh (.) but he's talking about the la:w
19        like the official law
```

In Extract 7.04, the class discusses the punishments for drug trafficking and murder in their countries. Before the extract, JOL had been activating/deactivating her audio to take turns during class activities. This explains the multiple audio activation/deactivation in this extract in lines 4 and 5, which was previously discussed in Chapter 5. Extract 7.04 begins with the teacher uttering a confirmation check to the information ROD had mentioned prior to the extract ("`the penalty is worse you mean↑`", line 1). ROD produces the SPP to the teacher's question by positively confirming that information ("`ye:s (.) the penalty is worse`", line 2). MAR also produces a low intonation SPP in line 3 by positively answering the teacher's question. The teacher follows this up with another confirmation check ("`for drug* trafficking than for killing people↑`", line 4). As the teacher asks the

confirmation check question in line 4, JOL activates her audio. ROD produces the confirmation ("yes (.) [yes", line 5) that the penalty is worse. JOL deactivates her audio as soon as ROD takes the floor in line 5. JOL's audio deactivation may be attributable to an inability to take the turn as the floor is already taken by ROD. Moreover, the shared screen may have contributed to the invisibility of the audio activation to the other participants. Following ROD's confirmation, the teacher gasps, indicating a change of state from not knowing to knowing that the penalty for drug trafficking is worse than that for murder (line 06).

The interactional floor at this point is open, and ROD self-selects and initiates a follow-up turn, which projects more elaboration on the information that he has just produced ("and i[even", line 7). However, MAR also self-selects for the next turn, which results in overlapping with ROD's turn in line 8. JOL also seems to self-select for the next turn by activating her audio in line 8. The overlap is resolved when ROD concedes and leaves the floor to MAR, whose turn also projects more talk ("[the* ((inaudible)) they have own^ (.) their own law", lines 8-9). Before MAR reaches a possible turn completion, a child's noise in JOL's physical space appears in the class shared space. A 2.0-second pause follows the interference of JOL's background noise before the teacher initiates an open class other-repair (Drew, 1997) ("what↑", line 10).

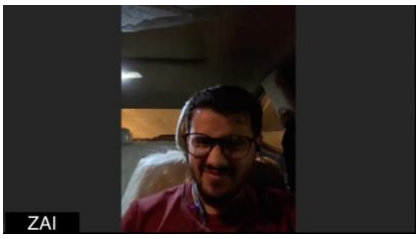
In response to the teacher's repair initiation, MAR carries out a self-repair by producing a repetition of his prior turn with slight modifications. At the onset of MAR's turn in line 13, the teacher moves the mouse cursor from the middle of the screen to JOL's audio deactivation button. The teacher then deactivates JOL's audio while MAR is carrying out his self-repair in line 13 (Figure 8). The teacher's audio deactivation action indicates that the trouble source here causing the hearing difficulty is not the overlap preceding it in lines 7-8, but the interference of the child's noise in JOL's physical space. MAR holds the floor after 0.9 seconds of silence in line 15 by picking up the repetition turn ("they killed", line 16). Although MAR's turn projects more talk to come, he pauses for 0.9 seconds in line 17, which is a relatively long pause. As a result, the teacher self-selects for the next turn and begins by producing a minimal response token ('Oh', line 18), indicating a shift in the epistemic stance (Heritage, 1984a) followed by a statement of disagreement ("but he's talking about the la:w", line 17). Reengaging with the suspended class activity marks the completion of the hearing trouble repair caused by the interference of JOL's background noises. It also marks the resumption of the momentarily suspended class activity.

In Extract 7.04, it can also be seen that participants hold JOL accountable (by deactivating her audio) for a noise that she does not personally produce, but she is deemed to be the trouble source producer since that disruptive noise comes from her end into the class shared space. The extract also demonstrates the participants' orientations to their preference for self-repair, as there is an opportunity for this both before and after the verbal indication of trouble. When the self-repair does not occur, other-repair is carried out by the teacher, who does not account for this decision afterwards. Not accounting for the audio deactivation by the teacher enables a quicker resumption of the progressivity of the class activity.

In contrast to the previous extracts, where other-repair prevails over self-repair, Extract 7.05 shows a different repair trajectory, with the trouble source producer initiating and carrying out self-repair.

Extract 7.05: [*dream home 00:50:31*]

```
01  TEA:  and would you like to [live abr-
02  ZAI:                        [because a:
03  TEA:  go ahead
04        (2.2)
05  ZAI:  yeah (.) but because in the mountain there is no service
06        (1.1)
07  ZAI:  there is no::*^
                       *loud dog barking noise
    joh                 ^startles
08        (1.0)&#9(0.5)
    zai        &frowns face
    joh               deactivates his audio((mouths sorry))
    fig
```



Line 8. #9: ZAI 'frowns.

```
09        there's $no signal for phones$+
    tea                                 +smiles
10        there is no a: a: no roads
11  TEA:  oh ↑really
12        it's [totally isolated
13  ZAI:       [you hafta
```

The class discussion revolves around where the participants live and if they would like to travel or live abroad. ZAI talks about a place in the mountains where he used to live before the beginning of the extract. In line 1, the teacher marks ZAI's talk as complete and moves on to the next question regarding living abroad. However, ZAI seems not to have completed his answer and opts for more elaboration ("`[because a:`", line 2). ZAI's extended answer

overlaps with the teacher's transition to the next question, which is resolved by the teacher's concession and by granting the green light to ZAI to extend his turn ("go ahead", line 3). After a 2.2-second silence, due to the overlap trouble, ZAI retakes the floor in lines 5-7. In line 7, ZAI's turn projects more talk to come ("there is no::"). However, a sudden loud dog barking noise occurs and hijacks the conversational floor. The dog barking noise startles JOH and suspends the ongoing class activity for 1.5 seconds in line 8. During this pause, ZAI frowns in orientation to the disruptive noise (#9).

After his startled reflex, JOH carries out self-repair by deactivating his audio and accounting for his action by apologising afterwards, as seen in line eight. Although his audio is deactivated, JOH mouths the word 'sorry' and makes it accessible to the other participants. This shows that the participants hold themselves accountable for any noises in their physical space which are heard in the class shared space. The suspension of ZAI's turn (line 8) projects the need and preference for self-repair to be carried out. The teacher does not move the mouse cursor to JOH's audio deactivation button despite having access to this feature. As the repair is then carried out, ZAI self-selects and retakes the conversational floor with a smile on his face ("there's $no signal for phones$+", line 9). The teacher mirrors ZAI's smile while he resumes the momentarily suspended class activity.

In all of the previous extracts in Chapter 6 and Chapter 7, the audio deactivation feature was used to repair trouble that has already occurred. Put differently; the ongoing class activity experienced a breakdown caused by interference from background noises. Another interesting pattern of using this feature shows that teachers use it to pre-empt potential trouble that can be caused by a participant's sound or background noise. The next section offers an examination of three extracts to demonstrate the teachers' use of the audio deactivation feature of Zoom as a preemptive procedure.

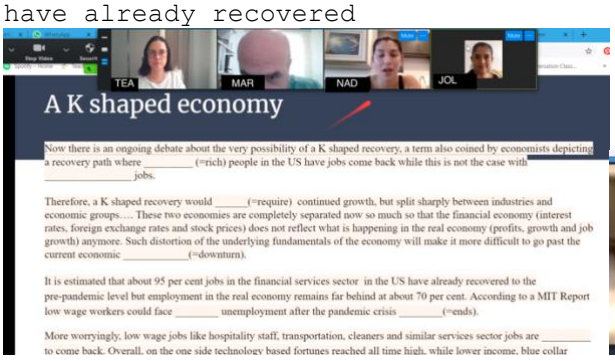
7.3 Deactivating Learners' Audio as a Pre-emptive Procedure

Pre-emptive procedures are practices used by the participants in anticipation and prevention of potential imminent trouble (Schegloff, 1986). The previous sections in the current chapter show that the features of audio activation/deactivation are used to repair an already existing trouble. However, the data also demonstrates another interesting use of these features (i.e. as a pre-emptive procedure, or a "preemptive move") (Schegloff, 1986, p. 133). They are used by the teachers in the data to pre-empt potential visibility trouble in a showing sequence or to help a learner to maintain the conversational floor. The current section presents extracts to illustrate this use.

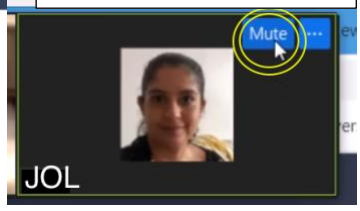
Extract 7.06 includes three learners with their teacher. The class is working on a task that requires each learner to read a paragraph from a passage which appears on the shared screen (#11). Prior to the extract, JOL had read a paragraph and at its completion the class stopped for a discussion of the content and grammatical forms. Lines 1-8 display the closure of JOL's paragraph discussion and show the teacher marking it complete.

Extract 7.06 *[my family: 00:34:06]*

```
01  TEA:  *you have to say (.)
           >>*tv noise in jol's background-->
02         it's [raining
03  JOL:       [it's raining
04  TEA:  ↑ye:s
05  JOL:  ah:ah
06  TEA:  Amhum
07  JOL:  it is raining
08  MAR:  [°it's raining°
09  TEA:  [go ahead ↑nadia
10  JOL:  °cool°
11         (3.2)
    nad    looks at the screen
    tea    drags cursor to the paragraph beginning
12  NAD:  okay (.)
13         it is estimated (.) that about (.) 95 percent jobs in the
14         financial #10&services sector*#11(.) in the united states (.)
    tea                &drags cursor to deactivate jol's audio
    tea                                deactivates jol's audio
                                      -->*background noise disappears
15         have already recovered
    fig
```



Line 14. #10 &#11: TEA drags the cursor to deactivate JOL's audio



In line 9, the teacher allocates NAD the turn to read the next paragraph ('`go ahead ↑nadia`'). JOL is heard making a low intonation assessment to the learnable she has just attained in line 7, before a 3.2-second gap occurs. During this gap, NAD appears to look at the screen, possibly to locate where to continue the reading. In this gap, the teacher also moves the cursor to the beginning of the paragraph that NAD has been allocated to read. NAD appears to locate where to start reading as she announces ('`okay`', line 12), and follows this up by reading the allocated paragraph. Meanwhile, JOL's background noise is still heard but not treated as trouble by the fellow participants as NAD reads (line 13). It is worth noting here

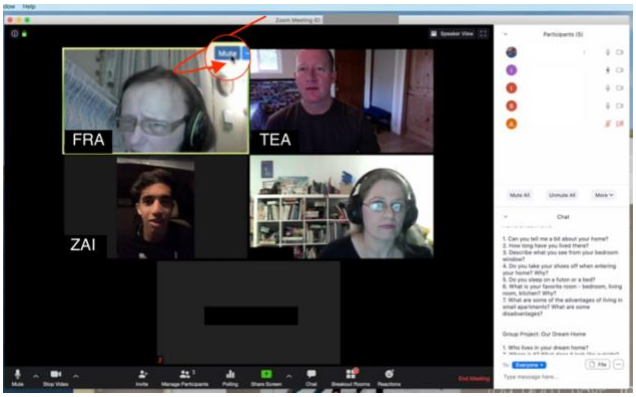
that JOL's background noise was also heard when she was participating in the previous lines. However, it is commonly observed in the data that when the person with the background noise is the current speaker, the teachers and fellow learners do not treat this as problematic. In addition, the way the platform works plays a role in this as it prioritises the sounds closer to the microphone. This means that when the participant speaks, his/her voice becomes prevalent and the background noise becomes louder only during pauses and gaps.

Since JOL is no longer speaking, the only audio coming from her end is the TV noise. It is not loud enough to cause a conversation breakdown, and NAD's reading is going smoothly. However, in line 14 the teacher drags the mouse cursor from the task area towards JOL's audio deactivation button (#10) and clicks on it (#11). As a result, the TV noise disappears from the class shared space. The question to ask here is *Why this now?* given that there is no breakdown in the interaction and the task work is carried out without interruption. The teacher's decision to deactivate JOL's audio after the end of her participation is for pre-empting potential hearing trouble, as well as to enable NAD to maintain the floor. Enabling NAD to maintain the floor can maximise her opportunities to use the target language, as opposed to having speaking or hearing trouble caused by interference from a fellow learner's background noise. Similar to the previous extracts, it is also possible to see the delay of undertaking this preemptive move as preference for self-audio deactivation. When JOL does not deactivate her audio after completing her participation, the teacher takes the initiative to do it. NAD completes her reading of the paragraph without trouble.

Another example of using the audio deactivation feature as a pre-emptive move to potential trouble is demonstrated by Extract 7.07, where one of the participants' sound is treated as a possible threat to the progressivity of the ongoing class activity.

Extract 7.07 *[reading the questions 00:02:21]*

```
01  TEA:  how *lo:ng have you lived there (.)
    fra       *reads questions quietly-->>
02        (0.9)
03  TEA:  ((snaps fingers and points to ZAI))
04        zaid (.) you wanna go with number ↑three:
05        (1.4)
06  ZAI:  ↑yes (.)
07        *#12describe what you see (.) from (0.3) your bedroom window
    tea   deactivates FRA's audio
    fra   -->* reading sound disappears
```

Fig 

Line 7. #12: The teacher drags the mouse to deactivate FRA's audio.

08 TEA: ↑yes

This extract is from the beginning of the class activity after the participants had been placed in breakout rooms. The teacher suggests that they take turns reading the questions from the list in the chat box. The teacher notices that FRA, who is reading the question, is also writing, so he notes that she does not have to write anything down. FRA takes the initiative to read the first question before the teacher reads the second (line 1). Despite the teacher's suggestion, FRA appears to be reading the questions in a quiet but remains audible voice in the class shared space, which overlaps with the teacher's turn (line 1). Following the teacher's turn completion, he snaps his fingers and points to ZAI, offering him the next slot and assigning him as the next speaker ('zaid (.) you wanna go with number ↑three', line 4). As noted in Extract 7.07, FRA is still reading the questions in a low voice during the teacher's assigning of the next participation slot to ZAI. However, this does not seem to cause a conversation breakdown, as ZAI takes the floor in line 6 and begins to read the question (line 7). The teacher then drags the mouse cursor and deactivates FRA's audio, enabling ZAI to complete his turn without risking a speaking trouble (#12).

Extract 7.07 presents another example where the class activity is ongoing, but the teacher pre-empts the potential emergence of a trouble source (i.e. quiet reading of the task questions) by temporarily deactivating FRA's audio. The teacher then assigns reading the next question to FRA and follows this by activating her audio to enable her to take the next turn (not in the transcript). Moreover, similar to Extract 7.06, a clear audio is instrumental to ensuring the progressivity of the ongoing activity as competing sounds may result in an interactional trouble (Jenks, 2014).

Another example of using the audio deactivation feature by the teacher as a preemptive move is presented in Extract 7.08, where a learner's background noise is threatening the visibility in a showing activity by a fellow learner.

Extract 7.08 *[fathers' day 00:25:11]*

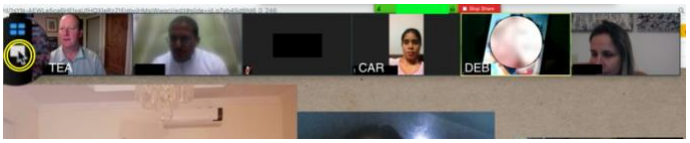
Extract 7.08 comes from a class that involves four learners with their teacher, who is sharing his screen to show a shared Google Slide on which the learners insert photos of their family members and comment on them. The task requires each learner to paste photos in a dedicated slide. The aim is that these photos will drive a discussion and create more speaking time for the learners. Prior to the extract, CAR had managed to insert some family photos and talk about them and address questions/comments by fellow participants. The turn then transitioned to JEN, who had trouble installing the Google Slides application; therefore, the teacher suggested that she can share the photos by bringing the phone closer to the webcam and showing them to the class. The platform's feature of 'speaker view', which displays only the current speaker (or possibly the one with background noises) on the screen while the teacher shares the screen is a potential threat to the progress of the showing sequence. To elaborate, CAR's background noise makes her video overlap with JEN's video when the photos are shown. To pre-empt a potential interaction breakdown, the teacher needs to deactivate CAR's audio. The reason the audio deactivation is seen as a preemptive move and not a repair initiation is that there is no orientation to this as problematic by any of the other participants.

```
01  TEA:  debbie (.) do you wanna sh- (.)
          Do you wanna show us* ↑some of your pictures (.)
    deb                         *looks down to her phone-->>
02        can you just show'em on your ↑phone
03        (1.5)
04  DEB:  ↑yes+
                +barking dog noise-->>
05        (1.2)*(1.2)
06  DEB:  give me one <second>•
07        *^(4.7)^
          -->*
    deb   ^gazes away and talks to someone^
08  DEB:  *#13well (.) i don't know if you can see:
          *holds the phone in front of webcam-->
    fig
```

Line 8. #13: DEB holds the phone in front of the webcam.

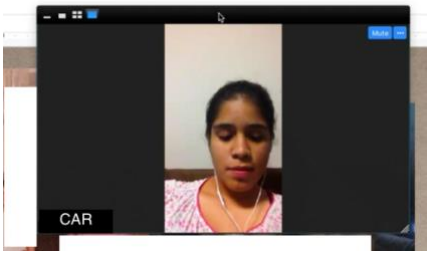
```
09        (1.3)
10  TEA:  >we can< see
11        (1.2)
12  TEA:  a:h [let me=
13  DEB:      [well (.) this is a (.)a day&#14 of fathers' day
    tea                                  &changes to speaker view-->>
14  TEA:  Yeah
```

fig



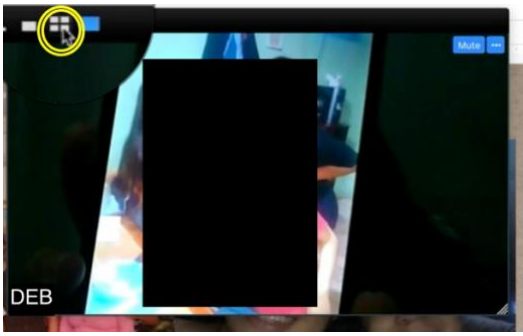
Line 14. #14: The teacher changes to speaker view.

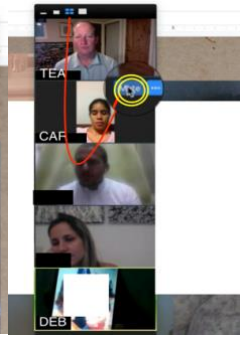
15 DEB: you can see- (.)
16 this is my fa:the:r (.) a:h%#15
%car image on-->
17 %my ((inaudible)) brother (.) with the: red shirt (.)
-->%



Line 16. #15: CAR image appears following background noise.

18 and my big sister* (.)% that's me (.)
*barking dog noise-->
%car image on-->
19 a:nd this little*% is my nephew
-->*
-->%
20 (4.0)
21 DEB: so:*% (.) at this day (.) was*%really (.) &#16cool because
----barking dog noise----
%------car image on-------%
tea -->& back to gallery





Line 23. #17: The teacher deactivates CAR's audio.

Line 21. #16: the teacher changes to gallery view.

22 it was fathers' day (.)
23 and this#17 day (.) we can pla- we play a lot of
tea #deactivates car audio-->>
24 the things (.) and we play (.) games (.) sweets
25 enjoy a lot of&(.) things
tea &changes to speaker view-->>
26 this is one of the pictures
27 TEA: ↑yeah
28 DEB: a:nd

In lines 1-2, the teacher addresses a question to DEB enquiring about her ability to show pictures from her phone by placing it in front of the webcam, and thus, assigning her as the next speaker. In the middle of the teacher's question, DEB gazes down at her phone. A 1.5-second gap follows the teacher's question before DEB responds positively in line 3, which

projects a showing action is relevant next. At the end of DEB's response, a barking dog noise occurs from CAR's background. A 2.4-second pause follows the response, during which DEB seems to prepare a picture to show from her phone. DEB accounts for this silence by asking for more time ('give me one <second>•', line 6), before another relatively long silence follows in line 7. During this 4.7-second silence, DEB appears to be talking to someone in her physical space but is not heard in class. Also, at the beginning of this silence, the barking dog noise stops. DEB launches the showing activity by holding the phone in front of the webcam in line 8 and asks if her fellow participants' can see the pictures on the phone's screen (#13). Such attendance to the frailty of interaction, where participants check if they can be seen and/or heard, seems common in such settings (Fornel, 1994). The fellow participants lean towards the screen following this question.

A 1.3-second pause occurs in line 9, before the teacher answers that they can see the pictures in line 10. A 1.2-second pause occurs while DEB holds the phone in front of the webcam before commenting on the pictures. DEB's initiation of the commentary overlaps with the teacher's accounting for his screen activity to change his Zoom layout to speaker view in line 12 (#14). The teacher then produces a continuer ('yeah', line 14). DEB initiates a multi-TCU turn in lines 15-26, where she comments on the family pictures taken on Father's Day (#15). During this long turn, a fleeting dog bark noise can be heard in CAR's background. The fleeting nature of this noise is highlighted by the speaker view feature on the teacher's screen and possibly on the other participants' screens as well, since it prioritises the video of the person with an active (current speaker) or louder audio at a particular moment. In line 21, the teacher switches to gallery view (#16) and deactivates CAR's audio in line 23 (#17). The teacher then switches back to speaker view and produces another continuer, enabling DEB to hold the conversational for a longer period.

The teacher's continuers are indications of engagement and are also indications that there is no breakdown in the conversation. The other participants do not seem to orient to a moment of trouble up to the point of the audio deactivation of CAR. This can be seen in their responses afterwards on DEB's pictures (not in the transcript). Thus, deactivating CAR's audio here has the potential to function as a preemptive move to anticipated trouble in visibility, should these noises become more persistent, as it can block DEB's video and hinder the ongoing showing activity.

7.4 Summary

This chapter has demonstrated the participants' use of the audio deactivation feature to repair hearing/speaking troubles caused by a known trouble source producer. The analysis in Section 7.1 showed that when the trouble source producer is known to the participants, the repair sequence is shorter and smoother, which results in less hindrance to the progression of the ongoing activity and a quicker resumption of the suspended activity. It also demonstrated the preference for self-repair by examining the repair space always made available before carrying out other-repair. Section 7.1 also showed that teachers account or do not account for deactivating the audio of the trouble producer. Section 7.2 offered examples of the teachers' use of the audio deactivation feature as a preemptive procedure ahead of potential trouble that can result from the interference of a participant's voice or background noises. The next chapter discusses the findings presented in the analysis chapters in light of the wider literature in the field.

Chapter 8. Discussion

8.1 Introduction

The current study examined the participants' strategic use of the audio activation/deactivation features of the Zoom videoconferencing platform to manage turn-taking and repair trouble caused by the interference of background noises in small groups, online, synchronous, L2 speaking classes. The current study used a multimodal CA methodology to conduct this examination, which provided a moment-by-moment multimodal analysis of the participants' use of these features and how this enabled or constrained their management of video-mediated classroom interaction.

The analysis chapters in the current study presented an examination of a number of points. Firstly, Chapter 5 showed how the 'on-mute' learners (OMLs) used the audio activation/deactivation features to organise their turn-taking. Secondly, Chapter 6 presented an analysis of how the teacher handled the interference of background noises using the audio activation/deactivation features to locate the trouble source producer. Thirdly, Chapter 7 offered an analysis of the teachers' use of the audio activation/deactivation features to handle the interference of background noises from a known trouble source producer. In addition, Chapter 7 shows how the teacher used these features as a pre-emptive move to prevent a potential interruption by background noises.

The main findings that emerged from the analysis are related to: (1) the relationship between the participants' use of the audio activation/deactivation features and the organisation of turn-taking; and (2) the relationship between the participants' use of these features and the organisation of repair in the online, synchronous, video-mediated L2 classroom interaction. These main findings will be interpreted and discussed in light of the existing literature in this chapter.

The current study draws on calls for more work to account for the participants' use of the medium's affordances (zoom in our case) for the organisation of mediated interaction (Arminen et al., 2016). In addition, the current study adds to the knowledge of the relationship between technology and the organisation of social interaction (Hutchby, 2001a, 2003b, 2014) as a detailed, moment-by-moment description of the teachers' and learners' use of the technology features in managing their interaction in video-mediated L2 speaking classrooms.

Moreover, the current study sheds light on the local management of turn-taking in audio/video-mediated L2 learning and teaching settings in general (Badem-Korkmaz & Balaman, 2022; Guo & Zhang, 2021; Jenks, 2014; Jenks, 2009; Malabarba et al., 2022; Stone & Brinham, 2022); and demonstrates how the participant' use of the technology features can play a role in this management. In addition, this study also sheds light on the organisation of repair in online audio/video-mediated L2 learning and teaching settings (Alzaidi, 2016; Brandt & Jenks, 2013; Cheung, 2021; Jenks, 2014; Rusk & Pörn, 2019), and the role of using the technology features in facilitating the management of trouble and maintaining/restoring progressivity of classroom activities.

This chapter is organised as follows: Section 8.2 summarises the findings of the study following a similar order to the analysis chapters. Next, Section 8.3 discusses the use of audio activation/deactivation by the OMLs in the management of turn-taking in light of the existing literature. Section 8.4 presents a discussion of the participants' use of the audio activation/deactivation features to manage trouble caused by the interference of background noises. This is then followed by a discussion of how these features were used beyond their intended design by the software developers, which is presented in Section 8.5. Finally, Section 8.6 provides a discussion of the online classroom interactional competencies that teachers and learners need to acquire for the successful management of interaction in online synchronous video-mediated L2 teaching and learning settings.

8.2 Summary of the Findings

This section summarises the findings that emerged from the analysis in this study. The section will be divided into three sub-sections following the order of the analysis chapters. Each sub-section will consider one of the analysis chapters.

8.2.1 Using the Audio Activation/Deactivation Features to manage Turn-taking

The first analysis chapter examined the learners' use of Zoom's audio activation/deactivation features to manage their turn-taking. In the classes recorded, some learners kept their microphones off either during some parts or the whole class. They only activated their audio to participate in an ongoing class activity. Following their participation, these learners deactivated their audio again. To distinguish these learners from the others, they were labelled as OMLs. This labelling was not to impose a category on these learners, but for the sole purpose of easing the readability of the analysis. The analysis showed that there is a reflexive relationship between using these features by the OMLs and the organisation of turn-taking

and repair in these classes. In what follows, the findings from each section in the first analytic chapter will be summarised.

Section 5.2 focuses on how the OMLs' use the features of audio activation projected the possible completion of the current speaker's turn and indicated that they are the potential next speakers. The OMLs in these sequences are not the primary speakers. Hence, the analysis in this section focused on tracking the position at which they activate their audio by examining: (1) what talk preceded the audio activation by the OMLs; (2) the embodied displays of engagement by these learners; (3) the audio activation by the OML; and finally, (4) what follows the audio activation by these learners. Following this, Sections 5.2.1 and 5.2.2 are organised in terms of the positions at which the OMLs activate their audio and self-select for the next turn. In both positions, the audio activation can be seen as a pre-beginning activity (Schegloff, 1996). A difference, however, can be drawn in terms of how early the OMLs activate their audio in relation to the turn in progress. The findings relating to the two positions identified in the data will be summarised next. Figure 8.1 below shows a post-analysis visualisation of these two positions.

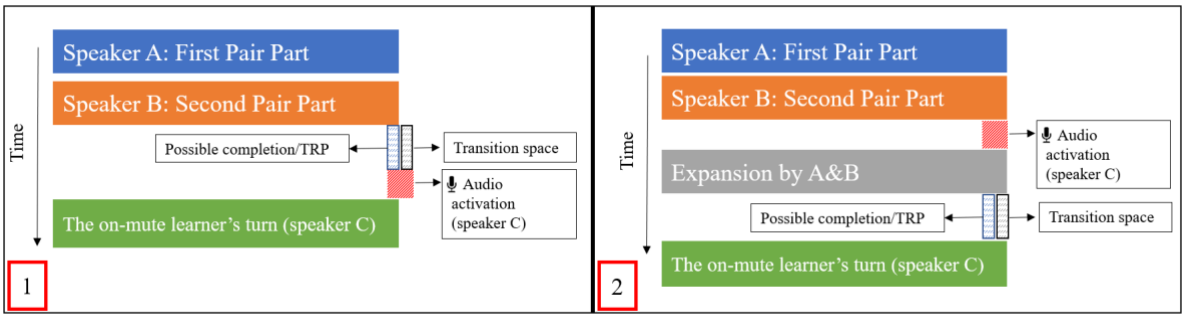


Figure 8.1: The Two Positions for Audio Activation by the OMLs

The first identified position at which the OMLs activate the audio to take the next turn is at the possible completion of the current speaker's turn (i.e. at the TRP). As Number 1 in Figure 8.1 demonstrates, the sequences analysed in this section consist of an FPP, SPP followed by the OML's initiated turn As shown Extract 5.01, for instance. The audio activation by the OML is well coordinated with the other participants' interaction and finely tuned with the possible completion of the current speaker's turn. Moreover, Section 5.2.2 presented an analysis of three extracts showing an earlier position for audio activation by the OMLs in relation to the turn in progress (see Number 2 in Figure 8.1). The turn initiation by the OMLs did not immediately follow the audio activation as in the position mentioned above in Section 5.2.1. Instead, the OML initiated his/her turn at the next possible TRP. This turn initiation's delay by the OML was due to expansions by either Speaker A, Speaker B or both. The OMLs

in the extracts analysed in this section activated their audio at a possible TRP. However, due to expansions by the current speakers, the turn launching was delayed to the next possible TRP. In both positions, the OMLs displayed engagement and close monitoring of the ongoing interaction using a number of resources, such as fixed or shifting gaze at the screen, smiles, or nodding with varying speeds (i.e. slow vs fast).

Extracts Section 5.2.1 show that the teachers recognised the audio activation by the OML as projecting their incipient speakership. This recognition can be seen in what the teachers did following the audio activation as evident in Extracts 5.01 and 5.03. The occurrence of background noises following the audio activation by the OML made it more salient and recognised by the teacher as shown in Extract 5.05, instance. It shows JOL's audio activation and delayed the turn-initiation until the next possible TRP. During this time, a TV noise was heard from JOL's background. The teacher held the mouse cursor over JOL's audio deactivation icon for a second, but did not click it. JOL then launched her turn, followed by the teacher moving the mouse cursor away from JOL's audio deactivation icon. Holding the mouse cursor over the audio deactivation icon for a while without clicking it can be seen as a recognition of the audio activation by JOL as projecting her incipient speakership.

Another observation presented in this chapter is relevant to the criticality of the timing of the audio activation by the OMLs in relation to the success of their self-selection for the next turn (Section 5.2.3). Activating the audio is an additional task that the OMLs have to undertake before launching the turn. Therefore, the mistiming of activating their audio can result in missing the transition space, and thus, losing the opportunity to participate in the ongoing class activity for the OMLs. It is observed in the data that following this, the learners deactivate their audio and return to their on-mute status, as shown in extracts 5.07 and 5.08.

The objective of the analysis in Section 5.3 was to investigate the positions at which the OMLs deactivate their audio after activating it to participate in an ongoing class activity. Figure 0.2 below shows the two positions that were identified in the data: deactivating audio at the transition space (number 1); or delaying it until the OMLs closed an extended sequence (number 2) as presented by the extracts in Section 5.3.1.

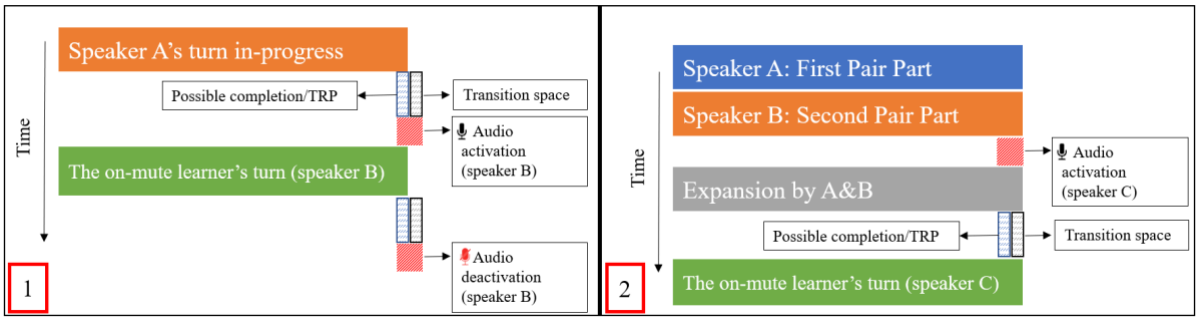


Figure 0.2: The Two Positions for Audio Deactivation by the OMLs

The second identified position for the audio deactivation by the OMLs was delayed beyond the completion of the turn following the audio activation (number 2 in Figure 0.2). The OMLs delayed the audio deactivation to accommodate an expansion to their participation in an ongoing class activity. Following the completion of the expansion, these learners deactivated their audio as demonstrated by the extracts in Section 5.3.2. Another observation is related to the consequentiality of the absence or delay of audio activation by the OMLs to the ongoing interaction. As the extracts in Section 5.4 show, the absence or delay of audio activation by these learners was noticeable and accountable by the other participants in the class.

All in all, the findings in Chapter 5 emerged from the analysis of the sequential positions of the OMLs' use of audio activation/deactivation features of Zoom, which projects their incipient speakership and shows their projection of the current speaker's turn possible completion. The analysis demonstrated the participants' recognition of the audio activation by the OMLs as projecting incipient speakership. Additionally, the OMLs marked the completion of their participation in the ongoing class activity by deactivating their audio. The analysis in this chapter also showed the consequentiality of mistiming, delay or absence of the audio activation by the OMLs.

The absence of audio deactivation by the learners upon the occurrences of background noises is also accountable by the other participants in the classroom, which results in initiating other-repair sequences. In handling these occurrences, the teachers and learners used the audio activation/deactivation features. The findings regarding this use are summarised in the following two sections.

8.2.2 Using the Audio Activation/Deactivation to Manage Repair: Unknown Noise Source

As the participants in the current study were geographically dispersed, the identification of the participant with background noise can be challenging. Chapter 6 presented an analysis of three occurrences of background noises and showed how the teacher used the audio

activation/deactivation features of Zoom to locate the learner feeding this noise. The analysis in this chapter divided each occurrence into three parts: (1) the emergence of the background noise; (2) searching for and locating the trouble source producer; and (3) resuming the progressivity of the suspended class activity.

As Number 1 in Figure 8 **0.3** below shows, the participants showed a preference for progressivity of the ongoing class activity as they did not immediately orient to background noise as disruptive. It seems that the persistence of a certain noise played a role in orienting to it as a trouble source, and thus, the participants needed to halt the progressivity of the ongoing activity to address the interference of the noise. Following the emergence of the trouble source, there was always a space for the participant having the background noise to deactivate his/her audio. If this window for self-audio deactivation was not exploited, a verbal repair initiation occurred by the teacher or in some cases by a learner as evident in Extract 6.02. The other-repair initiation made relevant next an audio deactivation by the trouble source producer, which did not occur in the cases in this chapter.

Interacting in such fractured ecologies (Luff et al., 2003, 2016) made it difficult for the participants in video-mediated multi-party interaction to identify the person feeding the background noise into the class shared space. This resulted in constraining the teacher's ability to handle the trouble and restore the progressivity of the class activity. To search for the trouble source producer, the teacher used the audio activation/deactivation features of Zoom. This was done by deactivating the audio of one learner at a time and holding for a couple of a second before reactivating it to check if the noise reappears (see Extract 6.01.03). This successive use of audio activation/deactivation features was in some cases accompanied by the teacher's confirmation checks, and collaborative work by the learners to identify the trouble source producer. The repair sequences were, thus, extended until the participant

feeding the background noise was identified or the noise disappears on its own.

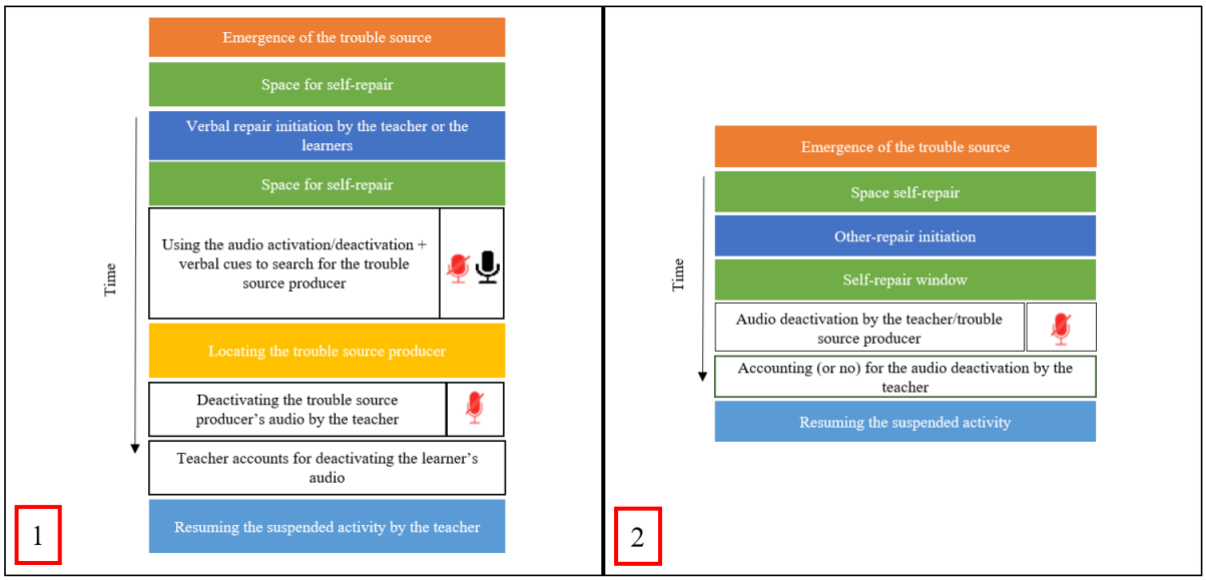


Figure 8 0.3: Trajectories of the Repair Sequences Shown in Chapters 6 and 7.

Figure 8 **0.3** demonstrates the trajectories of the repair sequences analysed in Chapters 6 and Chapter 7. The identification of the trouble source producer was followed by deactivating his/her audio. The teacher verbally accounted for his decision to deactivate the learner's audio. This accounting was mitigated with laughter (Glenn & Holt, 2013; Sert & Jacknick, 2015; Warner-Garcia, 2014) and its design orients to deactivating a learner's audio as a dispreferred action. Having successfully located the trouble source producer, the suspended class activity was resumed by the teacher. The use of the audio activation/deactivation features enabled the teacher to locate the source of the noise and restore the progressivity of the suspended class activity. The following section summarises the findings of examining the teachers' and learners' use of these audio activation/deactivation features to handle speaking or hearing troubles caused by the interference of background noises of a known learner.

8.2.3 Using the Audio Activation/Deactivation Features to manage Repair: Known Noise Source

This section summarises the findings that emerged from the analysis in Chapter 7. Two main findings emerged. Firstly, the teacher's epistemic access to the learner with background noises and the use of the audio deactivation features result in shorter repair sequences and a much quicker resumption of the progressivity of the class activity. Secondly, the teacher used the audio activation feature as a preemptive move (Schegloff, 1986, p. 133) to maintain the progressivity of the ongoing class activity.

The repair trajectory analysed in Section 7.1 is visualised in Number 2 in Figure 8 0.3 above. Similar to the trajectory shown in Number 1, the participants did not immediately treat an occurrence of background noises as a trouble source. There was a space for self-audio deactivation following the emergence of such noises, which is followed by orienting to this noise as a trouble source by halting the progressivity of the ongoing activity, as in Extract 7.01, for example. Another space for self-audio deactivation also followed such orientations. However, this did not occur in the extracts analysed in the chapter. The absence of self-audio deactivation resulted in the teacher's deactivation of the trouble source producer. Similar to the extracts in Chapter 6, teachers treated such audio deactivation as a dispreferred action by accounting for it. This is done by apologising and explaining the reason behind it, or only by justifying the decision. However, in other cases, there was no accounting for such decisions. Following the repair of the interference from the learner's background noise, the progressivity of the class activity was resumed.

Another observation was that the participants might not topicalise the trouble source; hence, the teacher deactivated the trouble source producer's audio and the current speaker resumed the momentarily suspended talk. This resulted in shorter repair sequences and less hindrance to the progressivity of the classroom interaction. Using the affordances of the medium played a role in enabling such a quick resolution of the trouble. To elaborate, Zoom places a green box around the video frame of the current speaker or the person with louder audio. In the occurrences of background noises, a TV noise emanating from a learner's background may not be louder than the current speaker's voice. However, it can become more audible during the pauses of the current speaker's talk and occupy the conversational floor. The green box feature then afforded the teacher quick access to the location of such disruptive noise and to deactivate that learner's audio.

Another finding presented in the chapter (see Section 7.3) relates to the use of the audio deactivation feature by the teacher to pre-empt the occurrence of potential trouble; thus, it maintains the progressivity of the ongoing class activity. The existence of background noises in the extracts analysed in this section did not cause interactional trouble. However, the teachers deactivated the audio of the learner with noises in their backgrounds, which enabled the class activity to progress without risking a potential interruption by such noises, as seen in the previous sections.

To recap, the analysis in the current study revealed a number of findings relating to the teachers' and learners' use of the audio activation/deactivation features in synchronous,

video-mediated L2 speaking classrooms. Moreover, the analysis shows how the participants' use of the medium's audio activation/deactivation features afforded them the ability to carry out multiple social actions pertaining to the organisation of turn-taking and repair in this setting. In the following section, these findings will be further discussed in relation to the relevant literature.

8.3 Audio Activation/Deactivation Features and Turn-Taking Management

This section discusses the main observations that emerged from the analysis in Chapter 5 in light of the relevant literature. These findings are related to the OMLs' use of audio activation/deactivation features for maintaining the boundaries between their physical and shared class environments, projecting their self-selection for the next turn, and projecting the completion of their participation. In addition, the section will discuss the consequentiality of the absence/delay of the audio activation by the OMLs.

8.3.1 Maintaining Boundaries between Physical and Shared Environments

Before the discussion moves to the sequential positions of audio activation/deactivation by the OMLs and the consequentiality of its absence or delays, a question regarding the reason behind some learners choosing to keep their audio deactivated. The answer to this question can be inferred from collecting the orientations of these learners to sources in their physical spaces that can make noises. For example, the learners in the current study joined the classes from home and during these lessons they mentioned having children, other family members, TVs, or Pets in their immediate environments. As the extracts analysed in Chapters 6 and 7 show, people, animals or objects in the learners' physical spaces made noises which caused a disruption to the class's ongoing activities. Thus, by keeping their audio off and only activating it to participate, the learners are actively maintaining the boundaries between their local environment and the class shared space. According to Fornel (1996), the participants in video calls are located in different physical environments, and should they want to maintain boundaries between their physical environments and common environments, they need to intervene separately. Fornel (1996) states that it is the responsibility of each participant to control the relationship between the two environments. The lack of collective control in this setting renders the interactional zone extremely fragile, as the intrusion of background noises, for example, can threaten it (Fornel, 1996).

Although Zoom communication software is significantly more sophisticated than the one used in Fornel's study, the principle is still the same when it comes to the importance of

maintaining boundaries between physical and public environments (i.e. the virtual space). Fornel (1996) provided an example that showed a participant halting the progressivity of an ongoing conversation to close his/her door because there were noises coming from outside the room. In the current study, the noises or sources of potential noises were in the participants' immediate physical environments. In order to maintain the boundaries between the two environments and "create an 'adequate' interactional frame" (Fornel, 1996, p.49), the OMLs used the audio activation/deactivation features of Zoom. Thus, it can be stated that the use of these Zoom features enabled the OMLs to actively participate in these video-mediated L2 class activities, while simultaneously isolating ongoing noises in their backgrounds (i.e. TV) or potential noises that can disrupt classroom interaction (i.e. children or pets). Having understood the logic behind the choice by some of the learners to keep their audio deactivated for some parts or the whole of class time, the following sub-section discusses how they self-selected and joined an ongoing class activity.

8.3.2 Self-selection using the Audio Activation Features

This sub-section offers a discussion of the OMLs' use of the audio activation feature to project their incipient speakership for the next turn (Section 5.2). The following subsections will discuss how the OMLs: (1) displayed their engagement before they activated their audio; (2) how they used the audio activation feature to project the completion of the current turn and indicate their incipient speakership; and (3) how the absence or delay of audio activation/deactivation can be noticeable and accountable by the other participants.

8.3.2.1 Displays of Engagement before the Audio Activation

This subsection discusses the observations relevant to the talk that preceded the audio activation by the OMLs and how they displayed their engagement and willingness to take the next turn leading up to the audio activation and launching the turn. Previous research revealed that learners displayed their engagement and willingness to participate in in-person classroom interaction using resources, such as: hand raising (Sahlström, 2002); summons (Maroni, 2012); gazing at the teacher (Mortensen, 2008); gaze and positioning (Evnitskaya & Berger, 2017), and body movements (Mortensen, 2009). Such studies showed that these displays of willingness to take the floor next were well-timed with the ongoing class activities. In video-mediated L2 classrooms, learners displayed recipiency during the turn in progress by positioning their bodies and gaze towards the screen (Stone & Brinham, 2022). The analysis in this study shows that the OMLs displayed engagement during the turn in progress prior to the audio activation and self-selection. For instance, Extract 5.01 demonstrates the OML

displaying her engagement by fixing her gaze on the screen, accompanied by smiling and nodding during the dyadic interaction between the teacher and her fellow learner.

In addition, the findings show the OMLs' constant gaze shifting to the right and left side of the screen by the OML before launching their turn (Extract 5.02). These findings align with Stone and Brinham's (2022) observations regarding the displays of engagement and recipiency in video-mediated L2 classrooms. Indeed, such displays of engagement by the OMLs in the data show how the participants fit their interaction into the affordances of the medium (Due & Licoppe, 2020) to overcome the constraint posed by interacting in fractured ecologies (Luff et al., 2003, 2016). Following these displays of engagement, the audio activation by the OMLs occurs; The sequential positions of the OML's audio activation are discussed next.

8.3.2.2 Audio Activation and Projecting Incipient Speakership

The discussion of the OMLs' use of the audio activation feature for projecting the possible completion of the turn in progress and projecting their incipient speakership will be from two levels. First, the audio activation and self-selection by the OML is discussed in relation to its sequential position in the IRF/IRE sequences (Mehan, 1979; Sinclair & Coulthard, 1975). Second, it will be discussed concerning its position in the turn in progress. It serves as a reminder that the extracts in this section demonstrated cases where the OMLs were not the primary speakers, meaning they self-selected and initiated their turns without being allocated by the teacher or a fellow learner.

The previous literature showed that a dominant pattern in classroom interaction is the IRF/IRE sequences. In this pattern, the teacher asks a question, for example, the learner's answer to that question is then followed by the teacher's feedback or evaluation. Some of the extracts that demonstrated the OMLs' use of the audio activation feature to join an ongoing conversation featured this pattern (Section 5.2). According to Waring (2009), a potential opportunity for the learners to self-select is at the completion of successive IRFs. Extracts 5.04 and 5.06 showed a similar position, where the OMLs self-selected and took the next turn. This, however, differs from Waring's study in terms of the self-selecting learner. In Waring's study, the learner who self-selected following a series of IRFs is the learner who was speaking in the last one, thus, benefitting from a preference for the next turn to be initiated by the last speaker. In the current study, the OML was not the last speaker, but managed to anticipate the completion of the IRF and projected themselves as the next speaker by activating their audio followed by initiating their turn at the next or next possible transition space.

Moreover, the current study presented cases that demonstrated the OMLs' self-selection following the completion of a single IRF. For instance, Extracts 5.01 and 5.02 showed that the OMLs activated their audio and launched the next turn immediately after the completion of the IRF sequence in which the teacher and a fellow learner were involved. The IRF sequences in this section consisted of an FFP and an SPP, and a third turn that closed the sequences (Ingram & Elliott, 2014), followed by the OMLs' audio activation and turn-initiation. Extract 5.05, comparatively, showed a slightly different pattern, as the OML self-selected and launched their turn in the third turn, which is usually occupied by the teacher.

Furthermore, Waring (2009) pointed out that the opportunities for learners' self-selection can be enabled by the teachers. This is consistent with what is observed in the extracts in Section 5.2. Following an audio activation by the OML, the teacher does not immediately reclaim the conversational floor. Instead, the teachers facilitated the opportunity for the OMLs' self-selection by increasing the wait time (Walsh, 2012). This is evident, for example, in Extract 5.05, where the teacher held the mouse cursor over JOL's audio deactivation icon and did not click it, despite the TV noise coming from her background.

Furthermore, as McHoul (1978) points out, the teacher has the first access to the turn following the student's answer. This access is observed in the extracts in this section as the teacher takes the third turn following the addressed learner's answer, either to reshape it (Walsh, 2012) as seen in Extract 5.01 or to display an agreement (Extract 5.02). This turn marked the closure of the dyadic interaction, and its completion was a possible point of departure for the OML to exploit in order to initiate their turns following the audio activation.

Having considered the sequential position of audio activation and self-selection by the OMLs in relation to the IRF/IRE sequences, the discussion now moves to its position in relation to the turn in progress. Previous studies have shown that participants in a conversation used a number of non-verbal practices to project self-selecting, such as clearing their throats, gaze, facial expressions, pointing, and head and body movements before the current turn reaches a possible completion (Mondada, 2007; Mortensen, 2009; Schegloff, 1996). The extracts in Section 5.2 showed two positions for audio activation: firstly, at the possible completion of the current turn in progress; and secondly, at an early position to the possible completion to the current turn in progress. This is similar to the positions identified in Mondada's (2007) study of the sequential positions of pointing gestures to project self-selection for the next turn. The use of the pointing gesture is similar to the audio activation in terms of how they are finely-tuned with the turn in progress and how the other participants adjust their interaction

accordingly in recognition of the use of these practices. Also, Mondada's study indicated that the participants successfully used the specificities of the setting (i.e. the table and the materials on it). This use of the setting's specificities is evident in the current study, as the participants managed to use the features of Zoom to project their incipient speakership.

In addition, the position of using the audio activation is consistent with Mortensen's (2009) study of using in-breaths, gaze, and body movement by the learners to claim incipient speakership in in-person classroom. Mortensen (2009) indicates that the students used these resources to claim incipient speakership and secure the other participants' displays of recipiency. Such use, as Mortensen reported, occurred as the teacher's turn was reaching a possible completion. They occurred in a pre-beginning position (i.e. before the verbal initiation of the turn by the self-selecting learner). In the current study, it is evident in the extracts presented in Section 5.2 that the OMLs activated their audio in a similar position.

Although keeping the audio off enabled the OMLs to maintain the boundaries between their local physical spaces and the class shared space, it constrained their ability to use resources such as clearing their throats (Schegloff, 1996) or using turn-entry devices, such as 'well', 'but', 'and', or 'so' (Sacks et al., 1974, p.719). This is because their audio was off, and the other participants in the class would not be able to hear them. In addition, other resources used in in-person settings, such as gaze and body movement (Mortensen, 2009) can lose a great deal of their significance in video-mediated interaction (Luff et al., 2003, 2016).

Similar to projecting self-selection, participants in a conversation can project the possible completion of their turns using different resources. The following offers a discussion of using the audio deactivation feature by the OMLs to project the possible completion of their participation.

8.3.3 Audio Deactivation and Projecting the Completion of Participation

As noted in Section 8.3.1, the OMLs managed to separate between their local environments and the class shared environment by keeping their audio off and only turning it on to participate in the class activity. It is also interesting to examine where the audio activation ends and how the OMLs register their participation as complete. This section discusses the findings related to how the learners' audio deactivation projected the completion of their participation. Previous researchers have shown that participants can project the completion of their turns using a number of verbal and non-verbal resources, such as syntax, prosody, pragmatic, gaze, and others (Bögels & Torreira, 2015; Hayashi, 2004; Levinson, 2012;

Mondada, 2006; Rossano, 2013; Streeck, 1995b). In the current study, the OMLs projected the completion of their participation by deactivating their audio. The analysis showed two positions at which the OMLs deactivated their audio following their participation.

First, the OMLs deactivated their audio at the transition space following the completion of the turn containing the action for which they activated their audio. Extract 5.09, for example, shows that LOR activated her audio to ask a question and deactivated it upon the completion of that question. Similarly, Extract 5.10 shows MAR self-selects to answer the teacher's question and deactivates his audio following the completion of the answer. By so doing, LOR and MAR marked their turns as complete and returned their on-mute status. Also, the audio deactivation here can be seen as a resource for projecting their turn's possible completion and yielding the conversational floor to the next speaker (i.e. to the teacher as in Extract 5.10 or a fellow learner in Extract 5.09).

Furthermore, the OMLs can delay their audio deactivation to accommodate an extension to their participation beyond the completion of the action for which they activated their audio to do. This is evident in the extracts analysed in Section 5.3.2, as the OMLs delayed the audio deactivation beyond the completion of a question-answer adjacency pair to greet a newcomer to the class, as seen in Extract 5.13 or elicit further information from the recipient of their question, as in Extract 5.14. On some occasions in the data, the OMLs seem to delay the deactivation of their audio after answering a question to make themselves available for a follow up. Similarly, they activated their audio to ask a question and delayed the deactivation until after receiving the answer.

The discussion, hitherto, has shown how the learners use the audio activation/deactivation features to project their incipient speakership and to mark the completion of their participation in the class activities. However, it is not always this straight forward as the learners' absence or delay of audio activation/deactivation is noticeable and accountable by fellow participants in the classroom. The following section discusses the observations regarding this consequentiality of the absence or delay of audio activation/deactivation.

8.3.4 Consequentiality of Absence/Delaying the Audio Activation/Deactivation

According to Hutchby (2014), using technology features can play a role in facilitating or constraining the structure of social interaction. In the previous section, it was evident that the audio activation feature enabled the OMLs to isolate the noises in their physical environment from the class shared space while still being able to actively participate in the class activities.

As shown in the analysis (Section 5.3), the mistiming of the audio activation can result in losing the opportunity to participate.

Mistiming of contributions by the participants in audio/video-mediated L2 interaction can be consequential and can result in overlaps (Jenks, 2009). The importance of timing contributions for the L2 learners was emphasised by Jenks (2009), as the learners lacked access to visual cues in audio chats. In this study, despite the participants' access to each other's videos, mistiming of contributions still occurred. This occurrence was due to the extra work the OMLs had to undertake prior to launching their turns (i.e. stretching their arms to reach the mouse or the touch the screen, dragging the mouse cursor, clicking the audio activation icon, or pressing the space bar before activating their audio). Therefore, activating the audio past the transition space resulted in losing the opportunity to participate because another participant had already occupied the conversational floor.

The Mistiming of audio activation can also result from transmission delays. Previous studies have presented evidence of the impact that transmission delays can have on turn-taking in video-mediated interaction (Malabarba et al., 2022; Ruhleder & Jordan, 2001b; Rusk & Pörn, 2019b). Therefore, it might be that the OMLs have activated their audio right at the transition space from their perspective, but they missed it, because they were experiencing transmission delays. It should be noted that the data in this study comes from the teachers' screens; thus, such a claim regarding the OMLs experiencing transmission delays remains speculation.

The absence/delay of the audio activation becomes more noticeable and accountable by fellow participants when the OML is the specified recipient and the nominated next speaker. When the teacher addresses a question to the OML, a response by the OML is relevant next. The absence of what is relevant next interrupts the progressivity of the interaction (Schegloff, 2007) and requires a repair initiation. As such, fellow participants orient to this absence/delay of audio activation and indicate hearing trouble using verbal and non-verbal resources (Extracts 5.15 and 5.16). The learners are expected to activate their audio to be able to participate in the class activity verbally, and not doing this is accountable by reference to the norm (Seedhouse, 2004). The norm in this case is to activate the audio before initiating a verbal turn. Extracts 5.15 and 5.16 showed that the participants treated the OMLs' absence/delay of audio activation as problematic. Thus, the participants initiated other-repair to alert the OML that their audio was not active. Following the repair initiation, the OMLs activated their audio and accounted for not activating it, as shown in Extract 5.16. Alternatively, the OML activated their audio and resumed the progressivity of the suspended

activity with accounting for the absence/delay of the audio activation. Having discussed the findings from Chapter 5 in light of the relevant literature, the discussion moves to the findings shown in Chapter 6 and Chapter 7.

8.4 Managing Background Noises Using Audio Activation/Deactivation Features

Chapters 6 and 7 presented analyses of a number of extracts in which the participants experienced interference from background noises and showed how they handled them. Previous researchers have shown the consequentiality of the interference of background noises for the ongoing interaction in audio/video-mediated interaction (Alzaidi, 2016; Cheung, 2021; Fornel, 1996; Jenks, 2014). For instance, Jenks (2014) contended that background noises could result in halting the progressivity in the ongoing interaction in voiced-based L2 chats. The findings in both Chapters 6 and 7 align with this observation and show that such background noises can cause trouble in interaction in video-mediated L2 learning and teaching settings. In addition, Jenks (2014) stated that the occurrences of background noises are not uncommon in such environments, and the participants can make them interactionally relevant by orienting to them. This is also the case for the setting examined in this study. The following sections offer a discussion of the main observations in Chapters 6 and 7.

8.4.1 Orienting to Background Noises as Disruptive

A question to ask at this point is relevant to what characterised the background noises treated as disruptive by the participants in this study. The analysis shows that participants did not orient to the background noises as disruptive immediately following their occurrence. A common characteristic of the background noises that the participants oriented to as disruptive was their persistence. Fleeting noises can occur, as the participants in this study joined the classes mostly from home, where many potential sources of background noises exist, such as TV, pets, and family members. However, persisting noises, such as a purring parrot and dog barking (Extract 6.01), people chatting in the background (Extract 6.02), TV noises (Extract 6.03, 7.02), or indefinite noises (Extract 7.01) were made interactionally relevant (Jenks, 2014), as the teachers halted the progressivity of the ongoing class activity to deactivate the trouble source producer (i.e. the learner having the background noise). As illustrated in Figure 8 **0.3**, the orientation to the background noises as disruptive did not immediately follow their occurrences. Instead, the participants maintained their talk flow for some time. This brings the discussion to the participants' orientations to preference for progressivity and self-repair. In what follows, these two types of preferences will be discussed.

Previous researchers indicated that there is a preference for the progressivity of talk. According to Schegloff (2007), progressivity refers to "moving from some element to a hearably-next-one with nothing intervening" (p.15). In the current study, the preference for progressivity was evident in the delay of orienting to the background noises as disruptive. For instance, Extract 7.01 shows the teacher's continuation of instruction-giving and trying to talk over the ongoing background noises. The continuation of talk over the ongoing noises at the beginning of their emergence is also evident in the other extracts in Chapters 6 and 7. In the case of fleeting noises, talking over the noise helps to maintain the progression of the activity. However, as noted above, the persistence of the noises can result in the momentary suspension of the class activity.

Moreover, the participants' orientation to the preference for progressivity manifested in the multiple repair initiations and the collaborative work by both the teachers and the learners to locate the trouble source producer. According to Hırçın Çoban and Sert (2020), progressivity entails "the resolution of interactional trouble and producing subsequent talk" (p.68), and the participants' engagement in repair sequences is an orientation to their preference for progressivity (ibid). Furthermore, in this study, not only the teachers initiated the repair sequences, but in some cases, the learners oriented to the background noises as disruptive and initiated repair sequences before the teacher did. Extract 6.02 provided an example of an other-repair sequence initiated by one of the learners following the persistence of disruptive noise. Following this repair initiation, collaborative work by the teacher and learners ensued in search of the noise source. When the noise disappeared on its own, the progressivity of the suspended activity was resumed by the teacher.

Moreover, the preference for progressivity was oriented to by the teachers as they skipped accounting for the audio deactivation decision. For instance, in the early instances of background noises occurrences, the teacher followed their audio deactivation decision with apologies or justifications, as seen in Extracts 6.01,7.01, and 7.02. However, in the latter occurrences, the teachers' skipped this accounting, which enabled a quicker resolution of the interference of the disruptive noises, as evident in Extracts 7.03, 7.04, 7.06, and 7.07.

The quicker resolution of the interference of disruptive noises to maintain/restore the progression of the suspended activity was enabled by the use of the audio deactivation feature of Zoom. For example, the teachers' use of the audio deactivation features as a pre-emptive move (Schegloff, 1986), as shown in Section 7.3, helped maintain the progression of the class activity. This use of audio deactivation is, of course, aided by the participants' epistemic

access to the identity of the learner having what is deemed to be disruptive noises. The green box surrounding the learner who was having the background noise also enabled a quicker identification of them. Therefore, it also enabled the teacher to maintain the progressivity of the class activity by pre-empting a potential interruption by disruptive noises.

8.4.2 Preference for Self-Repair

In the current study, the other-repair initiated sequences were prevalent in the Extracts analysed in Chapters 6 and 7. This observation aligns with the observations made by Rintel (2013, 2015) regarding repair in video-mediated interaction. According to Rintel, the participants may not be aware of being the trouble source producers, as they lack access to how they are being seen or heard; hence, the other-repair sequences were prevalent in his study. In the current study, the learners had TVs on, children and pets next to them while attending the class. Extract 6.03, for instance, shows that NAD was not aware that the noise that the teacher treated as disruptive was coming from her TV. Within these other-initiated repair sequences, the participants oriented to their preference for self-repair. According to Sidnell (2010), the preference for self-repair is not a matter of likes or dislikes, but it is seen in the positions of the initiation of repair by self or other speaker(s) in relation to the turn containing the repairable. Additionally, Liddicoat (2022) states that the trouble source producer structurally has the first opportunity to initiate repair. However, as aforementioned, the participants might lack epistemic access to what is being heard by fellow participants. Additionally, they might be aware of the noises, but they are experiencing internet connection troubles at that particular moment. This, in a way, offers the other participants the first access to initiate repair (Rintel, 2013). Nevertheless, the participants oriented to their preference for self-repair, as will be shown below.

The participants' orientations to the preference for self-repair can be seen in the delay of other-repair initiation following the emergence of the background noise. In both the minimal and non-minimal repair sequences in this study, there was always a space for the learner with the background noise to self-deactivate their audio (see Figure 8 **0.3**). As noted, the trouble source producer structurally has the first opportunity to initiate repair (Liddicoat, 2022). Also, in the current study, other-repair is always initiated by fellow participants past this first window for self-repair. For example, Extract 7.02 shows that the teacher continued with the class activity following the emergence of a TV noise. The absence of self-repair by the trouble source producer triggered the other-repair.

Moreover, another space for self-repair is offered to the trouble source producer following the verbal other-repair initiation by fellow participants. For instance, Extract 6.01 shows that the teacher offered CAR multiple self-audio deactivation opportunities before he finally deactivated her audio. A similar observation is illustrated by Extract 6.02, as the participants showed a preference for self-repair. However, the absence of self-audio deactivation led to multiple audio activation/deactivation trails to identify the trouble source producer. According to Liddicoat (2022), suggesting that self-repair initiation is preferred makes the other-initiation of repair dispreferred and, as such, they are usually produced in a mitigated manner, such as in a question form. The orientation to the dispreference of other-repair initiation is evident in the design of the verbal repair-initiation in Extracts 6.01, 6.02 and 6.03, for instance. The participants designed their verbal other-repair initiations as questions ("is anyone else getting that sound↑ (.)where is that coming from", Extract 6.02.01) and mitigated them with laughter ("$↑skyla what're you ↑watching$", Extract 6.01.02).

Previous research shows that teachers flagged the transgressiveness of certain behaviour in an indirect manner (Hazel & Mortensen, 2017; Klattenberg, 2021; Margutti, 2011). This is of course with the assumption that there is mutual epistemic access to the transgressive behaviour between the teacher and the learners (ibid). In the current study, the participants also verbally flagged the background noises as transgressive in a similar manner, as shown in Chapter 6. Nevertheless, the turns following this flagging showed that no mutual epistemic access was in place. The indirect verbal flagging of the noises as disruptive can be considered as an orientation to the dispreference of deactivating others' audio, especially when the teacher is unsure about who is feeding the noise.

Furthermore, the dispreference of deactivating the other participants' audio is also evident in the teachers accounting for the audio deactivation decisions. They accounted for deactivating the trouble source producer's audio either by justification ("↑SEL >i i muted< you ↑SEL because it's too loud the: (1.0) your sound (.) oka^y↑^", Extract 7.02) or by apologies and justifications ("sorry julia(1.0) julia i muted you because: (.) ahm (.) i (.) a- i feel like ahm: (.)>it might be< distracting↓", Extract 7.01), for instance.

The following section discusses how the participants managed the responsibility for audio deactivation following the occurrence of background noises.

8.4.3 Managing Responsibility for the Audio Deactivation

The participants in the current study held each other accountable for the noises coming from their backgrounds, even if they had not personally produced them. For instance, CAR was held accountable for the noises produced by her pets, as seen in Extract 6.01. This is opposed to holding NAD accountable for noises coming from her TV in Extract 6.03 or SEL, as shown in Extract 7.02, for the same reason. This shows that interactants in such fractured ecologies (Luff et al., 2003) are responsible for maintaining the boundaries between their local physical environment and the class-shared space (Fornel, 1996).

The way the participants oriented to the interference of background noises is similar to the teachers' turn design to orient to whom a transgressive behaviour belongs, which made the cessation of it relevant next (Hazel & Mortensen, 2017; Klattenberg, 2021). In this study, the participants talked over the persistent disruptive noises before halting the progression of the activity to initiate other-repair. In Extract 6.01, for instance, the teacher addresses a question to Skyla ("$↑skyla what're you ↑watching$", line 23), instead of deactivating her audio. According to Robinson (2006), some practices for initiating other-repair indicate whom the trouble source belongs to; thus, a self-repair is relevant next. In Klattenberg (2021), such orientations to the transgressive behaviour by the teachers led to its cessation next. In the current study, this is also the case, as shown in Extract 6.03, with NAD leaving her seat to turn off the TV following the teacher's orienting to its noise as disruptive. Thus, the participants in this study have shown that it is the responsibility of the trouble source producer to deactivate their audio upon the occurrence of background noises in their physical environments, regardless of their personal involvement in producing it.

So far, the discussion has shown how the teachers and learners used the audio activation/deactivation features to manage their turn-taking and repair in video-mediated L2 classrooms. The following section elaborates on how the participants created novel practices by using these features beyond their intended design.

8.5 Audio Activation/Deactivation as Novel Practices

In interacting in fractured ecologies, the participants can create novel practices using the medium's affordances to overcome the different constraints (Due & Licoppe, 2021; Hutchby, 2001). Moreover, Hutchby (2001) noted that the participants can use technology features beyond the intended use by the designers. In the current study, a number of novel uses of these two features were observed. Firstly, Section 8.3 discussed how the OMLs' use of the

audio activation/deactivation features enabled them to project their incipient speakership and mark the completion of their participation. Also, as noted in Section 8.3.1, the use of these features enabled the OMLs to maintain boundaries between their physical spaces and the class shared space and strategically use these features to actively participate in the class activities. According to Andrews (2020), being on-mute can deter the students from participation in video-mediated classes, as it is challenging for them to know when to join an ongoing conversation without risking overlapping with the current speaker.

Secondly, the teacher used the audio activation/deactivation features to search for the learner who had the disruptive noises. In Extracts 6.01, 6.02, and 6.03, the trouble source producer was unknown to both the teacher and the learners. As previously noted in Section 2.2, the participants offered a space for self-audio deactivation both before and after the verbal other-repair initiation. The absence of the self-audio deactivation by the trouble source producer, and the persistence of the noise increased the length of the class activity's suspension. Therefore, the participants needed to locate the whereabouts of the disruptive noise. The teacher is the only participant with access to activate/deactivate the audio of the others. Thus, he used them in the search for the trouble source producers by activating/deactivating the learners' audio (i.e. one at a time). This use of the audio deactivation/activation to search for the trouble source producers enabled the teacher to overcome a difficulty observed in the repair in video-mediated contexts. According to Ruhleder and Jordan (2001b), participants in video-mediated interaction may face difficulties in locating the trouble source, and they may be unable to repair the trouble, as "its origin is obscured" (p.132). This might be different from in-person classroom context as the participants are located in the same physical space and have access to what is being heard or seen by fellow interactants.

Thirdly, the teachers used the audio deactivation feature as a stronger repair device. In other-repair initiations, interactants move from weaker to stronger repair devices and the strength of the repair practice is tied to its precision in locating the trouble source (Pomerantz, 1984; Svennevig, 2008). During the occurrence of multiple other-repair initiations, the audio deactivation feature was used as a stronger repair device. This use is evident in the Extracts analysed in Chapter 6, which show that the participants used verbal other-repair initiation before using the audio activation/deactivation features. Also, the extracts in Chapter 7 show that the use of audio deactivation came after verbal indications of trouble, such as relatively long pauses and hesitation markers, as seen in Extracts 7.01, 7.02 and 7.03.

Fourthly, the teachers used the audio deactivation features as a pre-emptive move to potential trouble (Schegloff, 1986). Such use as a pre-emptive move, in turn, enabled the current speaker to maintain the conversational floor and secure access to more speaking time in the target language. This use of the audio deactivation feature is evident in the extracts analysed in Section 7.3.

As the analysis showed, the participants' use of the audio activation/deactivation has changed over the course of these lessons. The following section sheds light on the longitudinal development of using these features by both teachers and learners.

8.6 The Longitudinal Development of Using The Audio Activation/Deactivation Features in Interaction Management

Previous CA research has provided important insights into the examination of the longitudinal development of the participants' practices in L2 video-mediated learning and teaching settings (Balaman, 2016; Balaman & Doehler, 2022; Balaman & Sert, 2017a; Pekarek Doehler, 2021; Sert, 2017; Sert & Balaman, 2018). Similarly, the analysis in the current study showed how both the teachers' and learners' use of the audio activation/deactivation features developed over time. The longitudinal development is evident in the management of both turn-taking and repair using these features by the participants. In the early classes, the appearance of disruptive noises resulted in major holds to the progressivity of class ongoing activities. Moreover, the teachers' followed their deactivation of the learner's (trouble source producer) audio by accounting for such a decision and apologising (as in Extract 6.01, for instance). However, as the extracts in Chapter 7 showed, the repair sequences in the later classes are noticeably shorter and smoother. This is due to the teachers' routinisation of the audio deactivation as a solution to background noise interference. The routinisation of such a solution is manifested in skipping the accounting and apologising following audio deactivation by the teachers. Progressivity of the class activity was then restored in a quicker and smoother manner. Moreover, such routinisation is evident in the teachers' use of the audio deactivation feature to pre-empt potential trouble (see Section 7.3).

The longitudinal development in using the audio activation/deactivation features in managing interaction is not exclusive to the teachers. It is also evident in the learners' use of such features, as the extracts in Chapter 5 showed with the emergence of the OMLs and how these features are used to access an ongoing class activity. In the earlier classes in the data, the learners' background noises caused disruptions to the progression of the class activities. Therefore, the teachers handled such trouble by deactivating the learner's audio, as shown in

Chapters 6 and 7. However, the analysis in Chapter 5 shows that in the later classes, learners with potential sources of disruptive noises (i.e., pets, TV, children) kept their audio off and only activated it to access an ongoing conversation before deactivating it again. Such use of the audio activation/deactivation features evidences the learners' development in using the technology features over time to manage their interaction. As mentioned earlier, this developed use of the audio activation/deactivation features afforded the OMLs the possibility to isolate their background noises while still being able to actively participate in class activities.

The ability to use the technology features to overcome trouble in online synchronous classrooms and create more opportunities for participation is seen as an important competence that needs to be acquired by the participants. The following section discusses the importance of the online classroom interactional competencies and their role in maximising participation and learning opportunities for learners.

8.7 Online Classroom Interactional Competencies

Previous research in online language learning and teaching has emphasised the importance of online teaching and learning competencies (Grammens et al., 2022; Hampel & Stickler, 2005, 2015; Moorhouse et al., 2021, 2022). Relevant to the current study is the emphasis placed on the importance of acquiring the necessary classroom interactional competencies CIC (Walsh, 2011) and e-classroom interactional competencies (Moorhouse et al., 2021, 2022) for teachers. In their suggested e-CIC framework, Moorhouse et al. (2022) showed that technological competencies encompassed all the other competencies (see Section 2.5) since the lessons are mediated by technology. Therefore, the technological competencies are necessary for teachers to acquire in order to successfully manage online language learning classrooms.

In the current study, the teacher's technological competence played a role in overcoming the interactional trouble caused by the interference of background noises. In the early classes recorded in this study, both the teacher and the learners were new to small group classes on Zoom. In the extracts analysed in Chapter 6, the teacher spent a relatively long time to identify the learners feeding the background noises, despite the green box appearing around their video frames. Although the teacher and the learners successfully managed to locate the trouble source producers, a more developed technological competence could have contributed to a faster resolution of the trouble. The teacher's technological competence development can be seen in the latter recordings, as there was a quicker resolution of the trouble caused by the

background noises' interference; therefore, less hindrance to the progression of class activity as shown in Chapter 7.

However, this study argues that it is not only instrumental to acquire such technological competencies for the teachers, but it is also for the learners. For example, Chapters 6 and 7 showed that the participants' underdeveloped technological competence resulted in longer hindrances to the interaction in these classes. The learners were new to Zoom and did not have sufficient knowledge regarding using the different features. The teacher dedicated the beginning of the first classes to teaching the learners how to use Zoom's different features. Comparatively, the latter recordings, as shown in Chapter 5, showed that the learners developed technological competence and demonstrated this, as they successfully used the audio activation/deactivation features to maintain the boundaries between their local physical environments and the class-shared space. As a result, they were able to actively participate in the class activities, while isolating their background noises or potential noises from hindering the class activities. This, in turn, maximised their ability to participate more and secure more speaking time in the target language.

Nevertheless, the technological competencies are not the only ones that need to be acquired and developed by teachers and learners. This, of course, functions together with developing their CIC as well. It is not merely sufficient to know how to use the audio activation/deactivation features; teachers and learners also need to be able to know when to use these features to create more opportunities for participation and learning. As the analysis shows in Chapter 5, the OMLs' developed CIC and e-CIC afforded them the ability to use the audio activation/deactivation features effectively. They successfully projected the possible completion of the sequence/current speaker's turn and coordinated their audio activation with it, which resulted in successful access to the ongoing class activity and enhancing speaking time in the target language. Also, Chapter 5 showed how the teachers' developed CIC was evident in enabling the OMLs' to secure the conversational floor following the audio activation, thereby creating a potential learning opportunity. Similarly, another manifestation of a developed CIC and e-CIC is evident in the teachers' use of the audio deactivation feature as a pre-emptive move. This enabled the current speaker to maintain the conversational floor and gain more speaking time in the target language.

Overall, acquiring CIC and e-CIC is essential for both teachers and learners to successfully manage interaction in online, synchronous, video-mediated language learning classrooms. Furthermore, the successful management of the interaction and effective use of the platform's

affordances can maximise the learners' learning opportunities, which is the main objective of these classes.

8.8 Summary

This chapter began by summarising the research's findings in Section 8.2, where a similar order to the analysis chapters was adopted. Therefore, the findings from each section of the analysis chapter were reported in this section. Next, Section 8.3 discussed the main findings that emerged from the analysis of the OMLs' use of the audio activation/deactivation features to manage turn-taking in relation to the existing literature. Section 8.4 then discussed the main findings emerging from the participants' use of the audio activation/deactivation features to manage trouble caused by the interference of disruptive noises. The discussion is also done in relation to the existing literature on repair in general and repair in video-mediated interaction. Following this, Section 8.5 discussed how the participants in the study managed to create novel practices using the audio activation/deactivation features in the management of their turn-taking and repair. Finally, Section 8.6 demonstrated the importance of the online classroom interactional competencies for both teachers and learners to successfully manage the interaction and maximise learning opportunities for the learners.

Chapter 9. Conclusions

9.1 Introduction

The previous chapter summarised and discussed the current study's findings in light of the existing literature. This final chapter aims to revisit the study's objectives and show how these objectives have been achieved. Also, it aims to present an argument for the importance of the findings of the current study by showing its contributions to the existing knowledge and its implications. Moreover, the chapter offers recommendations for future research.

The current study draws on calls to understand the role of mediating technologies in facilitating or constraining the construction of social interaction (Arminen et al., 2016; Hutchby, 2001a, 2013, 2014). It also responds to calls for addressing the empirical question of how talk-in-interaction is managed in online computer-mediated platforms (Jenks, 2014). Another call that this study responded to was to the examination of how the on-mute learners (OMLs) join an ongoing conversation (Andrews, 2020). Oittinen (2020) also called for more work on larger datasets to extend the understanding of how participants coordinate their actions to reach the smooth running of interaction in different contexts mediated by technologies. Moreover, it takes into account the scarcity of literature on the examination of what teachers do in synchronous online lessons using CA as a research methodology (Moorhouse et al., 2022).

The study employed multimodal CA to explore the management of interaction in online, synchronous, video-mediated L2 speaking classes using Zoom's audio activation/deactivation features. Using CA analytical tools facilitated the rich and detailed description of the teachers' and learners'' use of these features in managing video-mediated L2 speaking classrooms. A total of 32 hours of video recordings of teachers' screens while teaching on Zoom constitute the data examined in the current study.

As pointed out in Section 1.3, the current study's overriding objective is to examine the teachers' and learners' use of the medium's features to manage interaction in online, synchronous, video-mediated L2 classrooms. More specifically, this study has aimed to examine the teachers' and learners' use of the audio activation/deactivation features of Zoom to manage turn-taking, and (2) to examine the teachers' and learners' use of these features to manage trouble caused by the interference of disruptive background noises. In what follows,

an explanation of how the two objectives are achieved will be presented. Additionally, how the findings of this study contribute to the existing knowledge is presented.

First, the objective of examining the teachers' and learners' use of the audio activation/deactivation features to manage aspects of turn-taking has been achieved through the detailed microanalysis, which revealed a reflexive relationship between using these features and the organisation of turn-taking in video-mediated L2 classroom interaction. The OMLs' use of the audio activation feature enabled them to project their incipient speakership. By tracking the positions at which the OMLs activated their audio, the analysis in Chapter 5 shows how the use of the audio activation feature by OMLs is well-coordinated and finely tuned with the turn in progress (Mondada, 2007). Moreover, it shows how the other participants in the class adjusted their conduct accordingly in recognition of the OMLs' self-selection for the next turn (ibid). Coordinating one's actions with those of others is one of the important concepts to studying multimodal interaction from a conversation analytic perspective. According to Hindmarsh and Pilnick (2002), witnessing the onset of an interlocuter "beginning to engage in a particular activity can project a trajectory of actions that routinely follow" (p.151).

In studying the OMLs' use of audio activation feature, the current study shows that its timing is essential to self-selection success. The analysis in Chapter 5 demonstrated the consequentiality of audio activation absence or delay on the OMLs' ability to participate in the ongoing class activity. The mistiming of audio activation (i.e. activated past the transition space) resulted in the OMLs missing their opportunity to take the next turn. Missing the opportunity to launch a turn means missing an opportunity to speak in the target language; thus, the right timing of audio activation becomes essential. The findings also show the participants' proactive work to maintain boundaries between their physical and class-shared space using the activation/deactivation features. This maintenance is done by keeping the microphone off and only activating it to partake in an ongoing classroom activity. Such maintenance is vital to successful interaction in a fragile transactional zone (Fornel, 1996). As the analysis showed, the absence of this maintenance could lead to the interference of disruptive background noises, which may interrupt the progression of the class activities.

The second aim of this study has been achieved by presenting a microanalysis of the participants' use of Zoom's audio activation/deactivation features to manage trouble caused by disruptive noises' interference in Chapters 6 and 7. The analysis in these two chapters revealed findings concerning the structure of repair sequences in video-mediated L2

classrooms, and the role that using the audio activation/deactivation features can play in constructing such sequences. Additionally, the use of these features by the participants played a role in the length of the time taken to restore the progression of class activity. Moreover, the study's findings show the participants' orientations to two types of preferences: for progressivity and self-repair, as illustrated in the analysis presented in Chapters 6 and 7. According to Jenks (2014), repair sequences, preference organisation, and other sequential aspects of talk are vital to the unfolding of all social interaction, including computer-mediated communication. The participants in this study oriented to their preference for progressivity, which manifested in: (1) the continuation of talk following the emergence of background noises; and (2) their multiple other-repair initiations in an effort to resolve conversational breakdowns caused by background noises' interference.

The participants also displayed their preference for self-repair by (1) delaying their orientations to the noises as disruptive and (2) offering a space for self-audio deactivation following the verbal other-repair initiation. The absence of self-audio deactivation by the trouble source producer triggered the teacher's decision to deactivate their audio. Moreover, the analysis revealed findings regarding the design of the turns in which the participants oriented to the interference of background noises as transgressive. Also, the analysis showed how the participants managed the responsibility for audio deactivation.

Based on the findings of the study, it is argued that the teachers' and learners' use of the audio activation/deactivation features has a reflexive relationship with the management of turn taking and repair in L2 video-mediated classroom interaction. Moreover, it is argued that the successful use of the audio activation/deactivation features enables the participants to participate in class activities while isolating their background noises. It is also argued that the unsuccessful use of these features is noticeable and accountable by the other participants and can hinder the on-mute learners' participation and thus learning. The development of participants' CIC and e-CIC plays a role in advancing the progressivity of interaction and creating more opportunities for participation and learning in L2 video-mediated classrooms.

The investigation of online data comes with its own methodological challenges and constraints. However, such challenges and constraints can be addressed, as will be shown in the following section.

9.2 Methodological Considerations

This section presents the methodological challenges and limitations faced when collecting and analysing video-mediated interaction in this study. The discussion below acknowledges these challenges and shows how the current study addressed them. As noted in Section 4.4.1, the data collected for this study are video recordings of the teachers' screens; it was not possible to record the learners' screens due to logistic issues. Acquiring the data from one perspective may result in limitations, such as the lack of access to the learners' on-screen activities. Such access would have benefited the understanding of the learners' use of the audio activation/deactivation to manage turn-taking. Moreover, it would have offered a better understanding of the causes of absence/delay of audio deactivation when their background noises interrupted the class activities.

A similar observation can be made regarding the inability to record the participants' physical environments. Such recordings would have offered a better understanding of their off-screen activities, especially before and after using audio activation/deactivation to manage turn-taking. Also, better access to the sources of background noises that interrupted the class activities would have been offered. It might help with access to the participant's off-camera efforts to handle background noises. However, the participants in video-mediated interaction do not have access to each other's on/off-screen activities unless shared or oriented to. Thus, the analysis in this study is only limited to what is available to all the participants in the video-mediated classroom.

As highlighted in Section 4.2.1, Zoom can take different layouts on different screens. These layouts are not static; meaning the participants can toggle between speaker view and gallery view, or where there is a screen share they might move the video frames of their fellow participants around the screen during the class. With the difficulty of recording the learners' screens, it might be helpful to conduct quick surveys to ask the learners regarding the Zoom layout they had and the location of the video frames. Another possible solution is to ask the learners to take multiple screenshots during the class to accommodate the fluid organisation of the Zoom layouts. This can boost the researcher's ability to make claims around aspects of conduct, such as gaze movement. According to Seedhouse (2022), CA can benefit from interviews in revealing significant aspects that may not manifest in the recorded interaction. Further, Seedhouse argues that it is good practice for conversation analysts to clearly state the ethnographic knowledge they integrated in their analysis. However, the analysts need to be

clear about which parts of their analysis come from examining the interaction and which parts are learnt from the ethnographic knowledge (i.e. interviews) (Seedhouse, 2022).

Another consideration is relevant to a skill that the researchers examining video-mediated data need to pay attention to. It can be helpful for the researcher to acquire sufficient knowledge of how the communication software works. This can help the researcher to make sense of the data and make decisions on which parts of the data to present in the study. Moreover, Rintel (2015) points out that such knowledge can help the researcher to develop (or not) some claims around aspects of conduct.

A challenge, nonetheless, that researchers analysing similar data to the current study might experience relates to transcribing it. The amount of details in this type of data is enormous, especially when there is a screen share, mouse cursor movements, clicks, typing, verbal and non-verbal conduct all occurring simultaneously. The researchers aspire to capture all the necessary details in their transcripts, but at the same time they are also aspiring to ease the readability of these transcripts. Thus, the researchers need to devise creative ways to represent and describe the different details in a way that creates a balance between their aspirations and the readers' needs. Good examples in this area are the additions to the transcription conventions can be found in a number of recent studies (Balaman, 2016; Balaman & Doehler, 2022; Meredith, 2016).

Moreover, when the researcher asks the participants to handle the recording of the classes themselves, this can create two challenges. The first challenge concerns the sufficiency of the hardware owned by the participant. In the current study, several participants experienced difficulties with their devices regarding processors and sufficient memory capabilities. Such difficulties make the simultaneous operation of Zoom, the screen recorder, and the web browser a challenging task. The second challenge concerns the logistics, which means uploading large files to the internet and sharing them with the researcher. The majority of learners in the current study experienced recurrent connection problems, showing low internet speed. Therefore, it will be highly challenging for them to share their screen recordings with the researcher.

Another methodological challenge in this study was to ensure the visibility of the OMLs' audio status indicator to fellow participants. The analysis in Chapter 5 shows the OMLs' use of the audio activation/deactivation features to project the possible completion of the current speaker's turn and project their incipient speakership. Zoom allows the participants to make their audio status to the other participants by placing a microphone with a slash crossing it

next to their names when it is deactivated. The appearance and disappearance of this icon, next to the participants' names represent their audio status. The challenge was that Zoom could have different layouts on each participant's screens depending on their preference (see Section 4.2.1). To ensure that the audio status indicator is available to all the participants, only instances where no participant shared their screen was included. In this case, the audio status of the OML was available to fellow participants, regardless of the Zoom layout on their screens.

9.3 Limitations of the Study

This section acknowledges potential critiques and the limitations of the current study. These limitations are related to the adequacy of the data collected and how it may impact the researcher's access to some of the details. In what follows, these limitations are listed with details of how they can be addressed in future research. Firstly, a possible limitation of the current study was in collecting the data from one perspective (i.e. the teachers' screens), as pointed out in Sections 4.4.1 and 8.7. This may limit the ability to produce rigorous arguments around points, such as the length of pauses, given that some participants may experience transmission delays (Lange, 2020). Thus, future researchers can also consider collecting data from the learners' screens. This should help them account for this shortcoming by comparing the recordings of participants' screens and producing more rigorous arguments around the areas mentioned above. In addition, collecting data from different perspectives in video-mediated interaction can help the researcher to gain access to the participants' screen activities.

Secondly, some computer users may connect two or more monitors to one computer and can use them simultaneously. This was the case for one of the teachers in the latter set of the data recorded in the current study. The recordings show only one screen on which the Zoom window appeared, which might limit the researcher's access to some of the teacher's screen activities at certain points. Nevertheless, the analysis focused on the screen activities related to the Zoom window, which was always recorded. Thus, it is recommended that future researchers can consider the possibility of recording both screens of the participants to ensure full access to their screen activities.

Thirdly, as noted in the previous section, Zoom can take different layouts on the participants' screens. The lack of recordings of the learners' screens limited access to how the Zoom window appeared on their screens. This limited the researcher's choice of instances to present

in the study and made it necessary to only include the instances where there was no screen share to ensure the visibility of the participants' audio status.

Despite the limitations, the current study provides several contributions to the existing knowledge by achieving the study's objectives and providing a rich multimodal microanalysis of the teachers' and learners' use of the audio activation/deactivation features to manage turn-taking and repair in video-mediated L2 speaking classrooms. The following sections outline these contributions and shed light on the current study's implications.

9.4 Contributions of the Study

The findings of this study contribute to the existing knowledge concerning a number of areas. Firstly, the current study adds to the knowledge regarding the use of the medium's affordances in L2 video-mediated interaction (Andrews, 2020; Cheung, 2021; Moorhouse et al., 2021; NurSürüç Şen, 2022). This study incorporated Hutchby's (2001, 2003, 2014) notion of affordances, which offers an approach to understanding the complex relationship between the technology features and the construction of social interaction (i.e. how the mediating technology affords or constraints the accomplishment of social interaction). This study's contribution lies in the detailed description of the participants' use of the audio activation/deactivation features to manage interaction in L2 video-mediated interaction. To knowledge, no previous research has been conducted on using these features to manage interaction in a video-mediated classroom context.

Secondly, the current study contributes to the knowledge about turn-taking in video-mediated L2 teaching and learning settings using CA as a research methodology. Although a number of CA researchers have investigated L2 video-mediated interaction (Badem-Korkmaz & Balaman, 2022; Balaman, 2019; Balaman & Doehler, 2022, 2022; Balaman & Sert, 2017; Malabarba et al., 2022; Rusk & Pörn, 2019), it remains an under-researched area, and more work needs to be undertaken to fully appreciate the uniqueness of turn-taking in such a setting, and the role of using the medium's affordances in managing it.

Moreover, the study adds to the knowledge about repair organisation in video-mediated L2 teaching and learning settings. According to Moorhouse et al. (2022), the CA literature on what teachers do to manage video-mediated L2 classrooms is scarce. This study offers insights into what teachers and learners do to manage turn-taking and interruptions caused by background noises' interference in video-mediated L2 classrooms. In addition, this adds to the knowledge regarding how the teachers and learners maintain and restore the progressivity of

interaction in video-mediated L2 classes. Also, the study contributes to the knowledge concerning preference organisation in video-mediated L2 classrooms by shedding light on the participants' orientations to preference for progressivity and for self-repair.

Thirdly, the study adds to the knowledge relating to the participants' work to maintain the boundaries between their local physical spaces and the class-shared environment. According to Fornel (1996, p.47), it is the participants' responsibility to maintain these boundaries to create "an 'adequate' interactional frame", as the interference of background noises, for example, can threaten this extremely fragile transactional zone. The study also showed how the OMLs actively used the medium's audio activation/deactivation features to maintain such boundaries between the local and shared environment. In addition, it showed how the OMLs' use of the medium's affordances helped minimise potential interruptions by background noises' interference and created more participation and speaking time in the target language.

9.5 Implications of the Study

Although the current study primarily focused on examining the participants' use of the audio activation/deactivation features to manage aspects of turn-taking and repair, its findings can shed light on the role this use can have in maximising participation in video-mediated classrooms. Chapter 5 showed that using the audio activation/deactivation features enabled the OMLs to actively engage in the ongoing class activities and simultaneously isolate their background noises. Moreover, by using the audio activation/deactivation features, the participants managed to resolve interactional trouble caused by the interference of background noises and maintained/restored the progressivity of the suspended activities. By using these features to resolve interactional troubles, the participants maximised the progressivity of the ongoing interaction and, thus, secured more speaking time and potential learning opportunities.

When speaking about the participants' use of the audio activation/deactivation features, it is worth noticing that the competent use of these features has proven beneficial to maximising learning opportunities. This highlights the importance of considering the development of teachers' and learners' technological competencies in training programmes. For teachers, developing such competencies in the online teacher training programme must be included (Hampel & Stickler, 2015). However, developing the teachers' technological competencies can help overcome such interruptions after they occur. While training the learners on using the platform's features effectively can help pre-empt such troubles, which can result in a smooth run of interaction in video-mediated classrooms. The learners' developed CIC and e-

CIC (Moorhouse et al., 2021, 2022; Walsh, 2011) have proven to be beneficial, as seen in the findings presented in Chapter 5. Thus, it can be argued that developing both teachers' and learners' technological competencies is essential for successful interaction management.

Although this study examines using Zoom's audio activation/deactivation features in the L2 classroom context, its findings can also have implications for other settings, as using these features is not exclusive to participants in the classroom context. Interruptions by the interference of background noises can occur in other contexts, especially with the rising shift towards working from home or other public venues, such as cafes. Thus, the findings of the current study may help participants in other video-mediated contexts to improve their practices while being in such environments.

Overall, this study has provided several insights into the management of interaction in video-mediated L2 classrooms. However, more work is required to explore interaction in such a setting. The following section offers recommendations for future researchers based on some of the current study's limitations.

9.6 Recommendations for Future Research

This study examined interaction management using the audio activation/deactivation features in video-mediated small-group L2 speaking classes. As was evident in the analysis in Chapter 6, the group size might play a role in the emergence of new patterns concerning using the audio activation/deactivation features. For example, Cheung (2021) drew attention to the teachers' use of these features in whole-class discussions. However, a multimodal microanalysis of such occurrences in larger classes by future researchers is a worthwhile quest.

Future researchers can also consider conducting longitudinal studies to examine the participants' development in using the communication platform's features to manage their classroom interaction over a period of time. Previous longitudinal studies that examined L2 video-mediated interaction (Balaman, 2016; Balaman & Doehler, 2022; Balaman & Sert, 2017; Sert & Balaman, 2018) have provided significant insights into examining the development of different practices in classroom interaction. Thus, carrying out such studies to understand the development of the participants' use of audio activation/deactivation features can be beneficial.

The current study does not primarily focus on examining the relationship between using audio activation/deactivation features and learning. Thus, future researchers may consider

examining how using these features can contribute to creating learning opportunities of the target language. This study examined the participants' use of the audio activation/deactivation features to manage interaction within the classroom contexts. Examining such use in other institutional contexts is worthwhile, given the different rules governing each context, such as courtrooms, medical consultations, business meetings, and conferences. Moreover, given the widespread use of video calls nowadays, examining the use of these features in mundane settings can be considered by future researchers.

References

Ajabshir, Z. F. (2019). The effect of synchronous and asynchronous computer-mediated communication (CMC) on EFL learners' pragmatic competence. *Computers in Human Behavior*, *92*, 169–177. https://doi.org/https://doi.org/10.1016/j.chb.2018.11.015

Alzaidi, H. O. G. (2016). *The practices of multiple other-initiated repair in online second language interaction* (Doctoral dissertation, Newcastle University).

Amar, C., Nanbu, Z., & Greer, T. (2022). Proffering absurd candidate formulations in the pursuit of progressivity. *Classroom Discourse*, *13*(3), 264–292.

Amiti, F. (2020). Synchronous and asynchronous E-learning. *European Journal of Open Education and E-Learing Studies*, *5*(2), 60-70.

Andrews, E. (2020). Disembodied Teaching and Learning: Contributions from Speech Acts, Peircean Sign Theory and Multimodal Approaches to Embodied Cognition. *Glossos*, *15*, 1-12

Antaki, C. (2008). Discourse analysis and conversation analysis. The SAGE Handbook of Social Research Methods, 431–446.

Antaki, C. (2012). Affiliative and disaffiliative candidate understandings. *Discourse Studies*, *14*(5), 531–547.

Arminen, I. (2017). *Institutional Interaction: Studies of Talk at Work*. Routledge

Arminen, I., Licoppe, C., & Spagnolli, A. (2016). Respecifying Mediated Interaction. *Research on Language and Social Interaction*, *49*(4), 290–309.

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

Badem-Korkmaz, F., & Balaman, U. (2022). Eliciting student participation in video-mediated EFL classroom interactions: focus on teacher response-pursuit practices. *Computer Assisted Language Learning*, 1–23.

Balaman, U. (2016). *A Conversation Analytic Study on the Development of Interactional Competence in English in an Online Task-Oriented Environment* (Doctoral dissertation, Hacettpe University).

Balaman, U. (2019). Sequential organization of hinting in online task-oriented L2 interaction. *Text & Talk*, *39*(4), 511–534.

Balaman, U., & Doehler, S. P. (2022). Navigating the complex social ecology of screen-based

activity in video-mediated interaction. *Pragmatics*, *32*(1), 54–79.

Balaman, U., & Sert, O. (2017). Development of L2 interactional resources for online collaborative task accomplishment. *Computer Assisted Language Learning*, *30*(7), 601–630.

Batlle, J., & Deal, M. (2021). Teacher epistemic stance as a trouble in foreign language classroom interaction. *Journal of Pragmatics*, *176*, 15–25.

Bögels, S., & Torreira, F. (2015). Listeners use intonational phrase boundaries to project turn ends in spoken interaction. *Journal of Phonetics*, *52*, 46–57.

Bokhove, C. (2018). Exploring classroom interaction with dynamic social network analysis. *International Journal of Research & Method in Education*, *41*(1), 17–37.

Brandt, A., & Jenks, C. (2013). Computer-mediated spoken interaction: Aspects of trouble in multi-party chat rooms. *Language@Internet*, *10*(5).

Bryman, A. (2016). *Social research methods*, 5th edition, Oxford University Press.

Carroll, D. (2004). Restarts in Novice Turn Beginnings: Disflyencies or Interactional Achievements? In R. Gardner & J. Wagner (Eds.), *Second Language Conversations* (1st ed.). Bloomsbury Academic, 201–220. http://www.bloomsburycollections.com/book/second-language-conversations/restarts-in-novice-turn-beginnings-disflyencies-or-interactional-achievements/

Cheung, A. (2021). Synchronous online teaching, a blessing or a curse? Insights from EFL primary students' interaction during online English lessons. *System*, *100*, 1-13.

De Fornel, M. (1996). The Interactional Frame of Videophonic Exchange. *Réseaux. The French Journal of Communication*, *4*(1), 47–72.

Del-Moral-Pérez, M. E., Villalustre-Martínez, L., & Neira-Piñeiro, M. del R. (2019). Teachers' perception about the contribution of collaborative creation of digital storytelling to the communicative and digital competence in primary education schoolchildren. *Computer Assisted Language Learning*, *32*(4), 342–365.

Dolce, F. M., & van Compernolle, R. A. (2020). Topic management and student initiation in an advanced Chinese-as-a-foreign-language classroom. *Classroom Discourse*, *11*(1), 80–98.

Drew, P. (2005). Conversation Analysis. *Handbook of Language and Social Interaction*, 71–102.

Due, B. L., & Licoppe, C. (2021). Video-mediated interaction (VMI): introduction to a special issue on the multimodal accomplishment of VMI institutional activities. Social Interaction. *Video-Based Studies of Human Sociality*, 3(3), 1-20

Due, B. L., Lange, S. B., Nielsen, M. F., & Jarlskov, C. (2019). Mimicable embodied demonstration in a decomposed sequ-292nce: Two aspects of recipient design in professionals' video-mediated encounters. *Journal of Pragmatics*, *152*, 13–27.

Duncan, S. (1972). Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology*, *23*(2), 283-292

Duran, D., & Sert, O. (2021). Student-initiated multi-unit questions in EMI classrooms. *Linguistics and Education*, *65*, 1-10.

Duran, D., Kurhila, S., & Sert, O. (2022). Word search sequences in teacher-student interaction in an English as medium of instruction context. *International Journal of Bilingual Education and Bilingualism*, *25*(2), 502–521.

Evnitskaya, N., & Berger, E. (2017). Learners' Multimodal Displays of Willingness to Participate in Classroom Interaction in the L2 and CLIL Contexts. *Classroom Discourse*, *8*, 1–24. https://doi.org/10.1080/19463014.2016.1272062

Fasel Lauzon, V., & Berger, E. (2015). The multimodal organization of speaker selection in classroom interaction. *Linguistics and Education*, *31*, 14–29.

Firth, A. (1996). The discursive accomplishment of normality: On 'lingua franca'English and conversation analysis. *Journal of Pragmatics*, *26*(2), 237–259.

Fornel, M. de. (1994). Le cadre interactionnel de l'échange visiophonique. *Réseaux. Communication - Technologie - Société*, *12*(64), 107–132.

Warner-Garcia, S. (2014). Laughing when nothing's funny: The pragmatic use of coping laughter in the negotiation of conversational disagreement. *Pragmatics. Quarterly Publication of the International Pragmatics Association (IPrA)*, *24*(1), 157–180.

Gardner, R. (2019). Classroom interaction research: The state of the art. *Research on Language and Social Interaction*, *52*(3), 212–226.

Gibbons, P. (1998). Classroom Talk and the Learning of New Registers in a Second Language. *Language and Education*, *12*(2), 99–118.

Gibson, J. J. (1979). The ecological approach to visual perception. Psychology Press

Glenn, P., & Holt, E. (2013). Introduction. In P. Glenn & E. Holt (Eds.). *Studies of Laughter*

in Interaction (pp. 1–22). London: Bloomsbury Academic. Retrieved June 22, 2023, from http://dx.doi.org.libproxy.ncl.ac.uk/10.5040/9781472542069.ch-001

González-Lloret, M. (2008). Computer-mediated learning of L2 pragmatics. *Investigating Pragmatics in Foreign Language Learning, Teaching and Testing*, *30*, 114–132.

González-Lloret, M. (2015). Conversation analysis in computer-assisted language learning. *Calico Journal*, 32(3), 569–594.

Goodwin, C. (1980). Restarts, Pauses, and the Achievement of a State of Mutual Gaze at Turn-Beginning. *Sociological Inquiry*, *50*(3–4), 272–302.

Goodwin, C. (1981). *Conversational Organization: Interaction Between Speakers and Hearers*.

Goodwin, C. (1986). *Gestures as a resource for the organization of mutual orientation*. 62(1-2), 29-50.

Goodwin, C. (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics*, *32*(10), 1489–1522.

Graham, C. R. (2006). *Blended learning systems: Definition, current trends, and future directions (w) Bonk CJ, Graham CR (red.), The Handbook of Blended Learning: Global Perspectives*. San Francisco: Pfeiffer Publishing.

Grammens, M., Voet, M., Vanderlinde, R., Declercq, L., & De Wever, B. (2022). A systematic review of teacher roles and competences for teaching synchronously online through videoconferencing technology. *Educational Research Review*, *37*, 1-16

Guichon, N. (2009). Training future language teachers to develop online tutors' competence through reflective analysis. *ReCALL*, *21*(2), 166–185.

Gunawardena, C. N., Nolla, A. C., Wilson, P. L., Lopez-Islas, J. R., Ramirez-Angel, N., & Megchun-Alpizar, R. M. (2001). A cross-cultural study of group process and development in online conferences. *Distance Education*, *22*(1), 85–121.

Guo, E., & Zhang, F. (2021). *From Offline to Online: The Reduced Embodiment in Teacher-Initiated Turn-Taking in GSL Instructions*, 357–369

Haddington, P., Keisanen, T., Mondada, L., & Nevile, M. (2014). *Multiactivity in social interaction : beyond multitasking*. AmsterdamPhiladelphia: John Benjamins Publishing Company.

Hampel, R., & Stickler, U. (2005). New skills for new classrooms: Training tutors to teach

languages online. *Computer Assisted Language Learning*, *18*(4), 311–326.

Hampel, R., & Stickler, U. (2015). *Developing online language teaching : research-based pedagogies and reflective practices* (1st ed. 20). Houndmills, Basingstoke, Hampshire.

Harper, R., Watson, R., & Licoppe, C. (Eds.). (2019). *Skyping the Family: Interpersonal video communication and domestic life*. John Benjamins.

Hayashi, M. (2004). Projection and grammar: notes on the 'action-projecting' use of the distal demonstrative are in Japanese. *Journal of Pragmatics*, *36*(8), 1337–1374.

Hazel, S., & Mortensen, K. (2017). The classroom moral compass – participation, engagement and transgression in classroom interaction. *Classroom Discourse*, *8*(3), 214–234.

Heath, C. (2002). Demonstrative Suffering: The Gestural (Re)embodiment of Symptoms. *Journal of Communication*, *52*(3), 597–616.

Heritage, J. (1984). *Garfinkel and ethnomethodology*. Cambridge CambridgeshireNew York: Polity Press.

Heritage, J. (2005). Conversation analysis and institutional talk. *Handbook of Language and Social Interaction*, *103*, 47.

Herring, S. C. (1996). Computer-mediated communication. *Computer-Mediated Communication*, 1–332.

Herring, S. C. (2019). *The Coevolution of Computer-Mediated Communication and Computer-Mediated Discourse Analysis BT - Analyzing Digital Discourse: New Insights and Future Directions* (P. Bou-Franch & P. Garcés-Conejos Blitvich (Eds.). Springer International Publishing, 25–67

Hindmarsh, J., & Pilnick, A. (2002). The Tacit Order of Teamwork: Collaboration and Embodied Conduct in Anesthesia. *The Sociological Quarterly*, *43*(2), 139–164.

Hırçın Çoban, M., & Sert, O. (2020). Resolving interactional troubles and maintaining progressivity in paired speaking assessment in an EFL context. *Papers in Language Testing and Assessment*, *9*(1), 64–94.

Hjulstad, J. (2016). Practices of Organizing Built Space in Videoconference-Mediated Interactions. *Research on Language and Social Interaction*, *49*(4), 325–341.

Hoey, E. M., & Kendrick, K. H. (2017). Conversation analysis. *Research Methods in Psycholinguistics: A Practical Guide*, 151–173.

Hosoda, Y., & Aline, D. (2013). Two preferences in question–answer sequences in language classroom context. *Classroom Discourse*, *4*(1), 63–88.

Hutchby, I. (2001a). *Conversation and technology : from the telephone to the internet*. Cambridge, UK .

Hutchby, I. (2001b). Technologies, texts and affordances. Sociology, 35(2), 441–456.

Hutchby, I. (2003). Affordances and the Analysis of Technologically Mediated Interaction:: A Response to Brian Rappert. *Sociology*, *37*(3), 581–589.

Hutchby, I. (2014). Communicative affordances and participation frameworks in mediated interaction. *Journal of Pragmatics*, *72*, 86–89.

Hutchby, I., & Wooffitt, R. (2008). *Conversation analysis*. Polity.

Ingram, J., & Elliott, V. (2014). Turn taking and 'wait time' in classroom interactions. *Journal of Pragmatics*, *62*, 1–12.

Ilomäki, S., Ruusuvuori, J., & Laitinen, J. (2021). Effects of transmission delay on client participation in video-mediated group health counseling. *Qualitative Health Research*, *31*(12), 2328–2339.

Jakonen, T. (2016). Managing multiple normativities in classroom interaction: Student responses to teacher reproaches for inappropriate language choice in a bilingual classroom. *Linguistics and Education*, *33*, 14–27.

Jakonen, T., & Jauni, H. (2021). Mediated learning materials: visibility checks in telepresence robot mediated classroom interaction. *Classroom Discourse*, *12*(1–2), 121–145.

Jefferson, G. (1984). Notes on a systematic deployment of the acknowledgement tokens "Yeah"; and "Mm Hm." *Paper in Linguistics*, *17*(2), 197–216.

Jefferson, G. (2004). Glossary of transcript symbols with an introduction. *Pragmatics and Beyond New Series*, *125*, 13–34.

Jenks, C. (2014). *Social interaction in second language chat rooms*. Edinburgh University Press.

Jenks, C. J. (2009). When is it appropriate to talk? Managing overlapping talk in multi-participant voice-based chat rooms. *Computer Assisted Language Learning*, *22*(1), 19–30.

Jenks, C. J. (2014). Contextual Variables in CMSI. In *Social Interaction in Second Language Chat Rooms*. Edinburgh University Press, 95–120

Jewitt, C., Bezemer, J., & O'Halloran, K. (2016). *Introducing multimodality*. Routledge.

Jordan, I. J. (1990). Self-selection in turn-taking: An approach to teaching Spanish. *Hispania, 73*(4), 1154–1157.

Kääntä, L. (2010). Teacher turn-allocation and repair practices in classroom interaction: A multisemiotic perspective. Jyväskylä Studies in Humanities, 137

Kääntä, L. (2012b). Teachers' embodied allocations in instructional interaction. *Classroom Discourse, 3*(2), 166–186.

Keegan, D. (2013). *Foundations of distance education*. Routledge.

Kendrick, K. H. (2015). *Other-initiated repair in English*. *1*(1), 164-190

Klattenberg, R. (2021). Question-formatted reproaches in classroom management. *Classroom Discourse, 12*(3), 214–232.

Kohnke, L., & Moorhouse, B. L. (2022). Facilitating Synchronous Online Language Learning through Zoom. *RELC Journal, 53*(1), 296–301.

Komter, M. (2012). Conversation Analysis in the Courtroom. In *The Handbook of Conversation Analysis*, 612–629.

Kurhila, S. (2004). Clients or Language Learners–Being a Second Language Speaker in Institutional. *Second Language Conversations: Studies of Communication in Everyday Settings*, 58.

Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia : University of Pennsylvania Press.

Lauzon, V. F., & Berger, E. (2015). The multimodal organization of speaker selection in classroom interaction. *Linguistics and Education, 31*, 14–29.

Lee, J. (2017). Multimodal turn allocation in ESL peer group discussions. *Social Semiotics, 27*(5), 671–692.

Lerner, G. (2002). *Turn-sharing: The choral co-production of talk-in-interaction. Chapter 9 of The Language of Turn and Sequence, ed. C. Ford, B. Fox, and S. Thompson, 225-256*. Oxford: Oxford University Press.

Lerner, G. H. (1991). On the syntax of sentences-in-progress. *Language in Society, 20*(3), 441–458.

Levinson, S. C. (2012). Action Formation and Ascription. In *The Handbook of Conversation Analysis*, 101–130.

Levinson, S. C., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology*, *6*, 10-26

Licoppe, C., & Dumoulin, L. (2010). The "Curious Case" of an Unspoken Opening Speech Act: A Video-Ethnography of the Use of Video Communication in Courtroom Activities. *Research on Language and Social Interaction*, *43*(3), 211–231.

Licoppe, C., & Morel, J. (2014). *Mundane video directors. Showing one's environment in Skype and mobile video calls*, 135–160

Licoppe, C. (2017). Showing objects in Skype video-mediated conversations: From showing gestures to showing sequences. *Journal of Pragmatics*, *110*, 63–82. https://doi.org/https://doi.org/10.1016/j.pragma.2017.01.007

Licoppe, C., & Morel, J. (2018). Visuality, text and talk, and the systematic organization of interaction in Periscope live video streams. *Discourse Studies*, *20*(5), 637–665.

Liddicoat, A. (2022). *An introduction to conversation analysis* (Third edit). London .

Lorenza, M. (2019). Transcribing silent actions: a multimodal approach of sequence organization. *Social Interaction. Video-Based Studies of Human Sociality*, *2*(1).

Luff, P., Heath, C., Kuzuoka, H., Hindmarsh, J., Yamazaki, K., & Oyama, S. (2003). Fractured Ecologies: Creating Environments for Collaboration. *Human-Computer Interaction*, *18*, 51–84.

Luff, P., Jirotka, M., Yamashita, N., Kuzuoka, H., Heath, C., & Eden, G. (2013). Embedded interaction: The accomplishment of actions in everyday and video-mediated environments. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *20*. https://doi.org/10.1145/2442106.2442112

Luff, P., Heath, C., Yamashita, N., Kuzuoka, H., & Jirotka, M. (2016). Embedded Reference: Translocating Gestures in Video-Mediated Interaction. *Research on Language and Social Interaction*, *49*(4), 342–361.

Macbeth, D. H. (1990). Classroom Order as Practical Action: the making and un-making of a quiet reproach. *British Journal of Sociology of Education*, *11*(2), 189–214.

Mahu, D.-P. (2012). Why Is Learning English So Beneficial Nowadays? *International Journal of Communication Research*, *2*(4), 374.

Malabarba, T., Mendes, A. C. O., & de Souza, J. (2022). Multimodal Resolution of Overlapping Talk in Video-Mediated L2 Instruction. In *Languages 7*(2). 1-22

Margutti, P. (2011). Teachers' reproaches and managing discipline in the classroom: When

teachers tell students what they do ‘wrong.’ *Linguistics and Education*, *22*(4), 310–329.

McAndrew, P., Foubister, S. P., & Mayes, T. (1996). Videoconferencing in a language learning application. *Interacting with Computers*, *8*(2), 207–217.

Mchoul, A. (1978). The organization of turns at formal talk in the classroom. *Language in Society*, *7*(2), 183–213.

McHoul, A. W. (1990). The Organization of Repair in Classroom Talk. *Language in Society*, *19*(3), 349–377.

Mehan, H. (1979). Learning Lessons: Social Organization in the Classroom. Cambridge, MA and London, England: Harvard University Press. https://doi.org/10.4159/harvard.9780674420106

Meier, A., & Reinecke, L. (2020). Computer-Mediated Communication, Social Media, and Mental Health: A Conceptual and Empirical Meta-Review. *Communication Research*, *48*(8), 1182–1209.

Melander Bowden, H., & Svahn, J. (2020). Collaborative work on an online platform in video-mediated homework support. *Social Interaction. Video-Based Studies of Human Sociality*, *3*(3).

Meredith, J. (2016). Transcribing screen-capture data: the process of developing a transcription system for multi-modal text-based data. *International Journal of Social Research Methodology*, *19*(6), 663–676.

Mlynář, J., González-Martínez, E., & Lalanne, D. (2018). Situated organization of video-mediated interaction: A review of ethnomethodological and conversation analytic studies. *Interacting with Computers*, *30*(2), 73–84.

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). Instructions in the operating room: How the surgeon directs their assistant’s hands. *Discourse Studies*, *16*(2), 131–161.

Mondada, L. (2015). *8. Introduction de nouveaux référents et gestion de l’attention conjointe : les apports de l’analyse interactionnelle multimodale*, 121-138.

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.

Moorhouse, B. L., Li, Y., & Walsh, S. (2021). E-Classroom Interactional Competencies: Mediating and Assisting Language Learning During Synchronous Online Lessons. *RELC Journal*, 1-15.

Moorhouse, B. L., Walsh, S., Li, Y., & Wong, L. L. C. (2022). Assisting and Mediating Interaction during Synchronous Online Language Lessons: Teachers' Professional Practices. *TESOL Quarterly*, *56*(3), 934–960.

Mortensen, K. (2008). Selecting next speaker in the second language classroom: How to find a willing next speaker in planned activities. *Journal of Applied Linguistics*, *5*(1), 55-79

Mortensen, K. (2009). Establishing Recipiency in Pre-Beginning Position in the Second Language Classroom. *Discourse Processes*, *46*(5), 491–515.

Mortensen, K. (2016). The body as a resource for other-initiation of repair: Cupping the hand behind the ear. *Research on Language and Social Interaction*, *49*(1), 34–57.

Mortensen, K., & Hazel, S. (2011). Initiating round robins in the L2 classroom - preliminary observations. *Novitas-ROYAL*, *5*(1), 55–70.

Mosa, A. A., Mahrin, M. N. bin, & Ibrrahim, R. (2016). Technological Aspects of E-Learning Readiness in Higher Education: A Review of the Literature. *Comput. Inf. Sci.*, *9*(1), 113–127.

Negretti, R. (1999). *Web-based activities and SLA: A conversation analysis research approach*. *3*, 75–87.

Norris, S. (2004). Analyzing multimodal interaction: A methodological framework. Routledge.

Norris, S. (2020). Multimodal theory and methodology: For the analysis of (inter) action and identity. Routledge.

NurSürüç Şen, S. (2022). Musicians of the Titanic: An EFL teacher's classroom management experiences during Covid-19 pandemic. *Journal*, *5*(1), 14–31.

O'Rourke, B., & Stickler, U. (2017). *Synchronous communication technologies for language learning: Promise and challenges in research and pedagogy*. *7*(1), 1–20.

Ockey, G. J., Timpe-Laughlin, V., Davis, L., & Gu, L. (2019). Exploring the potential of a video-mediated interactive speaking assessment. *ETS Research Report Series*, *2019*(1), 1–29.

Oittinen, T. (2020). Noticing-prefaced recoveries of the interactional space in a video-

mediated business meeting. *Social Interaction. Video-Based Studies of Human Sociality*, *3*(3), 1-25

Paul, S., & Asa, B. Turn taking in student Zoom discussions. Language, Cultute and Society, (20), 1-29.

Payne, J. S. (2020). Developing L2 productive language skills online and the strategic use of instructional tools. *Foreign Language Annals*, *53*(2), 243–249.

Pekarek Doehler, S., & Berger, E. (2019). On the reflexive relation between developing L2 interactional competence and evolving social relationships: A longitudinal study of word-searches in the 'wild.' *Conversation Analytic Research on Learning-in-Action: The Complex Ecology of Second Language Interaction 'in the Wild,'* 51–75.

Petitjean, C. (2014). Social representations of turn-taking in classrooms: from compulsory to post-compulsory schooling in French-speaking Switzerland. *Classroom Discourse*, *5*(2), 138–157.

Pirini, J. (2015). Research into tutoring: Exploring agency and intersubjectivity, PhD Thesis, Auckland University of Technology. Retrieved from https://openrepository.aut.ac.nz/handle/10292/9267

Pomerantz, A. (1986). Extreme case formulations: A way of legitimizing claims. *Human Studies*, *9*(2), 219–229.

Richards, K., & Seedhouse, P. (2004). *Applying conversation analysis*. Basingstoke : Palgrave Macmillan.

Rintel, S. (2013). Video calling in long-distance relationships: The opportunistic use of audio/video distortions as a relational resource. *The Electronic Journal of Communication/La Revue Electronic de Communication (EJC/REC)*, *23*(2), 1-19

Rintel, S. (2015). *Conversation Analysis of Video-Mediated Communication: Interactional Repair of Distortion in Long-Distance Couples' Video Calls*. SAGE Publications

Robinson, J. D. (2006). Managing trouble responsibility and relationships during conversational repair. *Communication Monographs*, *73*(2), 137–161.

Rodriguez, J. B., & Wilstermann, I. M. (2018). Learner initiative in the Spanish as a foreign language classroom: Implications for the interactional development. *Hacettepe University Journal of Education*, *33*, 113–133.

Rosenbaun, L., Rafaeli, S., & Kurzon, D. (2016). Blurring the boundaries between domestic and digital spheres. *Pragmatics. Quarterly Publication of the International*

Pragmatics Association, *26*, 291–304.

Rossano, F. (2013). Gaze in social interaction. *Handbook of Conversation Analysis*, 308–329.

Ruhleder, K., & Jordan, B. (2001a). Co-Constructing Non-Mutual Realities: Delay-Generated Trouble in Distributed Interaction. *Computer Supported Cooperative Work (CSCW)*, *10*(1), 113–138. https://doi.org/10.1023/A:1011243905593

Ruhleder, K., & Jordan, B. (2001b). Managing Complex, Distributed Environments: Remote Meeting Technologies at the "Chaotic Fringe." *First Monday*, *6*.

Rühlemann, C., & Gries, S. T. (2020). Speakers advance-project turn completion by slowing down: A multifactorial corpus analysis. *Journal of Phonetics*, *80*, 100976.

Rusk, F., & Pörn, M. (2019). Delay in L2 interaction in video-mediated environments in the context of virtual tandem language learning. *Linguistics and Education*, *50*, 56–70.

Sacks, H. (1984). On doing 'being ordinary.' *Structures of Social Action: Studies in Conversation Analysis*, 413–429.

Sacks, H. (1992). Lectures on conversation: Volume I. *Malden, Massachusetts: Blackwell*.

Sacks, H. (2020). On the preferences for agreement and contiguity in sequences in conversation. In *Language in Use*. Routledge, 8–22. (Original work published 1987)

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, *50*(4), 696–735.

Sahlström, J. F. (2002). The interactional organization of hand raising in classroom interaction. *The Journal of Classroom Interaction*, 47–57.

Satar, M., & Wigham, C. R. (2020). Delivering task instructions in multimodal synchronous online language teaching. *Alsic. Apprentissage Des Langues et Systèmes d'Information et de Communication*, *23*(1).

Satar, M., & Wigham, C. R. (2023). *Instruction Giving in Online Language Lessons: A Multimodal (Inter) action Analysis*. Routledge

Schegloff, E. (2010). Reflections on Quantification in the Study of Conversation. *Research on Language and Social Interaction*, *26*, 99–128.

Schegloff, E. A. (1979). The relevance of repair to syntax-for-conversation. In *Discourse and syntax* (pp. 261–286). Brill.

Schegloff, E. A. (1986). The Routine as Achievement. *Human Studies*, *9*(2/3), 111–151.

Schegloff, E. A. (1991). Reflections on talk and social structure. *Talk and Social Structure, 984*, 44–70.

Schegloff, E. A. (1996). Turn organization: one intersection of. *Interaction and Grammar*, *13*, 52–133.

Schegloff, E. A. (2007). Sequence organization in interaction: A primer in conversation analysis I, 1. Cambridge university press

Schegloff, E. A. (2013). Ten operations in self-initiated, same-turn repair. In G. Raymond, J. Sidnell, & M. Hayashi (Eds.), *Conversational Repair and Human Understanding* (pp. 41–70). Cambridge 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.

Seedhouse, P. (1996). Classroom interaction: possibilities and impossibilities. *ELT Journal*, *50*(1), 16–24.

Seedhouse, P. (2004a). *The interactional architecture of the language classroom : a conversation analysis perspective*. Malden, MA : Blackwell Pub.

Seedhouse, P. (2004b). Conversation Analysis Methodology. *Language Learning*, *54*(S1), 1–54.

Seedhouse, P. (2005). Conversation Analysis and language learning. *Language Teaching*, *38*(4), 165–187.

Sert, O. (2011). *A micro-analytic investigation of claims of insufficient knowledge in EAL classrooms* (Doctoral dissertation, Newcastle University).

Sert, O. (2013). 'Epistemic status check' as an interactional phenomenon in instructed learning settings. *Journal of Pragmatics*, *45*(1), 13–28.

Sert, O. (2015). *Social Interaction and L2 Classroom Discourse*. Edinburgh : Edinburgh University Press.

Sert, O. (2015). FROM TROUBLES TO RESOLUTIONMANAGEMENT OF DISPLAYS AND CLAIMS OF INSUFFICIENT KNOWLEDGE IN L2 CLASSROOMS. In *Social Interaction and L2 Classroom Discourse* (pp. 58–86). Edinburgh University Press.

Sert, O. (2017). Creating opportunities for L2 learning in a prediction activity. *System, 70*, 14–25.

Sert, O. (2019). 8. Mutual Gaze, Embodied Go-aheads and their Interactional Consequences in Second Language Classrooms. In J. Hall & S. Looney (Ed.), *TheEmbodied Work of Teaching* (pp. 142-159). Bristol, Blue Ridge Summit: Multilingual Matters.

Sert, O., & Balaman, U. (2018). Orientations to negotiated language and task rules in online L2 interaction. *ReCALL*, *30*(3), 355–374.

Sert, O., & Jacknick, C. M. (2015). Student smiles and the negotiation of epistemics in L2 classrooms. *Journal of Pragmatics*, *77*, 97–112. https://doi.org/https://doi.org/10.1016/j.pragma.2015.01.001

Sert, O., & Walsh, S. (2013). The interactional management of claims of insufficient knowledge in English language classrooms. *Language and Education*, *27*(6), 542–565.

Seuren, L. M., Wherton, J., Greenhalgh, T., & Shaw, S. E. (2021). Whose turn is it anyway? Latency and the organization of turn-taking in video-mediated interaction. *Journal of Pragmatics*, *172*, 63–78.

Sidnell, J. (2010). *Conversation analysis: An introduction* (Vol. 45). John Wiley & Sons.

Sinclair, J. M., & Coulthard, M. (1975). *Towards an analysis of discourse: The English used by teachers and pupils*. Oxford University Press, USA.

Sindoni, M. G. (2014). Through the looking glass: a social semiotic and linguistic perspective on the study of video chats. *Text & Talk*, *34*(3), 325–347.

Solem, M. S. (2016). Displaying knowledge through interrogatives in student-initiated sequences. *Classroom Discourse*, *7*(1), 18–35.

Stivers, T. (2015). Coding Social Interaction: A Heretical Approach in Conversation Analysis? *Research on Language & Social Interaction*, *48*, 1-19

Stivers, T., & Rossano, F. (2010). Mobilizing Response. *Research on Language and Social Interaction*, *43*(1), 3–31.

Stommel, W. J. P., & Stommel, M. W. J. (2021). *Participation of Companions in Video-Mediated Medical Consultations: A Microanalysis BT - Analysing Digital Interaction* (J. Meredith, D. Giles, & W. Stommel (Eds.), Springer International Publishing, 177–203

Streeck, J. (1995). On projection. In E. N. Goody (Ed.), *Social Intelligence and Interaction: Expressions and implications of the social bias in human intelligence*, Cambridge University Press, 87–110.

Svennevig, J. (2012). The agenda as resource for topic introduction in workplace meetings. In *Discourse studies.* 14(1), 53–66.

Takahashi, J. (2016). Contribution with Hand-Raising in Graduate Student Self-Selection: Bringing Legitimacy to the Focal Shift of Talk. *Working Papers in TESOL & Applied Linguistics*, *16*(2), 35–39.

Ten Have, P. (2007). Doing conversation analysis. In *introducing Qualitative Methods series* . Sage.

Ten Have, P. (2007). *Doing conversation analysis*. London : SAGE Publications.

Tsai, M.-H., & Kinginger, C. (2014). Giving and receiving advice in computer-mediated peer response activities. *CALICO Journal*, *32*(1), 82–112.

Tsui, A. B. M. (2001). Classroom interaction. In D. Nunan & R. Carter (Eds.), *The Cambridge Guide to Teaching English to Speakers of Other Languages*,. Cambridge University Press, 120–125.

Tudini, V. (2019). Conversation Analysis of Computer-Mediated Interactions. In *The Encyclopedia of Applied Linguistics*, 1–6.

Tudini, V. A. C. (2002). *The role of online chatting in the development of competence in oral interaction*. Griffith University.

Tutt, D., Hindmarsh, J., Shaukat, M., & Fraser, M. (2007). The Distributed Work of Local Action: Interaction amongst virtually collocated research teams. *In ECSCW 200: Proceedings of the 10th European Conference on Computer-Supported Cooperative Work, Limerick, Ireland*, Springer London, 199-218.

Uçar, H. & Kumtepe, A. T. (Eds.). (2021). *Motivation, Volition, and Engagement in Online Distance Learning*. IGI Global. https://doi-org.libproxy.ncl.ac.uk/10.4018/978-1-7998-7681-6

van Balen, J., Gosen, M. N., de Vries, S., & Koole, T. (2022). Taking learner initiatives within classroom discussions with room for subjectification. *Classroom Discourse*, 1–20.

Van Dijk, J. (2022). Zooming in on embodied social sensemaking: Mapping the design space in the context of videoconferencing. *Sixteenth International Conference on Tangible, Embedded, and Embodied Interaction*, 1–10.

Van Lier, L. (1988). *The classroom and the language learner : ethnography and second*

Language classroom research. London .

Walsh, S. (2011). *Exploring classroom discourse: Language in action*. Routledge.

Walsh, S. (2012). Conceptualising classroom interactional competence. *Novitas-Royal (Research on Youth and Language)*, *6*(1).

Waring, H. (2013). Managing Stacy: A case study of turn-taking in the language classroom. *System*, *41*, 841–851.

Waring, H. Z. (2009). Moving out of IRF (Initiation-Response-Feedback): A Single Case Analysis. *Language Learning*, *59*(4), 796–824.

Waring, H. Z. (2011). Learner initiatives and learning opportunities in the language classroom. *Classroom Discourse*, *2*(2), 201–218.

Wigham, C. R., & Satar, M. (2021). Multimodal (inter)action analysis of task instructions in language teaching via videoconferencing: A case study. *ReCALL*, *33*(3), 195–213.

Woods, D, (2021). Transana v4.00. https://www.transana.com. Madison, WI: Spurgeon Woods LLC.

Yousefi, S. (2014). Comparison of traditional and video mediated learning of English: Tracking a new approach. *Procedia-Social and Behavioral Sciences*, *98*, 1940–1944.

Zimmerman, D. H., & West, C. (1975). *Sex roles, interruptions and silences in conversation Language and sex: Difference and dominance.*Stanford, CA: Stanford University Press, 105-129.

Appendices

Appendix A: Transcription Conventions

Jefferson Transcription Conventions Adapted from (Jenks, 2011)

[ ]	Overlapping utterances – (beginning [) and (end])
=	Contiguous utterances (or continuation of the same turn)
(0.4)	Represent the tenths of a second between utterances
(.)	Represents a micro-pause (1 tenth of a second or less)
:	Elongation (more colons demonstrate longer stretches of sound)
.	Fall in pitch at the end of an utterance
-	An abrupt stop in articulation
?	Rising in pitch at utterance end (not necessarily a question)
CAPITAL	Loud/forte speech
__	Underline letters/words indicate accentuation
↑↓	Marked upstep/downstep in intonation
° °	Surrounds talk that is quieter
hhh	Exhalations
.hhh	Inhalations
he or ha	Laugh particle
(hhh)	Laughter within a word (can also represent audible aspirations)
> <	Surrounds talk that is spoken faster
< >	Surrounds talk that is spoken slower
(( ))	Analyst notes
( )	Approximations of what is heard
$ $	Surrounds 'smile' voice

Transcription Conventions for Multimodal Conduct (Mondada, 2018)

* *	Descriptions of embodied actions are delimited between
Δ Δ	that are synchronized with correspondent stretches of talk or time indications.
+ +	two identical symbols (one symbol per participant and per type of action)
*--->	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 in small caps in the margin.
fig	The exact moment at which a screen shot has been taken
#	is indicated with a sign (#) showing its position within the turn/a time measure.

Additional Conventions Adapted From (Balaman & Doehler, 2022)

Illustrations	Current screen of the participants who perform the screen-based activities
Circles	Points on the screen where the participants either click or hold the cursor still
Arrow	Direction of the cursor movements within the screen-based activity illustrations
Descriptions	Unanalytical descriptions of the illustrated screen-based activities

Other Additions Created in This Study

	Audio activation by the participant
	Audio deactivation by the participant
	Mouse clicks by the teacher
	Fast nodding
	Slow nodding

Appendix B: Information sheet (Teachers)

Newcastle University

School of Education, Communication & Language Sciences

Dear Participant (Teacher),

You are invited to take part in a research study entitled:

The Role of the Teacher in Managing Collaboration in Synchronous Online Collaborative Language Learning: A Multimodal Analysis

- Please read this document carefully and ask any questions you may have before agreeing to take part in the study.
- The study is conducted by **Ali Mohammad A Alghamdi** as part of his Integrated PhD in Educational and Applied Linguistics at Newcastle University.
- This research project is supervised by Dr Muge Satar, Dr Spencer Hazel and Dr Peter Sercombe from the School of Education, Communication & Language Sciences at Newcastle University.
- The purpose of this study is to research is to examine English language teachers' interaction in online collaborative language learning sessions.
- You have been invited to take part in this study because of your expertise in teaching English language in online settings.
- If you agree to take part in this study, you will be asked to use a screen recorder software to video record your online English language sessions, record your online lessons via zoom using its recording function, and share the videos with the researcher.
- It is your right to be informed of the research findings once the study is complete. If you would like to receive a summary of the findings, please email the researcher.
- You are free to decide whether or not to participate. If you decide to participate, you are free to withdraw at any time without any negative consequences for you. In case of withdrawal, you can choose to grant and deny permission for the researcher to use the data collected up to the point of the withdrawal or not.
- All responses you give or other data collected will be kept anonymous and confidential. The records of this study will be kept secure and private on the University servers. All files containing any information you give will be password protected and/or locked on the university servers. In any research report that may be published, no information will be included that will make it possible to identify you individually. There will be no way to connect your name to your responses at any time during or after the study. All the videos recorded shall be securely stored by the researcher and will only be used for research purposes.
- All the data gathered for this study will be used for research purposes only in this study, data sessions, conference presentations and future publications.
- If you have any questions, requests or concerns regarding this research, please contact me

This study has been reviewed and approved by the School of Education, Communication & Language Sciences Ethics Committee at Newcastle University (date of approval:20/08/2019)

Faithfully yours,

Ali Mohammad A Alghamdi

Appendix C: Information Sheet (Students)

Newcastle University

School of Education, Communication & Language Sciences

Dear Participant (Student),

You are invited to take part in a research study entitled:

The Role of the Teacher in Managing Collaboration in Synchronous Online Collaborative Language Learning: A Multimodal Analysis

- Please read this document carefully and ask any questions you may have before agreeing to take part in the study.
- The study is conducted by **Ali Mohammad A Alghamdi** as part of his Integrated PhD in Educational and Applied Linguistics at Newcastle University.
- This research project is supervised by Dr Muge Satar, Dr Spencer Hazel and Dr Peter Sercombe from the School of Education, Communication & Language Sciences at Newcastle University.
- The purpose of this study is to research is to examine English language teachers' interaction in online collaborative language learning sessions.
- You have been invited to take part in this study because of your interest in learning English language in online settings and because you are a student in a teacher's class who is willing to participate in this study.
- If you agree to take part in this study, you will be asked to take screenshots of your screen, and agree to the use of the Zoom recordings of your online lessons for this study.
- All the videos recorded shall be securely stored by the researcher and will only be used for research purposes, future publications, conference presentations and data sessions.
- It is your right to be informed of the research findings once the study is complete. The main findings will be sent by email or short meetings online.
- You are free to decide whether or not to participate. If you decide to participate, you are free to withdraw at any time without any negative consequences for you. In case of withdrawal, you can choose to grant permission for the researcher to use the data collected up to the point of the withdrawal or not.
- All responses you give or other data collected will be kept anonymous and confidential. The records of this study will be kept secure and private on the University servers. All files containing any information you give will be password protected and/or locked on the university servers. In any research report that may be published, no information will be included that will make it possible to identify you individually. There will be no way to connect your name to your responses at any time during or after the study.
- All the data gathered for this study will be used for research purposes only in this study, data sessions, conference presentations and future publications.
- If you have any questions, requests or concerns regarding this research, please contact me via email

This study has been reviewed and approved by the School of Education, Communication & Language Sciences Ethics Committee at Newcastle University (date of approval20/08/2019)

Faithfully yours,

Ali Mohammad A Alghamdi

Appendix D: Informed Consent

Newcastle University

School of Education, Communication & Language Sciences

Declaration of Informed Consent

- I agree to participate in this study, the purpose of which is to observe teachers' interaction in online collaborative language learning.
- I have read the participant information sheet and understand the information provided.
- I have been informed that I may decline to participate or withdraw from the study at any point without penalty of any kind.
- I have been informed about the types of data, including personal data that the researcher will elicit from me and for which purposes these data will be used. The lawful basis for processing my personal data is consent.
- I have been informed that data collection will involve the use of recording software.
- I have been informed that all of my responses will be kept confidential and secure, and that I will not be identified in any report or other publication resulting from this research.
- I have been informed that the investigator will answer any questions regarding the study and its procedures. The investigator's email is a.alghamdi2@newcastle.ac.uk and they can be contacted via email or by telephone on
- I will be provided with a copy of this form for my records.

- **Please complete the following (Circle Yes or No for each question):**

I give permission for extracts of the recordings (moving images, i.e. videos) from the online exchanges where I appear to be used in academic presentations and publications:

With my face in clear but with all other identity markers (e.g. username) masked out Yes No

With my face blurred and all other identity markers masked out Yes No

I give permission for snapshots of the recordings (still images) from the online lessons where I appear to be used in academic presentations and publications:

With my face in clear but with all other identity markers masked out Yes No

With my face blurred and all other identity markers masked out Yes No

I give permission for audio in the videos when I appear to be used in academic presentations and publications:

With my voice in clear but other identity markers masked out Yes No

With my voice altered/ muted and only use subtitles Yes No

Any concerns about this study should be addressed to the School of Education, Communication & Language Sciences Ethics Committee, Newcastle University via email to ecls.researchteam@newcastle.ac.uk

_____

Date Participant Name (please print) Participant Signature

I certify that I have presented the above information to the participant and secured his or her consent.

20/08/2019 Ali Mohammad Alghamdi

Date Signature of Investigator