Articles

Managing Turn-Taking and Student Response Through a Microphone Gesture in an EFL Classroom

Aya Watanabe
Fukui Prefectural University

Despite the growing interest in examining the roles of multimodal practices in L2 interaction and language learning (Hall & Looney, 2019; Jacknick, 2021; Lilja, 2022), few studies have been conducted on tracking down teacher's use of recurrent embodied practices utilized in an educational setting over lessons and how students orient to it. This study examines a teacher's systematic use of a specific gesture and embodiment through closely observing classroom interactions between an experienced EFL teacher and young learners in Japan. The analysis focuses on a recurrent hand gesture, which will be termed as *a microphone gesture*, that is utilized mainly as an interactional resource to allocate turns and moderate speaker shifts. The aim of the study is twofold: a) to describe the orderliness of the embodied practice employed by the teacher in terms of managing turn-taking and b) to show how the gesture is used to achieve pedagogical goals.

　教室会話におけるマルチモーダルな実践の記述への関心が高まっているにも関わらず　(Hall & Looney, 2019, Jacknick, 2021, Lilja, 2022)、教育現場で教師が授業中に使用するジェスチャーを追跡し、学習者がそれに対してどのように志向しているかについての研究はこれまであまり行われていない。そこで、本稿では教師と生徒間のやりとりを詳細に分析することで、教師がマルチモーダル実践を体系的に使用していることを検証する。特に、本教室で繰り返し使われるハンドジェスチャー：マイクロフォン・ジェスチャーに焦点を当て、話者の順番交替を調整するための相互作用的資源として、どのように利用されているかを分析する。特に、a) 相互行為における順

番交替の観点から、教師が採用する身体的実践の秩序性を記述すること、b) 同時にジェスチャーがどのような教育目的を達成しているのかを明らかにすることを目的とする。

Keywords: EFL classroom interaction; gesture; multimodal conversation analysis; turn-taking; young learners

How teachers and students use their body movements in conjunction with their utterances in classroom interaction has been of interest among researchers. Previous studies utilizing conversation analysis (CA) have revealed that embodied actions, such as *gaze direction, body posture, head nods,* and *hand gestures*, play an important role in terms of organizing turn-taking and turn-allocation in classrooms. For instance, it has been reported that teachers use *gaze*, *head nods*, and *pointing* to nominate students as next speakers (Kääntä, 2012; Mortensen, 2008, 2009; Sert, 2015) along with personal address terms (Lerner, 2003) and students display their willingness to be selected as the next speaker through a *hand raise* and *mutual gaze* (Mortensen, 2009). Thus far, the value of using gestures in second language (L2) instruction has been recognized by researchers studying L2 use and teaching (Allen, 2000; Stam & McCafferty, 2008). Studies reveal that gestures are used by L2 teachers to increase comprehensible input (Allen, 2000), explain vocabulary (Lazaraton, 2004) and provide corrective feedback (Taleghani-Nikazm, 2008) to L2 learners. However, not many studies have focused on how a specific gesture gets recurrently employed by a teacher as a resource to manage turn-taking and achieve pedagogical functions for teaching young learners over time. Eskildsen and Wagner (2013) focused on how a teacher reused a shared gesture to elicit particular vocabulary in an adult ESL classroom. Tozlu Kılıç and Balaman (2023) also reported on teachers' repeated use of a target expression combined with a gesture which served to visually scaffold L2 learning for very young learners. The present study contributes to this line of research by examining a recurrently used hand gesture over time in the context of a teacher-fronted classroom interaction.

When humans interact with one another, they coordinate not only the way they talk but also their body movements along with their utterances (Goodwin, 2000; see also Atkinson, et al., 2007). These bodily movements play an important role as key interactional resources with which participants make sense of each other. How embodied actions act as interactional resources to achieve both professional and mundane social activities remains a relevant topic of investigation among scholars who are interested

in analyzing multimodality in human interaction (Deppermann, 2013; Heath & Luff, 2013; Streeck et al., 2011). Previous studies have revealed that bodily actions are organized in an orderly manner to achieve specific interactional goals (Sacks & Schegloff, 2002). Focusing on the use of a specific gesture in work meetings, Mondada (2007) reports how pointing gestures were used as a method to display a shift in speakership and project the emergence of possible next speakers. Describing how participants' bodily conduct gets employed in a systematic and recognizable way to achieve certain interactional and institutional aims is still an open-ended question that poses questions such as the following. *How does a specific embodied action get formulated and utilized as an interactional resource? When is it utilized and what does it achieve? How do the participants recognize and display or not display their understanding of the embodied action in the subsequent turn?*

In this article, I examine a classroom interaction to illustrate how an experienced English as a foreign language (EFL) instructor manages turn-taking and distributes speakership to possible next speakers as she interacts with her nine young learners. It focuses on a recurrently observed hand gesture which will be referred to as *the microphone gesture* (see figure 1).[1] and how it gets utilized over time.

Figure 1
The microphone gesture being used in the focal classroom

Literature Review
Studies on Turn-taking and Next Speaker Selection in Classroom

Turn-taking practices in conversations have been closely observed and examined. When two people talk on the phone, for instance, the order of turn-taking and speakership shift are rather simple. One person takes the role of the speaker and the other becomes the listener. When the speaker

signals the completion of a turn, for instance with a turn-final falling intonation or the completion of a word, phrase, or sentence, the listener projects the completion of the turn and is expected to take the next turn at a transition relevance place (TRP) (Sacks et al., 1974). In their seminal paper, Sacks et al. (1974) laid out the basic rules for the turn-taking organization and revealed the systematicity observed in ordinary conversation. The rules of who takes the turn, how long each turn takes, and when each will take a turn, are not predetermined, but instead locally managed by the participants themselves through careful monitoring and projection. When the number of people involved in the interaction increases and the interaction takes place face-to-face in ordinary situations or institutional contexts, naturally, the turn-taking organization and speaker shifts become more complicated.

Turn-taking organization observed in multiparty classroom interactions has often been studied based on participants' verbal conduct and has been described as unequally distributed among participants (Gardner, 2013; Markee, 2000, 2015; Seedhouse, 2004). In teacher-fronted classrooms, the teacher is in essence the only one who is entitled to allocate turns and select the next speaker, which is often a student, and the nominated speaker can only select the teacher after they complete their turn (Mehan, 1979; McHoul, 1978). Even if the teacher is not selected by the student, the teacher can continue to self-select, and the process gets repeated. Basically, as McHoul (1978) states, "only teachers can direct speakership in any creative way" (p. 188). This unequal distribution of turns and teacher control of turn-taking are a reflection of the asymmetric nature of knowledge and the difference in the social roles and expectations designated to teachers and students (Drew & Heritage, 1992). The teacher is expected to be knowledgeable about the content and carries the responsibility to assess students of their performances, which is reflected in the triadic dialogue known as the IRE sequence (Mehan, 1979). Thus, this characteristic is reflected in the turn-taking organization and sequence organization in classroom and makes the classroom interaction different to other institutional interactions and ordinary conversations.

The turn-taking organization in classroom is influenced by multiple participants and the use of embodied actions. Speaker shift and next-speaker selection in multiparty interaction can be a complicated act to be managed and negotiated verbally and nonverbally (Hayashi, 2013). Especially relevant for this study are interactions where the current speaker selects the next. In ordinary multiparty conversation, Lerner (2003) reported that next-speaker selection occurred through the current speaker

explicitly addressing a specific speaker by *gaze* and personal address terms, or else tacitly addressing them by formulating a turn specifically designed for an individual. Stivers and Rossano (2010) further reported that speakers design their actions to increase the accountability and relevance of a coparticipant's response by simultaneously combining multiple turn-design features like interrogative prosody, sensitivity to recipient's epistemic domain, and speaker *gaze*. Recent studies investigating multimodal aspects of turn-taking in teacher-fronted classrooms show that teachers' *gaze*, *body orientations*, *pointing*, and *head nods* are utilized to allocate response turns to students (Kääntä, 2012; Mortensen, 2008, 2009; Sert, 2015). Examining the context of Content and Language Integrated Learning (CLIL) classrooms in Finland, Kääntä (2012) reported how teachers employ *gaze* and *pointing* to select the next speaker, and how students use *gaze* and *hand raising* after teacher-initiated questions to bid for the next turn. She points out that to accomplish smooth speaker shift, it is important that the teacher and the potential next speaker establish mutual gaze and clearly see each other's visual and bodily conduct.

Furthermore, the mechanisms of speaker shift and multimodal methods of speaker selection in classroom have been reported to be a collaborative act between teachers and students. Based on the data from Danish L2 classroom, Mortensen (2009) demonstrated that before teachers select a specific student to be the next speaker, they *gaze* towards the students in order to find a participant willing to answer a teacher-initiated question. Students willing to be selected as the next speaker display their willingness to take the next turn through *hand raising* or/and *gazing* towards the teacher. Similarly, Lauzon and Berger's (2015) study revealed that students play a significant role in locally managing their participation by displaying availability/ unavailability to respond to teacher initiations in French L2 classrooms in Switzerland. Thus, both teachers and students systematically employ talk and embodied action as interactionally meaningful resources to negotiate and manage turn-taking. These studies challenge the traditional notion of dominance in the teacher's role of controlling the classroom participation by revealing that speaker selection is in fact, jointly accomplished by all participants (Lauzon & Berger, 2015) through collaborative adjustments and orderly use of multimodal resources. In sum, it is useful to focus on the nonverbal resources utilized by both teachers and students when observing and analyzing turn-taking in classroom interaction.

Studies on Gestures and the Focal Gesture: The Microphone Gesture

Gestures are visible actions that are "used as an utterance or as a part of an utterance" (Kendon, 2004, p. 7) indicated through a movement of our body, especially through our hands. Gesture scholars have classified gestures into various categories and Applied Linguists have used these categories to investigate teacher gestures observed in classrooms. For instance, investigating a teacher's nonverbal actions in a foreign language classroom, Allen (2000) adopts Burgoon et al.'s (1989) five categories. They are *emblems*, *illustrators*, *affect displays*, *regulators*, and *adaptors*. *Emblems* are symbolic body movements specific to a culture, e.g., a thumbs-up gesture. *Illustrators* are movements that illustrate the utterance. *Affect displays* are facial expressions which displays emotions. *Regulators* are body movements that manage and regulate the flow of speaking and listening between two or more interactants. *Adaptors* are movements that individuals perform to satisfy their physical or psychological needs. Among these five categories, the most relevant category for the focal gesture of this paper, *microphone gesture*, is the fourth category, *regulators*. Analyzing Spanish as a foreign language classroom interaction, Allen (2000) reported that the teacher used gestures categorized as *regulators* to maintain turn-taking and have students repeat after her or continue to talk. Eliciting repetition, managing turns, and pursuing student utterances were some of the functions achieved by using the *microphone gesture*. Thus, under this categorization, the microphone gesture can be classified as a *regulator*.

Another influential categorization comes from a seminal work by McNeill (1992). He offers four categories: *iconic gestures*, *metaphoric gestures*, *deictic gestures*, and *beat gestures*. First, *iconic gestures* are gestures that depict the content of talk and represents both objects and bodily actions. *Iconic gestures* can be further categorized as *kinetographic*, if the gesture depicts bodily movements, or *pictographic*, if the gesture represents the actual form of an object. Second, *Metaphoric gestures*, in contrast, are gestures that describe abstract concepts. *Deictic gestures* are pointing gestures that indicate something specific in the environment or abstract concept. Lastly, *beat gestures* are hand moves, like a flick of the hand, which has a rhythmical pulse that goes along with the speech. Based on this classification, the *microphone gesture* belongs to a type of *iconic gesture*, specifically, *pictographic gesture*, as it represents an actual form of an object. Thus, this gesture can be understood to carry the functional characteristic of the represented object.

The *microphone gesture* is not commonly seen or used in our daily lives; however, the use of a microphone as a tool to address public audiences at large events, such as concert halls, to record sounds electronically at musical or political gatherings, or to interview people for a television show (Ponomareva, 2011) is commonly observed. According to Ponomareva (2011) who studied media interviews, a microphone served as a transmission device and fulfilled the role to determine the turn-taking through turn-initiation, continuation of the turn, and turn ending. Specifically, speaker nomination occurred through the microphone transition operated by the interviewer, who managed the course of interview. Thus, possessing the microphone signaled such a transactional state and had strong impact on the way interactants organized institutional talk. When a microphone itself is absent from the context and is substituted by a hand gesture, and if the gesture is recognized by participants as a representation of a microphone, the gesture can serve similar interactional functions as the microphone, for example, signaling a transactional state and thereby becoming a turn organizational resource. Gestures that substitute tools with specific functions can serve as an interactional resource and achieve similar interactional goals as the object. For instance, Mortensen's (2016) study on *cupping the hand behind the ear* gesture, once described as a "non-electric aid to hearing" (Stephens & Goodwin, 1984, p. 215), revealed that the participants treated the gesture as a display of a problem in hearing and this served as an interactional resource to achieve other-initiation of repair utilized by the teacher. Studying the same gesture, Amar (2022) showed how teachers used this gesture to pursue students' response when they fail to provide an answer in a timely manner.

When analyzing how gestures get deployed in interaction, it is important to attend to the timing of its appearance as well as the overall structure and phases of how they develop. The whole gesturing from the moment the limb begins to move and ends when the limb returns to the original position is referred to as a *gesture unit* (McNeill, 1992). The *gesture unit* consists of one or more *gesture phases* which include a *preparation*, a *pre-stroke hold*, the *stroke*, a *post-stroke hold*, and the *retraction* (Kendon, 2004; McNeill, 1992). Among the different phases, the *preparation*, *pre-stroke hold*, *post-stroke hold*, and *retraction* can be optional but the *stroke* marks the peak of the movement that expresses the meaning of the gesture, thus is obligatory. Previous studies have documented how gestures often get deployed and developed over turns with temporal progressivity (Kendon, 2004; Mondada, 2007; Sikveland & Ogden, 2012). In terms of bodily movements in naturally

occurring interaction, it has also been reported that a "very large number of moves and sequences of moves in interaction end where they begin" (Sacks & Schegloff, 2002, p. 137). This is also known as the "home position" (Sacks & Schegloff, 2002). Therefore, it is important to pay attention to the timing of when and where the gesture begins and ends, as well as how these stages develop in coordination with the talk. The *microphone gesture* gets deployed and developed through a series of phases: (a) the preparation – making a fist with either the right or left hand, departing from the original position; (b) the peak – placing the fist in front of a specific student; and (c) the retraction – withdrawing the fist back to the original *home position*. Moreover, the microphone gesture often co-occurred with the teacher's *gaze* towards a specific participant. Thus, the coordination of gaze direction, the temporal development, the position of body, and the verbal production should be taken into account when the gesture is being analyzed.

Eliciting Repetitions and Eliciting Responses in Language Classrooms

Eliciting repetitions from young novice L2 learners is pervasive in language classrooms (Chaudron, 1988; Duff, 2000; Kanagy, 1999). Repetition can be defined as "the act of copying or reproducing verbal or nonverbal behavior produced by self or other in communicative situations" (Piirainen-Marsh & Alanen, 2012, p. 2825). Repetitions in language classrooms can be observed in various forms, for instance, repetitions of pronunciation, prosody, vocabulary or phrase, grammatical features, and even nonverbal actions. Teachers use repetitions to provide corrective feedback on student utterances (Chaudron, 1988), to provide uptake and draw students' attention to a specific form, and to encourage students to become engaged in interaction (Duff, 2000). For learners, repetition is beneficial because it allows them to hear and practice problematic turns, and to join with other classmates in the common activity of learning. Therefore, eliciting repetitions play an important role in language classrooms.

Eliciting responses to teacher-initiated questions is also a common practice conducted by language teachers. When teachers initiate questions, and if there are no response in the next turn, the silence becomes noticeable and needs to be addressed. With the absence of response, the teacher moves on to prompt students to fulfill their obligations to answer the question. Under these circumstances, teachers use various techniques to elicit responses from learners. For instance, teachers might use designedly incomplete utterance (Koshik, 2002) with a rising intonation to mark the

absence and provide hints to students (Sert & Walsh, 2013). Moreover, teachers may use gestures to encourage and facilitate student response. Sert (2015) calls this an "embodied elicitation" (p. 102) and argues that teachers good use of combining gesture and speech has the potential to facilitate learners displays of understanding and lead to language learning. Based on data from a first-year EFL classroom in a Japanese university, Amar (2022) reported how teachers used the ear cupping gesture to pursue students' response when an answer was inapposite or not provided in a timely manner. Eliciting responses or pursuing a response is widely observed in language classrooms with learners of various levels of linguistic ability. Building on the prior research, this study aims to describe how an experienced teacher recurrently employs the microphone gesture to manage turn-taking and elicit student responses from young EFL learners. Moreover, it aims to show how the microphone gesture achieves different pedagogical goals as it gets recurrently employed over different lessons.

The Data and Method

The study is based on approximately 450 minutes of video recordings of EFL classroom interactions at an after-school English program in Japan. The participants are an experienced EFL instructor with over 30 years of teaching experience and nine young learners attending a lesson held once a week (one lesson consists of 60 minutes). All students in this class were aged 5 or 6 at the start of data collection and were beginning level learners of English with almost no experience being exposed to English prior to coming to this school. Based on several years of observation of this instructor's classes, the instructor followed an English only policy in all her classes from day one and uses a great deal of nonverbal resources. It was clear that the teachers' rich use of her body orientation, gaze, facial expressions, and frequent use of gestures served as clues for students to recognize patterns, understand and follow the activities. Thus, her classes were videorecorded over time and became the focus of investigation. This study has been reviewed and approved by the Committee of Human Studies at The University of Hawaii in June 2008. The participants and their parents were informed of the research in July 2008. The purpose of the research, expectations of the participants, their rights, and benefits were explained explicitly to the participants in Japanese. Consent forms, which includes the use of data collected in 2006, were signed by the students and their guardians in both Japanese and English.

The excerpts analyzed in the current study were taken from the following data sources: April 2006 (Excerpt 1), June 2006 (Excerpt 2), and June 2009

(Excerpt 3). These excerpts were selected because they contained the variety of use of the microphone gesture, although the data were collected in different time periods. The microphone gesture appeared in every class and was recurrently used from day one of the English lesson represented in Excerpt 1, which occurred 37 times in total. In the lesson recorded 2 months later (Excerpt 2), microphone gesture was observed 23 times. Lastly, in the lesson recorded 4 years later (Excerpt 3), the same gesture was utilized 15 times. Collections of the microphone gesture were made and analyzed to describe how the teacher utilized the gesture when addressing young learners. The data collected over lessons allowed the author to observe how this recurrent gesture was utilized and how it was received by the learners in different time periods.

The data were analyzed using conversation analysis (CA) with an aim to explicate the underlying mechanism of social interactions from the participants perspective. The data were transcribed using multimodal transcription conventions developed by Mondada (2018) (see Appendix 1 for the symbols designated for each speaker.). Thus, the detailed transcription includes gaze symbols and gesture designated to a specific participant to describe when a particular action is in preparation, reaching its apex, or under retraction. All the names of the participants in the transcripts were changed to pseudonyms.

Analysis

In the following section, I will describe three examples of how and when the *microphone gesture* is being utilized, as well as how the gesture is being oriented to by the young learners who are in the very early stages of learning English. The first excerpt come from the very first day of the English lessons when all the participants met in the classroom for the first time. The second excerpt comes from a lesson that occurred 2 months later, and the third excerpt 38 months later. Note that they consist of mostly same group of students being taught by the same teacher.

Eliciting Repetition in the Target Language

This segment occurred at the beginning of the very first lesson, when the teacher had just collected tape recorders from each student and put them on the table in front of her. The excerpt starts right after Eisaku's tape recorder has been placed on the table by the teacher (TEA) who says, *here's eisaku's*, (in line 1) while gazing at the other students. Towards the end of the turn,

the teacher gazes at a specific student, Shizuka (SHI), who is sitting next to Eisaku.[2]

Excerpt 1 Tape recorder [T1_3_4_Shizuka]

```
1  TEA    +here's eisaku'*s.+
   tea                    *…>
          ▪gazes at Ss▪  ▪gazes towards shi▪
   shi    +holding tape recorder, moves body to tea+
          †gazes at tea†
```



```
2         + (0.5) +
   tea    reaches RH towards shi-->
   shi    +holding tape recorder with both hands+
3  TEA    AH (.) this is ↑mi↓ne*
          takes the recorder, tries to put it in pocket*
          ▪smiles and gaze at other Ss▪
   shi    †gazes at tea†
4  SHI    *+chigau+
                 no
           +gaze and body leaning towards TEA+
   tea    *RH in pocket and moves RH towards…>
          ▪gaze at shi▪
```

```
5  TEA   →    no?
       ⇒    mic--->
```



```
   shi        †gazes at tea†
6  SHI        NO!
   tea        -->
7  TEA        ∘say∘  (0.3)  it's *mine
   tea        circular hand......*mic-->
8  SHI        it's +mi*ne
   shi        +puts RH to chest-->
   tea        -----,,,*
```


At the end of line 1, the teacher shifts her gaze from the whole class to focus specifically on an individual student to indicate a transition. While looking at the teacher, Shizuka, the gaze-selected student, holds her tape recorder with both hands and moves her body towards the teacher during the teacher's utterance. Shizuka's shift in posture, the movement of her body and her gaze towards the teacher can be understood as soliciting teacher attention (Cekaite, 2008), and in fact, results in establishment of mutual-gaze with the teacher. By holding her tape recorder towards the teacher, Shizuka indicates her orientation to the topic of the activity and attentiveness towards the teacher's actions. After the mutual gaze, a gap follows and the teacher reaches her right hand towards Shizuka's tape recorder (line 2).

In line 3, the teacher's loud *AH* token draws the participants' attention and acts as a display of noticing something worthy of reporting, as she

takes Shizuka's tape recorder with her right hand. After a micropause, the teacher claims the ownership of the object by stating, *this is mine*, with a rather exaggerated upward and downward intonation, and demonstrates her statement with the embodied action of trying to put the tape recorder in her pocket (note that the tape recorder is larger than her pocket). Her facial expression, the smile, and gaze towards other students, as well as the animated intonation and management of the object in line 3, indicate the playful nature of this statement and action. Furthermore, the series of actions: taking a student's tape recorder, claiming ownership of it, and trying to put it in her pocket, were performed in a previous exchange with Eisaku prior to this sequence. This repetition or establishment of multimodal routine could serve as a clue for other participants to recognize the similar pattern that the teacher might be performing and help them project or anticipate what might follow next (Kanagy, 1999; Watanabe, 2016).

In the next turn, Shizuka responds to the teacher's previous embodied conduct and a statement in line 3 by denying it using one word in her first language (L1) Japanese, *chigau*, meaning *no* or *wrong*, along with a gaze and leaning towards the teacher. As soon as Shizuka finishes the verbal utterance in line 4, the teacher moves her right hand from her pocket towards Shizuka simultaneously forming a fist (the preparation stage). In line 5, the teacher reformulates the previous Japanese utterance into the target language, English, with rising intonation, *no?* Co-occurring with this reformulation of the word, the teacher places her right fist position as if it were holding a microphone in front of Shizuka's mouth (the peak). Gazing back at the teacher, Shizuka repeats the reformulated word in a rather loud volume in line 6. Following this, the teacher continues to take the next turn by uttering a directive, ∘*say*∘ *it's mine*, with a circular hand motion (the preparation) and formulating another microphone gesture towards Shizuka (the peak). Shizuka then repeats the phrase, *it's mine*, with the embodied action of placing her right hand to her chest. Here it is important to mention that the sentence, *it's mine*, emerged in the previous interaction with Eisaku to claim his ownership of his tape recorder. Shizuka is displaying her understanding of this interaction through not only through repetition, but also with her embodied action. In line 8, the teacher reverses the microphone gesture (the retraction) as Shizuka completes the repetition and the teacher's right hand returns to the home position.

Given that this is the very first day of English class for these students, their knowledge of English is close to nil. As mentioned earlier, students in this excerpt were 5 to 6-year-old preschool students who probably had limited exposed to English prior to coming to this school. Allen (2000) who

studied nonverbal foreign language teacher talk, mentions the importance of nonverbal resources "especially for learners in lower-level classes who, because their knowledge of the language form is limited, rely on extra-linguistic cues to close the gaps in comprehension" (pp. 169-170). Although the teacher follows an English only policy in her lessons, her nonverbal cues like body movements, gaze, facial expressions, and frequent use of gestures serve as clues for students to recognize patterns, understand and follow the activity, and help produce what they are expected to do in the interaction. It is also important to note that all students are displaying their attentive participation in various ways. More specifically, the selected next speaker, Shizuka, also participates by using her bodily conduct, gaze, and repetition of teacher utterances, as prompted by the microphone gesture. The other students who are not producing utterances and being listeners are attentively monitoring the interaction with their gaze (Goodwin, 1980), and make public their understanding of their current role as ratified overhearers by not taking the next turn.

From this excerpt, we can see that the microphone gesture serves at least three purposes: gaining attention, allocating a turn, and achieving a pedagogical goal. First, the microphone gesture is produced to gain joint attention from the young learners. Studies have reported on how children solicit attention from teachers using artifacts and embodied actions (Cekaite, 2007, 2008), but it is equally important for the teachers to gain attention from students. The microphone gesture combined with the teacher utterance is effectively performed to obtain attention from all participants including the selected next speaker. Second, the microphone gesture serves to allocate a turn and establish a specific student as a next speaker through publicly displaying the selection. The deselected students continue to orient to this action by monitoring and gazing at the focal student and the teacher. Third, the teacher's pedagogical aim of having students repeat the phrase in the target language (line 7) and trying to have them use English as much as possible is enabled with the use of microphone gesture. In this excerpt, we can see that the simultaneous production of the microphone gesture and the teacher utterance serves as an indication to prompt a student to repeat the reformulated word and target phrase that is produced with the microphone. Moreover, the position of its placement, e.g., right after the student's L1 production (line 5) and the teacher's intra-turn pause (line 7), serve as indications of which words to repeat. Furthermore, the return of the gesture to the home position marks the completion of the pedagogical goal being achieved. The microphone gesture occurred 37 times in total in the first day of instruction. The limited space does not allow other examples

to be shown, however, the next section shows a summary of the sequential positioning of the microphone gesture used to elicit repetition.

Formalization to Achieve Repetition: Sequence Organization

In the early stages, especially in the first lesson, the microphone gesture frequently occurred as a means not only to seek a response, but also to prompt a repetition from students, who were selected as the next speaker. Previous studies on repetitions in FL classrooms have pointed out that teacher repetitions are pervasive in relatively low-level foreign language classroom interactions (Duff, 2000). As discussed earlier, teachers utilize repetition to provide uptake on student utterances and to encourage students to become engaged in interaction. In addition to increasing participation, repetition benefits learners by allowing them to practice articulating problematic turns. By providing a candidate response along with the microphone gesture, the teacher prompts students to use the target word, without specifically verbalizing "repeat after me." Table 1 is the formalization of the phenomenon which describes speaker shifts, turns, embodied actions, and actions the turn achieves in the sequence.

Table 1

Formalization of the Microphone Gesture to Achieve Repetition

	Teacher (TEA), Selected Student (S) talk	Action	TEA's Gesture and embodied action	S's embodied action
Turn 1	TEA: Teacher Initiation	Initiates sequence		Displays recipiency (gaze)
Turn 2	S: Student Response in L1	Responds to TEA's initiation	Orients to student's contribution (gaze)	Answers in non-verbal action (nod)
Turn 3	TEA: Candidate Response	Provides candidate response in L2	The Microphone Gesture	Displays recipiency (gaze)
Turn 4	S: Repetition	Repeats candidate response	Microphone (Turn-final retraction)	Displays recipiency

In many cases, the microphone gesture serves the interactional purposes of obtaining attention from young learners, allocating a turn to a specific speaker, mobilizing a response (Stivers & Rossano, 2010), and prompting repetition of the candidate response produced by the teacher. In addition to this, there are other functions the microphone gesture carries out to achieve pedagogical purposes. Next, we will observe other pedagogical functions of the gesture in lessons that took place after the first lesson.

Legitimatizing Student's Turn

The following excerpt comes from a lesson that occurred about 2 months after the previous excerpt. Here, the teacher is reading a picture book entitled *A Beautiful Butterfly* (Nakamoto, 2011) to the students and invites them to participate by asking questions based on the pages of the picture book. The teacher uses two types of voices: a regular voice and an animated voice. The animated voice is recognizable and is differentiated from the regular voice in the transcript by embedding the talk in at-marks (@). The animated voice is produced when the teacher is reading out loud from the book and enacting the main character's voice, which is that of a caterpillar (lines 1-7).

Excerpt 2 Something blue [T3_1_2_3_2.49]

```
01   TEA   @*I want to be a (0.3)  blue butterfly*@
            *p at  the  words in  the picture book*
02         (0.5)
03   TEA   tch  @I *have to eat*  (0.5)  something ↑blue::@
                 *eating gesture*
04         (0.7)
05   TEA   @some[thing blue:@]
               ▪gazes at Ss▪
06   EIS        [blueberry::]
07   TEA   @*bl↑ue, blue, [blue.*@
            *gazes and p at the picture book*
08   EIS                  [blueberry ai
09   TEA   *something ↑blue,*  >every*body<
            *snaps fingers twice*    *cupping ear--->
            ▪gazes at Ss▪
```

```
10   Ss      something blue* ((EIS not included))
     tea        ------------->*
11           (0.2)
12   TEA     what is something blue=
13   EIS     =*BULUE: BE[RRY*
             gazes at tea
     tea      *mic towards eis*
```



```
14   TEA                    [oh ↑blueberry is blue yes,
15           blueberry is bl- blue, blueberry is blue,
16           (0.4)
```

In lines 1, 3, and 5, the teacher is reading the book with an animated voice that represents the caterpillar looking for something blue to eat. While using the animated voice, pointing at the book, pausing, and gesturing to entertain and support the students' understanding, the teacher utters, *something blue*, in line 5. Eisaku, the focal student in this excerpt, gives the name of a blue fruit, *blueberry::*, in the middle of the teacher's turn (line 6), elongating the final vowel in overlap with her utterance. As observed in the following turn, this self-selected turn which displays his understanding and projection based on his attentive listening, does not receive any uptake and fails to obtain mutual gaze as the teacher continues to look at the other students. In the next turn, the teacher shifts her gaze and continues to read the book in an animated voice as she points to it (line 7). Again, during the teacher's turn, Eisaku overlaps his talk with the teacher's repetition in a second attempt to initiate a self-selected turn in line 8 which again results in no uptake. In line 9, the teacher repeats the key phrase, *something blue*, switching to a normal voice, along with a rhythmical finger snap and opens the floor to the whole group by shifting her gaze and addressing the students with, *everybody*, and employing the cupping ear gesture (Mortensen, 2016; Sert, 2015). In

this sequence, the cupping ear gesture serves as a prompt for repetition of the key phrase, something blue, which is responded to and achieved by the students (line 10), other than Eisaku. After the choral repetition and a gap, a question gets initiated by the teacher in line 12.

Immediately after the teacher's initiation in line 12, without a gap, Eisaku provides a response in line 13 in a clear and loud volume. Here it is interesting to note the timing of the teacher's microphone gesture. The gesture was produced at the same time as Eisaku produced the response and the peak was placed in front of him. The placement of the gesture and the timing of the response without a latch is almost as if it was planned. However, this is not surprising because the gesture is positioned after the two failed attempts of Eisaku taking self-selected turns that were not taken up by the teacher. The teacher's utterance overlaps in the middle of Eisaku's answer in line 14, which indicates that she can project what his answer is. By producing the "oh" token in the middle of a known answer (Hosoda, 2015), as well as taking up and repeating student's answer several times, the teacher is reinforcing the appropriateness of Eisaku's response produced at this particular point, i.e., after the teacher initiation. The teacher returns her microphone gesture to the home position at the end of the student utterance (the retraction) in line 13. It displays the completion of the expected action, which is producing a response turn at the right timing.

In comparison to Excerpt 1, the microphone gesture does not co-occur with the teacher's production of a candidate response and does not prompt a repetition of the teacher utterance. Instead, it serves to manage legitimate participation through official allocation of a turn. The gesture was produced right after the teacher's production of the teacher-initiated question in a normal voice and functioned to elicit a student response in the appropriate sequential position. Eisaku's self-selections in lines 6 and 8 were placed in the middle of the teacher's animated utterance, and as a result, interrupted the on-going activity of the book reading and did not receive any attention or uptake from the teacher. Although the response itself fulfilled the role of providing a valid answer, as we can observe from the positive treatment to the same answer received in line 14, the earlier attempts were not taken up due to their sequential positioning. Thus, in this excerpt, the microphone gesture did not function to prompt a repetition, but instead served to ratify the selected student as a legitimate speaker and allowed the student's turn to be officially included in the main interaction. The microphone gesture makes it visible for all participants, including the speaker, that turns must

be taken in a specific sequential position, that is after a teacher-initiated question was produced in a normal voice. In other words, the microphone gesture in this excerpt served as a means to distinguish non-legitimate self-selected turns from legitimate teacher-selected turns. The gesture served to display appropriate turn-taking rules in this particular context, a picture book reading, and carries out a disciplinary action without direct confrontation to maintain the progressivity of the talk.

Pursuing Expected Utterance

The final excerpt comes from a lesson recorded 38 months (about 3 years) after the first lesson. Prior to this excerpt, the students were reporting their weekend activities to the teacher, which is a routine interaction that is carried out in every lesson (Watanabe, 2016). This sequence comes after the completion of the reporting activity and opens with the teacher initiation of a subsequent activity related to the weekend report.

Excerpt 3.1 Who question [T9_4_Eisaku]

```
01 TEA      okay, *∞who question*
                  *raises RH*
   eis             ∞raises RH--->
02          (0.5)
   eis      ---->
03 TEA      who?  (. )   *hm hm hm*
                         *snaps fingers*
   eis      ------------------->
04          *(1.0)
   tea      *……->
   eis      ---->
05 TEA      okay (.) eisaku,∞
            #mic to eis----->
   eis      ------------,,,∞
```

```
    fig           #fig4
```



```
06            (1.5)
    tea       mic and a head nod-->
                       ▪gazes at eis▪
07 EIS        kanoko?*=
    tea       mic--,,*
08 TEA        =who no [no no. ]   *>you have to say,<*
                                    *hand to her mouth*
09 EIS              [◦ah no◦]
10 TEA        who [*stayed home*
                   *both hands moves for each word*
11 EIS             [who ]
12 EIS        ah
```

In line 1, the teacher starts with a positive acknowledgment, *okay*, and initiates a sequence by uttering a *who question*, while raising her right hand, a gesture that provides an embodied example of what the students should do to express their readiness to respond. Eisaku, the focal student in this excerpt, also raises his right hand to show his availability and willingness to be selected as the next speaker (Mortensen, 2009). This phrase itself does not function to initiate a specific action. However, from the participants' perspective (as demonstrated by Eisaku's hand raise) it is publicly available that this is a turn which can be treated as initiating an action that makes a response conditionally relevant. The routineness of this teacher initiation is also reflected in Eisaku and the teachers' simultaneous hand raising. While Eisaku continues to raise his hand in line 2, a 0.5 silence follows and the teacher produces an additional turn starting with an upwardly intoned, *who?*, a micropause and, *hm hm hm*, with snapping fingers to indicate that some words should follow. Snapping fingers instead of giving specific words

is a technique often used by this teacher and observed elsewhere across data. After a 1.0 second silence, the teacher allocates a turn to Eisaku by gaze selecting him, addressing him by name, and simultaneously placing the microphone gesture in front of him (the peak). He has been raising his hand throughout multiple teacher turns (lines 1-5) and retracts his hand raise as his name is spoken.

After the next speaker is selected by the teacher's embodied turn allocation, the microphone gesture is maintained in the same position (peak hold) throughout a rather long silence. Eisaku then takes the next turn by producing the name of one of the students with a rising intonation (line 7). The microphone gesture is reversed soon after Eisaku's utterance (the retraction) and the teacher immediately produces, *who no no no*, repeating a negation three times, and treating the previous turn as inadequate. After this, she displays an explicit instruction using, *you have to say*, accompanied with a hand in front of her mouth to emphasize the verb. It is interesting to point out that the microphone gesture retraction served to indicate a shift and the end of the speakership. Overlapping with the teacher's negation, Eisaku quietly produces an ah token and a negation which aligns with the teacher utterance (line 9). The teacher then goes on to provide an example of the expected utterance, producing an interrogative sentence starting with *who* followed by a past tense verb (line 10). In overlap with the model question, Eisaku produces the repeated key word, *who*, and utters *ah*, after the teacher's model question is completed. It has been reported that Japanese a-prefaced response tokens display a change of state and receipt the received information as new (Endo, 2018). Considering that Eisaku is a native speaker of Japanese, he might be displaying a change of state (lines 9 & 12) to show his updated understanding of what he is expected to do by receipting the negative assessment and explicit instruction of the teacher.

In Excerpt 3.1, the use of the microphone gesture not only allocates a turn to pursue a response from the recipient, but also signals a speaker shift by its removal. In order to achieve the teacher's pedagogical goal, that is to have the students produce a question starting with who, the microphone gesture also functioned to show whether the student's production was in line with the teacher's expectation. In this excerpt, when the microphone is utilized, students are not expected to produce a repetition of the teacher utterance as in the first excerpt, but to produce and formulate a certain question pattern, i.e., who-initial question. The following is a continuation of the previous excerpt.

Excerpt 3.2 Who question [T9_4_Eisaku]

```
13 TEA     okay, >everybody< *who question
                             *…..#mic----->
   fig                            #fig 5
```



```
14 EIS     ah who:: did (2.3) who did- (2.0) nn:::
   tea     mic---------------------------------->
15 EIS     who::::: [did
   tea     mic drops-->
16 TEA             [did (English) no?
                     mic-------------->
17 EIS     see-  see::
   tea     mic------->
18         (0.7)
19 EIS     basket[ball
   tea     mic------->
20 TEA           [who went to
                   mic------->
21 EIS     who went to s-[ see a basket [ball ga-
   tea     mic---------------------------------->
22 TEA                   [see
                           mic-->
23 TEA                                  [ball game
                                         mic------>
24 EIS     game*
   tea     mic,*
   kan     raises hand
   ken     raises hand
25 TEA     oh who went to see a basketball game
```

In line 13 the teacher again initiates a sequence as she gazes at students and renews an action that is similar to the one in line 1. Although the teacher seems to invite the whole class with *everybody*, the teacher raises the microphone gesture and is placed in front of Eisaku towards the end of the teacher's turn. Given a second chance, Eisaku then responds with elongation and intra-turn pauses to formulate a question starting with *who* in line 14. Instead of using the grammatical format that the teacher provided, which is treating who as a subject and the verb following it, i.e., who (subject) + verb, Eisaku formulates a question using who and the auxiliary verb *did*. (The use of this grammatical form could be because the students are used to making question forms using the auxiliary verb did.) With the help of the teacher's clues (McHoul, 1990) accompanied with the embodied microphone cue, from lines 15 through 24, the teacher and Eisaku collaboratively formulate the who question until the teacher sums it up in line 25 saying, *oh who went to see a basketball game*. Here, the use of an *oh* token to a known utterance is a form of positive assessment that encourages the students (Hosoda, 2015). Towards the end of the co-production of the who question (line 24) and as the teacher produces the question in one turn (line 25), student bids begin from two students already raising their hands to answer this question.

In these excerpts, the microphone gestures functioned to gain attention, allocate a turn, pursue a response and to achieve the pedagogical task of producing a specific question type: who-initial question. The microphone gesture combined with the teacher clues allowed the respondent to sustain the speakership until the task is completed. The microphone is not only used in the pursuit of response, but also to maintain the relevance of that speakership over multiple turns. The use of turn-moderating microphone gesture over stretches of turns enables the teacher to pursue responses to her questions, and to thus structure the interactional organization of the class. Furthermore, the students' ability to monitor and adjust to the teacher's verbal and non-verbal action and instruction is essential to participating and achieving the task appropriately.

Discussion and Conclusion

The study has shown that the teacher's recurrent use of a specific hand gesture, the microphone gesture, was not randomly produced, but is in fact utilized as a meaningful multimodal resource to achieve interactional as well as pedagogical goals. Interactionally, the microphone gesture was systematically and recurrently employed by the teacher to organize and modulate turn-taking and turn-allocation (Allen, 2000). Repeated use of

a particular gesture could often lead to recipients adopting the gesture in their talk, a phenomenon known as "return gestures" to remedy troubles in interaction (de Fornel, 1992; Eskildsen & Wagner, 2013). However, interestingly, this gesture was never used by the students throughout the data, which shows how the teacher demonstrated the right and responsibility to manage turn-taking and how learners oriented to this right (Kääntä, 2012; McHoul, 1978). Selecting a next speaker could be a complicated task in multiparty classroom interaction where many students are present, and the teacher attempts to distribute opportunities for every student to participate. The microphone gesture combined with gaze carried out a smooth shift in speakership in turn transitions (Kendrick et al., 2023) by publicly displaying an orientation to who the next speaker will be. Secondly, it also served to obtain attention from the other students and have them focus on the interaction. Keeping eight to ten young learners attentive and having them participate in classroom activities for 60 minutes is not an easy task. By fully utilizing the microphone gesture and other embodied conduct, the teacher was able to gain joint attention (Kidwell & Zimmerman, 2007) from the learners and have them focus on the selected interactant by monitoring the activity and displaying their understanding when being selected. Having peripheral participants focus and monitor the on-going activity could lead to peripheral learning and serve as an important device for language learning (Okada, 2010). Thirdly, the microphone gesture served to pursue response when the response to teacher-initiated first pair parts were missing or delayed. Together with the use of verbal production of the teacher initiations, as well as gaze and rising intonation, the microphone gesture made it relevant for a student response to be produced by a selected individual.

Pedagogically**,** the microphone gesture was employed by the teacher to achieve instructional goals to produce certain types of responses to teacher-initiated turns and manage participation. In Excerpt 1, the microphone gesture served to prompt a repetition of the teacher's utterance co-produced with the gesture after the student's L1 utterance (see Table 1). The microphone gesture functioned to obtain an L2 response from the students by having them repeat the teacher's reformulation and produce expected utterances with appropriate timing and linguistic forms. This sequential position was recurrently used by the teacher and helped to achieve the pedagogical goal of having students produce repetition. Second, the microphone gesture managed to legitimize participation (Excerpt 2). The student's self-selected turns are a display of learner initiative and demonstrate willingness to

participate. However, when they are overlapped with the teacher's utterance, or provide answers before the question is initiated, they become disruptive. The microphone gesture served to differentiate the self-selected turns by legitimizing only those turns it allocated. The last excerpt showed that the microphone gesture enabled the teacher to pursue expected utterance. The length of microphone gesture was managed by the teacher and the prolonged use of the gesture functioned to maintain the role of speakership for the selected speaker. In sum, the teacher's use of the microphone gesture provides an example of the human body being utilized as a resource for achieving both interactional and institutional goals at the same time.

Lastly, the study contributes to our understanding of how a specific embodied practice gets recurrently utilized and oriented to by participants in educational settings (Eskildsen & Wagner, 2015). By tracking a very specific gesture, the microphone gesture, the study reveals how turn-taking and speaker shift is organized and coordinated by focusing on embodied practices employed by a teacher and her students. How teachers use recurrent gestures and embodied conducts when teaching a foreign language and studying its effectiveness has a lot to offer when reflecting on teaching practices or training new teachers. As the aim of this study is not to make any generalizations but to describe the orderly and systematic use of this gesture, I hope it contributes to our understanding of how gestures and embodiments are used as a meaningful resource in classrooms. Future studies should continue to investigate the systematic use of recurrent embodied practices employed by participants to reveal the world of embodiment in language classrooms.

Notes

1. Here it is important to point out the idiosyncratic use of the microphone gesture by this particular teacher. In this study, I am not trying to generalize the use of the microphone gesture in all EFL classrooms, but to reveal the systematic use of the gesture in this classroom through analyzing how it is treated and oriented to in the interaction by the participants.
2. All names in the excerpts are pseudonyms.

Acknowledgments

Aya Watanabe is an Associate Professor at the Research Institute for Regional Economics at the Fukui Prefectural University. Her research interest includes applying conversation analysis to analyze and explore interactions in English as a Foreign Language classrooms, especially focusing on the use of embodied resources like gesture, gaze, and body posture. Her work has been published in *Classroom Discourse* and other academic journals.

Appendix

Appendix 1 is available from the online version of this article at https://jalt.org/main/jj.

References

Allen, L. Q. (2000). Nonverbal accommodations in foreign language teacher talk. *Applied Language Learning*, *11*(1), 155–176.

Amar, C. (2022). Ear cupping in EFL classroom interaction: An embodied means of pursuing students' response. *Text & Talk, 42*(6), 779–799. https://doi.org/10.1515/text-2020-0127

Atkinson, D., Churchill, E., Nishino, T., & Okada, H. (2007). Alignment and interaction in a sociocognitive approach to second language acquisition. *The Modern Language Journal*, *91*(2), 169–188. https://doi.org/j.1540-4781.2007.00539.x

Burgoon, J., Buller, D. B., & Woodall, W. G. (1989). *Nonverbal communication: The unspoken dialogue*. Harper & Row.

Cekaite, A. (2007). A child's development of interactional competence in a Swedish L2 classroom. *The Modern Language Journal*, *91*(1), 45–62. https://doi.org/10.1111/j.1540-4781.2007.00509.x

Cekaite, A. (2008). Soliciting teacher attention in an L2 classroom: Affect displays, classroom artefacts, and embodied action. *Applied Linguistics*, *30*(1), 26–48. https://doi.org/10.1093/applin/amm057

Chaudron, C. (1988). *Second language classrooms: Research on teaching and learning*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524469

De Fornel, M. (1992). The return gesture: Some remarks on context, inference and iconic gesture. In P. Auer & A. Di Luzio (Eds.), *The contextualization of language* (pp. 159–176). John Benjamins. https://doi.org/10.1075/pbns.22

Deppermann, A. (2013). Multimodal interaction from a conversation analytic perspective. *Journal of Pragmatics*, *46*(1), 1–7. https://doi.org/10.1016/j.pragma.2012.11.014

Drew, P., & Heritage, J. (1992). *Talk at work: Interaction in institutional settings*. Cambridge University Press.

Duff, P. (2000). Repetition in foreign language classroom interaction. In J. K. Hall & L. S. Verplaetse (Eds.), *Second and foreign language learning through classroom interaction* (pp. 109–138). Routledge.

Endo, T. (2018). The Japanese change-of-state tokens *a* and *aa* in responsive units. *Journal of Pragmatics, 123*, 151–166. https://doi.org/10.1016/j.pragma.2017.06.010

Eskildsen, S. W., & Wagner, J. (2013). Recurring and shared gestures in the L2 classroom: Resources for teaching and learning. *European Journal of Applied Linguistics*, *1*, 139–161. https://doi.org/10.1515/eujal-2013-0007

Eskildsen, S. W., & Wagner, J. (2015). Embodied L2 construction learning. *Language Learning*, *62*(5), 268–297. https://doi.org/10.1111/lang.12106

Gardner, R. (2013). Conversation analysis in the classroom. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 593–611). Wiley-Blackwell. https://doi.org/10.1002/9781118325001.ch29

Goodwin, C. (1980). Restarts, pauses, and the achievement of a state of mutual gaze at turn-beginning. *Sociological Inquiry*, *50*(3-4), 272–302. https://doi.org/10.1111/j.1475-682X.1980.tb00023.x

Goodwin, C. (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics*, *32*(10), 1489–1522. https://doi.org/10.1016/S0378-2166(99)00096-X

Hall, J. K., & Looney, S. D. (2019). *The embodied work of teaching*. Multilingual Matters. https://doi.org/10.21832/HALL5495

Hayashi, M. (2013). Turn allocation and turn sharing. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 167–190). Wiley-Blackwell. https://doi.org/10.1002/9781118325001.ch9

Heath, C., & Luff, P. (2013). Embodied action and organizational activity. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 283–307). Wiley-Blackwell. https://doi.org/10.1002/9781118325001.ch14

Hosoda, Y. (2015). Teacher deployment of 'oh' in known-answer question sequences. *Classroom Discourse*, *7*(1), 58–84. https://doi.org/10.1080/194630 14.2015.1099111

Jacknick, C. M. (2021). *Multimodal participation and engagement: Social interaction in the classroom*. Edinburgh University Press. https://doi.org/10.3366/edinbur gh/9781474455183.001.0001

Kääntä, L. (2012). Teachers' embodied allocations in institutional interaction. *Classroom Discourse*, *3*(2), 166–186. https://doi.org/10.1080/19463014.2012. 716624

Kanagy, R. (1999). Interactional routines as a mechanism for L2 acquisition and socialization in an immersion context. *Journal of Pragmatics*, *31*(11), 1467–1492. https://doi.org/10.1016/S0378-2166(98)00113-1

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge University Press. https://doi.org/10.1017/CBO9780511807572

Kendrick, K. H., Holler, J., & Levinson, S. C. (2023). Turn-taking in human face-to-face interaction is multimodal: Gaze direction and manual gestures aid the coordination of turn transitions. *Philosophical Transactions of the Royal Society B, 378*(1875), 20210473. https://doi.org/10.1098/rstb.2021.0473

Kidwell, M., & Zimmerman, D. H. (2007). Joint attention as action. *Journal of Pragmatics*, *39*(3), 592–611. https://doi.org/10.1016/j.pragma.2006.07.012

Koshik, I. (2002). Designedly incomplete utterances: A pedagogical practice for eliciting knowledge displays in error correction sequences. *Research on Language and Social Interaction*, *35*(3), 277–309. https://doi.org/10.1207/ S15327973RLSI3503_2

Lauzon, F. V., & Berger, E. (2015). The multimodal organization of speaker selection in classroom interaction. *Linguistics and Education*, *31*, 14–29. https://doi. org/10.1016/j.linged.2015.05.001

Lazaraton, A. (2004). Gesture and speech in the vocabulary explanations of one ESL teacher: A microanalytic inquiry. *Language Learning*, *54*(1), 79–117. https:// doi.org/10.1111/j.1467-9922.2004.00249.x

Lerner, G. H. (2003). Selecting next speaker: The context sensitive operation of a context-free organization. *Language in Society*, *32*(2), 177–201. https://doi. org/10.1017/S004740450332202X

Lilja, N. (2022). Multimodal conversation analysis as a method for studying second language use and learning in naturally occurring interaction. In G. Stam. & K. Urbanski (Eds.), *Gesture and multimodality in second language acquisition: A research guide* (pp. 204–228). Routledge. https://doi. org/10.4324/9781003100683

Markee, N. (2000). *Conversation analysis*. Routledge. https://doi.org/10.4324/9781410606471

Markee, N. (Ed.). (2015). *The handbook of classroom discourse and interaction*. Wiley Blackwell. https://doi.org/10.1002/9781118531242

Mehan, H. (1979). *Learning lessons: Social organization in the classroom*. Harvard University Press. https://doi.org/10.4159/harvard.9780674420106

McHoul, A. (1978). The organization of turns at formal talk in the classroom. *Language in Society*, *7*(2), 183–213. https://doi.org/10.1017/S0047404500005522

McHoul, A. (1990). The organization of repair in classroom talk. *Language in Society*, *19*(3), 349–377. https://doi.org/10.1017/S004740450001455X

McNeill, D. (1992). *Hand and Mind: What gestures reveal about thought*. University of Chicago Press.

Mondada, L. (2007). Multimodal resources for turn-taking: Pointing and the emergence of possible next speakers, *Discourse Studies*, *9*(2), 194–225. https://doi.org/10.1177/1461445607075346

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. https://doi.org/10.1080/08351813.2018.1413878

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. https://doi.org/10.1558/japl.v5i1.55

Mortensen, K. (2009). Establishing recipiency in pre-beginning position in the second language classroom. *Discourse Processes*, *46*(5), 491–515. https://doi.org/10.1080/01638530902959463

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. https://doi.org/10.1080/08351813.2016.1126450

Nakamoto, M. (2011). *A beautiful butterfly*. Apricot Publishing.

Okada, Y. (2010). Learning through peripheral participation in overheard/overseen talk in the language classroom. In T. Greer (Ed.), *Observing talk: Conversation analytic studies into second language interaction* (pp. 133–149). JALT.

Piirainen-Marsh, A., & Alanen, R. (2012). Repetition and imitation: Opportunities for learning. In N. M. Seel (Ed.). *Encyclopedia of the sciences of learning* (pp. 2825–2828). Springer. https://doi.org/10.1007/978-1-4419-1428-6_657

Ponomareva, Y. (2011). *Microphone transitions as a gestural practice in dyadic television interviews*. [Master's thesis, Linköping University]. Upsattser. https://www.uppsatser.se/uppsats/e449da6d2a/

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50*(4)*,* 696-735. https://doi.org/10.2307/412243

Sacks, H., & Schegloff, E. A. (2002). Home position. *Gesture*, *2*(2), 133–146. https://doi.org/10.1075/gest.2.2.02sac

Seedhouse, P. (2004). *The interactional architecture of the language classroom: A conversation analysis perspective*. Wiley-Blackwell.

Sert, O., & Walsh, S. (2013). The interactional management of claims of insufficient knowledge in English language classrooms, *Language and Education*, *27*(6), 542–565. https://doi.org/10.1080/09500782.2012.739174

Sert, O. (2015). *Social interaction and L2 classroom discourse*. Edinburgh University Press. https://doi.org/10.1515/9780748692651

Sikveland, R. O., & Ogden, R. A. (2012). Holding gestures across turns: Moments to generate shared understanding. *Gesture*, *12*(2), 166–199. https://doi.org/10.1075/gest.12.2.03sik

Stam, G., & McCafferty, S. G. (2008). Gesture studies and second language acquisition: A review. In S. G., McCafferty & G. Stam (Eds.), *Gesture: Second language acquisition and classroom research* (pp. 3–24). Routledge. https://doi.org/10.4324/9780203866993

Stephens, S. D. G., & Goodwin, J. C. (1984). Non-electric aids to hearing: A short history. *Audiology*, *23*, 215–240. https://doi.org/10.3109/00206098409072836

Stivers, T., & Rossano, F. (2010). Mobilizing response. *Research on Language and Social Interaction*, *43*(1), 3–31. https://doi.org/10.1080/08351810903471258

Streeck, J., Goodwin, C., & LeBaron C. (Eds.). (2011). *Embodied interaction: Language and body in the material world*. Cambridge University Press.

Taleghani-Nikazm, C. (2008). Gestures in foreign language classrooms: An empirical analysis of their organization and function. In M. Bowles., R. Foote., Perpinan, S., & R. Bhatt (Eds.). *Selected proceedings of the 2007 Second Language Research Forum* (229–238). Cascadilla Proceedings Project.

Tozlu Kılıç, E., & Balaman, U. (2023). Facilitating students' learning of a target construction through teacher interactional resources in EFL kindergarten classrooms. *TESOL Quarterly*, *57*(2), 656–669. https://doi.org/10.1002/tesq.3209

Watanabe, A. (2016). Engaging in an interactional routine in EFL classroom: The development of L2 interactional competence over time. *Novitas Royal*, *10*(1), 48–70.

Appendix 1

Transcription Conventions

[	The point where overlapping talk and/or gesture starts
]	The point where overlapping talk and/or gesture ends
(0.0)	length of silence in tenths of a second
(.)	micro-pause less than 2/10 of a second
underlining	relatively high pitch
CAPS	relatively high volume
::	lengthened syllable
–	cut-off; self-interruption
=	'latched' utterances
?/./,	rising/falling/continuing intonation respectively
!	animated tone, not necessarily an exclamation
( )	unintelligible stretch
(word)	transcriber's unsure hearings
> <	increase in tempo, as in a rush-through
< >	decrease in tempo
°°	a passage of talk quieter than the surrounding talk
↑	higher pitch begins
↓	lower pitch begins
@	animated voice

Multimodal Transcription Conventions (Adapted from Mondada, 2018)

* *	delimits gestures and actions done by the **Teacher (TEA)**
▪ ▪	gaze by the Teacher
+ +	delimits gestures and actions done by **Shizuka (SHI)**
† †	gaze by Shizuka
∞ ∞	delimits gestures and actions done by **Eisaku (EIS)**

Abbreviations

Ss: Students choral response   S: Unidentified student
TEA: Teacher   SHI, EIS: Identified student
RH/LH: Right hand / Left hand   p: Pointing