



Social Interaction. Video-Based Studies of Human Sociality.
2026 Vol. 9, Issue 2
ISSN: 2446-3620
DOI: 10.7146/si.v9i2.162868

Social Interaction

Video-Based Studies of Human Sociality

Mobilizing Students' Choral Responses in Chinese L2 Classrooms: The Temporal Coordination of Gesture, Gaze, and Successive Prompts

Yingsheng Liu

University of California, Los Angeles

Abstract

Drawing on video-recordings of Chinese L2 classroom interaction, this study investigates how teachers mobilize students' choral responses through the coordinated deployment of multimodal interactional resources, including materially anchored pointing gestures, gaze reorientation, and successive prompting. Fine-grained multimodal Conversation Analysis (CA) reveals that these multimodal resources perform analytically distinguishable yet interactionally complementary functions in the mobilization of choral responses. Materially anchored pointing gestures functions to project and delimit the content of anticipated responses; gaze reorientation signals the appropriate moment for collective entry; and successive prompts together with teachers' reformulations serves to establish recognizable response formats when initial attempts at mobilizing whole-class participation prove unsuccessful. These findings highlight that the successful accomplishment of choral responses depends not merely on linguistic projection but on the dynamic coordination of verbal, embodied, material, and sequential resources that jointly organize both the projectability and the temporality of collective participation. This study also extends research on Classroom Interactional Competence (CIC) by showing that whole-class participation constitutes a collaborative interactional achievement, emerging through teachers' strategic employment of multiple modalities and students' timely orientation to those multimodal resources.

Keywords: choral response; multimodality; projectability; temporality; classroom interaction;
Conversation Analysis

1. Introduction

Choral response refers to the case in which two or more participants produce the same or substantially similar utterance simultaneously, which constitutes a recurrent form of collective participation in multiparty interaction. While it has been widely noted that choral production is far more than an incidental instances of simultaneous talk but an interactionally organized accomplishments (e.g., Lerner, 2002; Pfänder & Couper-Kuhlen, 2019), this phenomenon bears particular significance in classroom settings. Particularly in second language (L2) classrooms, choral responding represents one of the most pervasive participation formats, routinely employed in pronunciation practice, pattern drills, vocabulary learning, and grammar instruction. Previous research has identified a range of pedagogical benefits of mobilizing students' choral responses, including maximizing learners' opportunities for oral production, reinforcing pronunciation and form-focused repetition, and fostering collective engagement in classroom activities (e.g., Waring, 2011; Zhou, 2024). Despite the pedagogical centrality of choral responses in L2 instruction, remarkably little is known about how teachers accomplish their mobilization in the moment-by-moment unfolding of classroom interaction. Existing studies of everyday interaction have demonstrated that speakers frequently mobilize choral responses from interlocutors through specific syntactic and sequential designs that enable multiple recipients to jointly complete an emerging action (De Stefani & Mondada, 2014; Lerner, 2002; Tekin, 2023). However, classroom interaction constitutes a fundamentally different participation framework from ordinary conversation, which is institutionally organized and interactionally asymmetrical, with teachers routinely controlling topic development, allocating speaking rights, and regulating students' access to the conversational floor (Candela, 1999; McHoul, 1978; Seedhouse, 2004). In other words, the opportunities for learner self-selection remain comparatively constrained in classroom settings and the mobilization of choral responses by teachers cannot be understood simply as an extension of collective elicitation practices observed in everyday interaction.

Despite the limited body of research explicitly examining how teachers mobilize whole-class responses with multiple modalities, previous studies of classroom interaction have identified a variety of interactional practices through which teachers elicit learner participation, including designedly incomplete utterances (DIUs), explicit nomination, repetition, questioning, and gesture (e.g., Hazel & Mortensen, 2019; Koshik, 2002; Mehan, 1979; Waring, 2011; Wang, 2023). These studies have considerably advanced our understanding of how nonverbal modalities was employed by teachers to pursue learner responses and organize classroom participation. What remains unclear is how multiple modalities are dynamically coordinated in the moment-by-moment organization of classroom interaction to mobilize choral responses and the distinctive yet complementary contributions of different modalities to the projection of choral response. To address this issue, the present study draws on multimodal Conversation Analysis (CA)

and video-recorded data of Chinese L2 classroom interaction. This approach enables a fine-grained examination of the moment-by-moment unfolding of interaction, which is helpful to reveal how teachers and students display their understandings of turn design, sequential organization, and nonverbal conducts. As the analysis will show, the orchestration of choral responses constitutes a collaborative interactional accomplishment achieved through the systematic coordination of multiple multimodal resources, including materially anchored pointing gestures which functions to project the content of anticipated responses, gaze reorientation which signals the appropriate moment for collective entry, and successive prompting practices which serves to progressively establish recognizable response formats. By explicating the interactional functions served by specific multimodal resources, the present study contributes to our understanding of how whole-class participation is organized as an observable interactional achievement in L2 classroom interaction.

2. Literature Review

Conversation Analysis (CA) has long demonstrated that the organization of turn-taking is grounded in the principle that “overwhelmingly, one party talks at a time”, with participants orienting to the minimization of both gaps and overlaps in the sequential unfolding of interaction (Sacks et al., 1974, p. 699). According to the turn-taking model proposed by Sacks et al. (1974), transition at possible completion points of a turn-constructional unit (TCU) is systematically organized through three basic mechanisms: the current speaker may select the next speaker, another participant may self-select, or the current speaker may continue speaking. The orderly organization of turn-taking, however, does not preclude the occurrence of overlap, which typically arises when two or more participants anticipate the possible completion of an ongoing turn and coordinate their own entries (Jefferson, 1973, 1986). One recurrent type of such orderly overlap is *choral production*, also termed as *choral speech* or *simultaneous response*, in which multiple participants produce substantially similar utterances concurrently (Lerner, 2002; Pfänder & Couper-Kuhlen, 2019; Walker, 2016). Choral production has been documented across a wide range of social activities, including laughter (Glenn, 2003; Jefferson, 1979), booing (Clayman, 1993), cheering (Kerrison, 2018; Tekin, 2023), and chanting (Cummins, 2018; Cummins & Longo, 2023; Yagi, 2022). Across this body of literature, one consistent observation particularly pertinent to the present study is that choral production is typically grounded in participants’ shared orientation to those interactional resources that render both the timing of turn transitions and the projected content of the next turn publicly recognizable. These resources encompass a range of syntactic, prosodic, and sequential practices through which participants collaboratively coordinate the production of collective action. Lerner’s (2002) seminal study, for example, demonstrates that speakers recurrently design their turns with highly projectable

syntactic trajectories, which enables multiple recipients to anticipate the next turn and results in the orderly accomplishment of choral production. Extending this line of inquiry beyond verbal interaction, De Stefani and Mondada (2014) examine the orchestration of choral production during guided tours, finding that choral production is typically achieved through the coordinated deployment of talk, pointing gesture, and changes in bodily trajectory, and the establishment of new objects of joint attention. A more recent study by Tekin (2023) on multiplayer video gaming further demonstrate that choral vocalization from game players is systematically occasioned by unfolding contextual contingencies and particular events in the local environment. Collectively, these studies converge on the observation that choral production is far more than an accidental consequence of simultaneous speaking. Rather, it constitutes an interactional accomplishment through which participants collaboratively coordinate the timing and design of collective entry by orienting to multiple interactional resources that make subsequent action publicly projectable, including syntactic organization, embodied conduct, sequential positioning, and locally emergent contextual configurations.

Although previous studies have substantially advanced our understanding of the sequential and linguistic organization of choral production across a variety of interactional settings, considerably less attention has been paid to how such choral productions are orchestrated in second language classroom, a distinctive site from everyday interaction because of its institutional organization. Different from ordinary interaction where opportunities for speaker selection are distributed among participants, classroom interaction is characterized by an asymmetrical participation framework in which teachers typically initiate and close instructional sequences and allocate speaking rights to students (McHoul, 1978; Mehan, 1979; Seedhouse, 2004). While such institutional asymmetry ensures the orderly management of participation within large groups, it also substantially limits students' opportunities for self-selection. This renders teachers' interactional competence in mobilizing choral production especially consequential and deserves our closer investigation. Despite the paucity of research explicitly examining teachers' temporal organization of multimodal resources in mobilization whole-class response, previous studies of classroom interaction have identified a variety of interactional practices through which teachers elicit learner response. These include diverse linguistic formats, such as explicit nominations (Candela, 1999; Sari, 2020), yes-no questions (Lee, 2008), repetition (Hosoda & Aline, 2013), and designedly incomplete utterances (Hazel & Mortensen, 2019; Koshik, 2002; Wang, 2023, 2024), which have been shown to establish projectable response slots and mobilize student participation. Beyond these verbal practices, recent research has also documented a variety of embodied and material resources employed by teachers in pursuing learner responses. Such resources include gesture hold (Wang, 2023), screen-based textual and visual cues (Balaman, 2018), visible pedagogical materials (Hazel & Mortensen, 2019; Wang, 2023, 2024), and the spatial organization of the classroom (Hall & Looney, 2019; Wang, 2024). These studies have considerably advanced our

understanding of how specific linguistic, embodied, and material resources was recurrently used by teachers to pursue learner responses and organize classroom participation.

Nevertheless, existing research examining the ways teachers elicit responses from students has largely focused on individual learner responses and repair initiation, whereas studies addressing the mobilization of choral responses and the temporal organization of multiple modalities have been predominantly conducted in everyday interaction settings. What remains unclear is how multiple modalities are dynamically coordinated in the moment-by-moment organization of classroom interaction to mobilize choral responses and the distinctive yet complementary contributions of different modalities to the projection of choral response. Addressing these questions is of both theoretical and pedagogical significance, as they bear directly on the development of Classroom Interactional Competence (CIC), namely, teachers' and learners' ability to employ interaction as a resource for mediating and facilitating learning (Walsh, 2012, 2014). As is argued by Sert (2015), learning opportunities are fundamentally tied to participants' management of turn-taking, repair, instructional activities, and embodied conduct. Situated within this line of inquiry, the present study will draw on the multimodal Conversation Analysis to conduct a fine-grained examination of the moment-by-moment unfolding of interaction in Chinese L2 classrooms, aiming to reveal how teachers and students orient to the temporal unfolding of choral responses in Chinese L2 classrooms and display their understandings of turn design, sequential organization, and nonverbal conducts. This paper first discusses how materially anchored pointing gestures project the content of projected choral responses. It then turns to teachers' strategic deployment of gaze reorientation in signaling the appropriate moment for collective entry. Finally, it explores how teachers secure recurrent patterns of choral responding through successive prompting practices, particularly in environments where initial attempts to mobilize a collective response prove unsuccessful. By doing so, this study aims to explicate the temporal coordination of multiple multimodal resources, each distinctive yet complementary interactional functions in the organization of collective classroom participation.

3. Data and Method

This study draws on a corpus of publicly available video-recorded face-to-face interactions from Chinese L2 classrooms collected from several online platforms, primarily YouTube and Bilibili. The recordings were identified using English and Chinese search terms such as 'Chinese L2 classroom' and 'Chinese as a Foreign Language (CFL) classroom'. The videos were uploaded by educational institutions, language schools, or individual instructors for educational purposes and depict naturally occurring classroom

teaching. Because the recordings were produced independently of the present research, they preserve the naturally unfolding organization of classroom interaction and provide access to a broad range of instructional settings and pedagogical practices. During data collection, recordings that were clearly edited were excluded and only videos documenting continuous classroom interaction suitable for CA investigation were included. The resulting corpus comprises 64 recordings, each ranging from approximately 15 minutes to over one hour in duration. Although learners' proficiency levels vary from elementary to advanced, Mandarin Chinese serves as the primary medium of instruction throughout the dataset.

Data analysis proceeded in two stages. The first stage involved identifying all instances of students' choral responses within the corpus. Given the multiparty nature of classroom interaction, the scope of a choral response necessarily varies according to both the number of participating students and the degree of temporal overlap among their responses. For the purposes of this study, an instance was included as a *choral response* in the dataset when (a) two or more students produced a substantially similar response to a teacher's prompt; and (b) their responses overlapped at least partially in time. Applying these criteria, a total of 36 analyzable cases of choral response in which the overlapping productions were sufficiently audible for detailed interactional analysis were identified in our dataset. The second stage of our analysis involved a fine-grained multimodal Conversation Analytic (CA) examination of these cases. Teachers' embodied conduct was annotated using ELAN (Version 6.6), while the prosodic features of teachers' designedly incomplete utterances were examined using Praat (Version 6.3.10). All excerpts were transcribed according to the Conventions of Conversation Analysis by Jefferson (2004). Because Chinese orthography does not employ capitalization for emphasis, stressed syllables are indicated through underlining following Du Bois et al. (1993). Conventions for representing embodied conduct follow Mondada (2024), with asterisks (*---*) denoting gesture and plus signs (+---+) indicating gaze on separate transcript lines.

4. Multimodal Resources for Mobilizing Students' Choral Responses

This section examines how teachers mobilize students' choral responses through the coordinated deployment of three key multimodal resources. The analysis is organized in a way to take a fine-grained look at the interactional functions of each of the three modalities through which teachers accomplish the mobilization of choral responses. The first subsection examines how the content of choral responses is projected by materially anchored pointing gestures; the second subsection investigates how the appropriate moment for collective entry is established by teachers' gaze reorientation; and the third subsection explores how teachers secure recurrent patterns of choral responding

through successive prompting practices, particularly in environments where initial attempts to mobilize a collective response prove unsuccessful.

4.1 Pointing gestures and materially anchored choral responses

Previous research has widely noted that gesture constitutes an important resource in the organization of turn-taking, such as the alignment of gesture withdrawal and possible turn completion (Duncan, 1972; Kendrick et al., 2023). The recurrent gestural practice identified in the present data, however, differs in an important respect. Rather than functioning primarily as a turn-yielding device, teachers' pointing gestures are systematically laminated with pedagogical materials to render the projected content of choral responses publicly recognizable. Through this coordination of embodied conduct and material artifacts, teachers delimit the space of relevant responses, narrow the range of possible completions, and increase the projectability of the anticipated collective production. The following excerpt illustrates such an interactional practice.

Excerpt 1. ((The teacher displays several Chinese phrases related to language-learning activities on the screen and asks students to construct sentences using them.))

01 T: 今天 你们 有 中文 课 吗?
jīntiān nǐmen yǒu zhōngwén kè ma?
today 2PL have Chinese lesson PRT
Do you have Chinese lesson today?

02 (0.8s)

03 来(.) 大家 一起 说:.
lái(.) dàjiā yìqǐ shuō:.
come guys together say
Let's say it together.

04 我们: [今天: 有: 中文 课].
wǒmen: [jīntiān: yǒu: zhōngwén kè].
1PL today have Chinese lesson
We [have Chinese lesson today].

05 Ss: [今天: 有: 中文 课].
 [jīntiān: yǒu: zhōngwén kè].
 today have Chinese lesson
 [have Chinese lesson today].

06 T: 我们 得-
 wǒmen děi-
 1PL have.to
 'We have to-

07 → 我们 得:?
 wǒmen děi:?
 1PL have.to
 We have to:?

08 → (1.0s)
 ---point to the Chinese phrase 'learn Chinese' on the screen---



Figure 1a.

09 Ss: → [学: 中:文:].
 [xué: zhōng:wén:].
 learn Chinese
 'Learn Chinese:.'

10 T: → [学: 中:文:].
 [xué: zhōng:wén:].
 learn Chinese
 'Learn Chinese:.'

- 11 好.
 hǎo.
 good
 Good.
- 12 今天 的 中文 课 我们 做 什么?
 jīntiān de zhōngwén kè wǒmen zuò shénme?
 today ATT Chinese lesson 1PL do what
 What will we do in today's Chinese lesson?
 -----change to another slide-----
- 13 (0.9s)
- 14 °今天 的 中文 课° 我们 做 什么?
 °*jīntiān de zhōngwén kè*° wǒmen zuò shénme?
 today ATT Chinese lesson 1PL do what
 What will we do in today's Chinese lesson?'
- 15 Ss: 复习:-
 fù.xí:-
 review
 'Review-'
- 16 T: → 我们:?
 wǒmen:?
 1PL
 'We:?'
 --point to the Chinese phrase 'review weekly quiz' on the screen--



Figure 1b.

- 17 Ss: → [复:习: 周:考:].
 [fū:xí: zhōu:kǎo:].
 review weekly.quiz
 ‘Review the weekly quiz.’
- 18 T: → [复:习: 周:考:].
 [fū:xí: zhōu:kǎo:].
 review weekly.quiz
 ‘Review the weekly quiz.’
- 19 这 个 星 期 有 周 考 吗?
 zhè ge xīngqī yǒu zhōukǎo ma?
 this CLF week have weekly.quiz PRT
 ‘Will there be a weekly quiz this week?’
- 20 (0.7s)
- 21 S2: 这 个 星 期 没 有 周 考.
 zhè ge xīngqī méi yǒu zhōukǎo.
 this CLF week NEG have weekly.quiz
 ‘These will not be a weekly quiz this week.’

In this excerpt, the class is engaged in practicing vocabulary and sentence patterns with the target lexical items visibly displayed on the presentation slide. At line 7-8, the teacher produces a designedly incomplete utterance terminating in the lengthened modal *dei* ‘have to’ followed by a pointing gesture directed toward the displayed phrase *xue zhongwen* ‘learn Chinese’. After a brief pause, teacher and students enter the target phrase in overlap (lines 9-10). Importantly, the students’ response here is not simply occasioned by the teacher’s withdrawal from the turn but constitutes a collaboratively launched choral production grounded in the teacher’s embodied orientation to the displayed lexical item. The projected completion is rendered publicly recognizable through the lamination of embodied conduct and material display, with the teacher’s sustained pointing gesture making the target response visually available and projectable. A comparable interactional configuration recurs at line 16, where the teacher similarly employs a designedly incomplete utterance while pointing to the phrase *fxi zhoukao* ‘review the weekly quiz’ displayed on the screen, following which the class subsequently produces the projected completion in full overlap with the teacher (lines 17-18). Here, the teacher’s sustained pointing gesture serves not merely to direct visual attention but to

delimit the space of relevant responses by specifying precisely which lexical item constitutes an appropriate completion. Notably, the teacher's simultaneous production of the target phrase does not diminish students' participation but rather functions as feedback that simultaneously ratifies the students' collective response as the relevant next action. Taken together, these two cases demonstrate that the mobilization of choral responses depends on the coordinated deployment of syntactic projection, materially anchored pointing gestures, and visible pedagogical artifacts, which makes the content of choral response publicly recognizable.

Excerpt (2) presents a related interactional practice in which material anchoring similarly constrains the response space but does so through a markedly different configuration of embodied and environmental resources.

Excerpt 2. ((Prior to the excerpt, the teacher performs several actions and asks the students to describe what he is doing. In this instance, the teacher performs a new action by moving a chair from the inner part of the classroom toward the door.))

- 01 T: 很 好.
 hěn hǎo.
 very good
 ‘Very good.
- 02 现在 .hhh 老师: 再 做 一 个 动作.
 xiànzài .hhh lǎoshī: zài zuò yí gè dòngzuò.
 now teacher also do one CLF action
 Now let me do another thing.
- 03 你们 看.
 nǐmen kàn.
 2PL look
 ‘You guys take a look.
- 04 (0.9s)

05 老师: 在 做 什么?
lǎoshī: zài zuò shénme?
 teacher at do what
 'What am I doing?'

06 (0.6s)

07 Ss: 搬 椅子.
bān yǐzi.
 move chair
 'Moving a chair.'

08 T: °好°.
 °Hǎo°.
 good
 'Good.'

09 → 搬: 到:
 bān: dào:
 move to
 Move to,
 *-----move a chair---->>

Figure 2a.



10 → .hhh 哪儿 了?
 .hhh nǎr le?
 where PST
 where?
 >>-----*

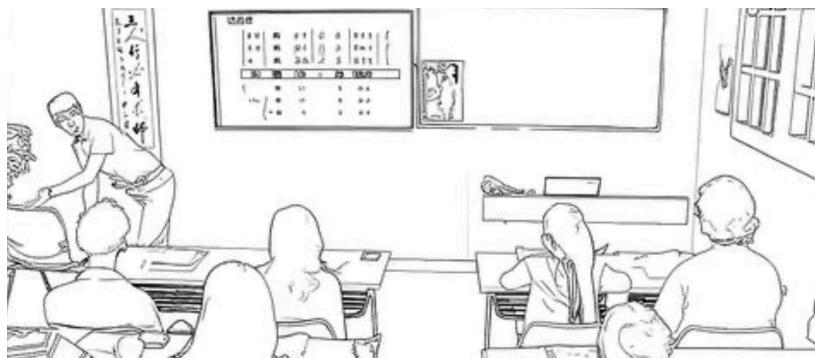


Figure 2b.

11 (0.5s)

12 Ss: → 教室 门口.
Jiàoshì ménkǒu.
 classroom door
 ‘The door of our classroom.’

13 T: 嗯: 用 把 怎么 说?
en yòng bǎ zěnmē shuō?
 INJ use BA how say
 ‘Em how to say it with the ‘ba-structure’?’

14 (0.4s)

15 Ss: 老师 把 椅子 搬 到 教室 门口 了.
lǎoshī bǎ yǐzi bān dào jiàoshì ménkǒu le.
 teacher BA chair move to classroom gate PST
 ‘The teacher moves the chair to the door of our classroom.’

In this excerpt, the teacher demonstrates the *ba*-construction through bodily enactment by physically moving a chair across the classroom while inviting students to describe the unfolding action. At lines 9, the teacher produces the designedly incomplete utterance *ban dao* ‘move to’ with audible prosodic lengthening and the embodied movement of the chair toward the classroom door. Notably, unlike Excerpt (1), the projected completion in this case is not directly retrievable from a written display but instead grounded in the evolving material environment, which requires students to derive the relevant response from the teacher’s embodied conduct and the chair’s newly established spatial position. In this sense, the material environment functions not as a repository of lexical items but

as an interactionally organized referential field from which the projected completion must be inferred. Following the designedly incomplete utterance, an additional verbal practice from the teacher emerges at line 10, which is the production of the wh-word *nar* 'where'. This verbal conduct further constrains the response space by making explicit the semantic domain within which the projected completion is to be formulated. Rather than supplying the target answer, the teacher incrementally increases its projectability by combining syntactic projection, embodied demonstration, and interrogative specification.

Taken together, a comparison between Excerpts (1) and (2) reveals that, despite relying on the coordinated deployment of embodiment and materiality to mobilize choral responses, teachers' practices may differ systematically in the degree to which the projected response is specified. In Excerpt (1), materially anchored pointing gestures and visible textual displays delimit a highly constrained response space in which the target lexical item is directly available for collective production, while in Excerpt (2), the teacher's embodied demonstration together with the wh-format establishes a projectable response slot while leaving the lexical formulation to the students themselves. These findings highlight that material anchoring operates along a continuum of response specification from rendering the projected completion directly readable from visible pedagogical materials to simply indicating the relevant referent which requires students to transform perceptual information into an appropriate linguistic formulation. This continuum reveals that the interactional contribution of materiality lies in calibrating the degree of projectability through which teachers mobilize students' collective responses.

4.2 Gaze reorientation and the temporality of collective entry

In addition to materially anchored pointing gestures, teachers' gaze plays a crucial role in organizing the temporality of choral responses. Previous CA research has adequately shown that gaze constitutes an important interactional resource for turn-taking and speaker transition by displaying the current speaker's orientation toward a prospective next speaker (Goodwin, 1981; Kendon, 1973). In classroom settings, however, gaze operates under a participation framework markedly different from ordinary conversation, in which teachers typically distribute their attention between pedagogical materials, a specific student, and the whole class as a collective recipient. This subsection examines the systematic redirection of the teacher's gaze from pedagogical materials, such as slides, the blackboard, or classroom objects, back to the students, and shows how this embodied shift contributes to establishing when the projected response becomes actionable for the class as a collective next speaker, which can be illustrated by the excerpt given below.

Excerpt 3. ((Prior to the excerpt, the teacher shows the students a picture of a place covered with ice and snow. In this instance, the teacher asks whether the students have ever been to places like this, in order to practice sentence construction using the Chinese word for “ever”.))

01 T: 你们(.) 去 过 那么 冷 的 °地方° 吗?
nǐmen(.) qù guò nàme lěng de °dìfang° ma?
2PL go ever that cold ATT place PRT
‘Have you ever been to a very cold place?’

02 (0.6s)

03 Ss: 从来 没 去 过.
cónglái méi qù guo.
ever NEG go ever
‘I have never been to a place like that.’

04 T: °哎°.
ai.
INJ
‘Ah.

05 林一斌 去 过 吗?
lín yībīn qù guo ma?
Linyibin go ever PRT
Linyibin, have you ever been to a place like that?’

06 (0.5s)

07 S1: 没 去 过.
méi qù guo.
NEG go ever
‘I have never been to a place like that.’

- 08 T: °没 去 过°.
°méi qù guo°.
NEG go ever
'You have never been to a place like that.'
- 09 一 次 也 没 去 过?
yí cì yě méi qù guo?
one CLF even NEG go ever
You have never been to a place like that even once?
- 10 (0.4s)
- 11 一 次 也 没 去 过.
yí cì yě méi qù guo.
one CLF even NEG go ever
You have never been to a place like that even once.
- 12 → 我:?
wǒ:?
1SG
I: ?
*-----point to *conglai* 'ever' on the blackboard-----*
+ ---gaze switched from the blackboard to students---+



Figure 3a.



Figure 3b.

13 Ss: → 从来 也 没 去 过:-
 cónglái yě méi qù guo:-
 ever ever NEG go ever
 'Have never been to a place.'

14 T: 那么:?
 nàme:?
 very
 very: ?

15 (0.6s)

16 Ss: 冷 的 地方.
 lěng de dìfang.
 cold ATT place
 'Cold place'

17 T: 对了.
 duìle.
 right
 'Right.'

In this excerpt, the class is engaged in practicing sentence production using target lexical items displayed on the blackboard. Following prior discussion about whether participants have ever visited a very cold place, the teacher points to the target word *conglai* 'ever' while initially orienting her gaze toward the board and producing a designedly incomplete utterance with audible lengthening on wo 'I' (line 12). Crucially, while maintaining the pointing gesture, the teacher subsequently redirects her gaze from the blackboard to the

students, right after which the class makes the projected choral response. Here, the students' response is oriented to the teacher's multimodal packaging of the projected action in which verbal design, pointing gesture, and gaze are sequentially coordinated. Whereas the displayed lexical item and the pointing contributes primarily to the projectability of response content, the teacher's gaze reorientation functions to establish the moment at which collective entry becomes interactionally relevant. This case makes visible the differentiated interactional work performed by materiality and gaze, with the latter organizes the temporality of collective participation.

The function of gaze reorientation in mobilizing students' choral response is also apparent in Excerpt (4) with the sequential configuration differs in several important respects.

Excerpt 4. ((Prior to the excerpt, the teacher displays the topic "Thanksgiving" on the screen. In this instance, she elaborates on the topic and provides instructions on what the students should discuss in the following activity.))

- 01 T: 感恩节 的 °时候° hhh
 gǎnēnjié de °shíhou° hhh
 Thanksgiving ATT time
 During Thanksgiving,
- 02 → 你:(.) [一:定: 会: 买: 什 么 东 西],
 nǐ:(.) [yí:dìng: huì: mǎi: shénme dōngxi],
 2SG definitely will buy what thing
 What will (you) buy for sure?
 ----point to the material-----
 +-----gaze from screen to students-----+

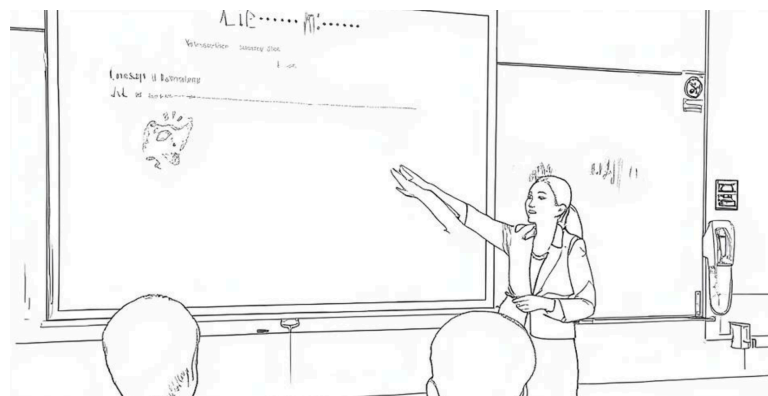


Figure 4a.

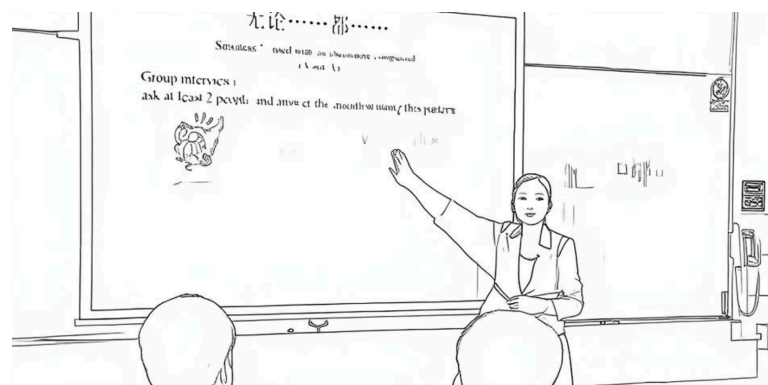


Figure 4b.

- 03 Ss: → [一:定: 会: 买: 什: 么 东: 西]
 [yí:dìng: huì: mǎi: shénme dōngxi]
 definitely will buy what thing
 'What will (you) buy for sure.'

- 04 T: → (0.8s)
 +-----gaze from students back to screen-----+

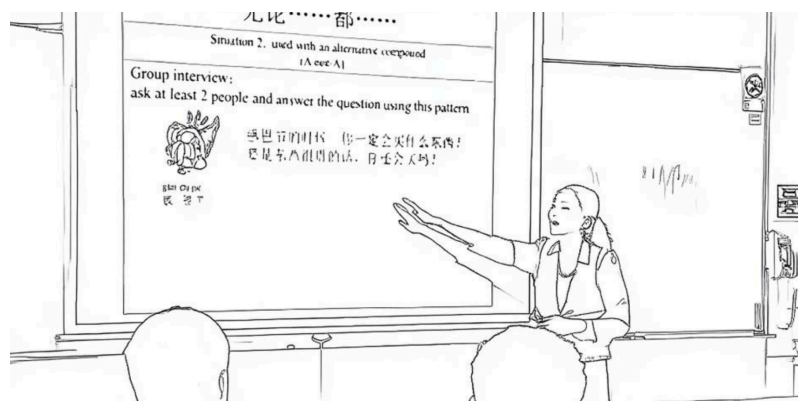


Figure 4c.

- 05 → 要是 °这 个 东西° 很 贵,
yàoshì °zhè gè dōngxi° hěn guì,
 if this CLF thing very expensive
 ‘If this thing is very expensive,
- 06 (1.2s)
- 07 → 你: 还 会 买 吗?
nǐ: hái huì mǎi ma?
 2SG still will buy PRT
 will you still buy it?
- 08 (0.7s)
- 09 用 °这 个° 句型.
yòng °zhè ge° jùxíng.
 use this CLF sentence.type
 Use this sentence.
- 10 你 看.
nǐ kan.
 2SG see
 Let’s see.
- 11 你(.) 一定 会 买 什么 °东西°.
nǐ(.) yídìng huì mǎi shénme °dōngxi°.
 2SG definitely will buy what thing
 What will you buy for sure.’

In this sequence, the teacher introduces a sentence frame related to Thanksgiving while simultaneously pointing toward the displayed instructional material. While the lengthened production of *nǐ* ‘you’ at line 2 functions to establish a syntactically projectable frame that is locally available for repetition by the class, particularly noteworthy is that students initiate their choral repetition precisely when the teacher’s gaze reaches them (line 3). Once again, the displayed material and pointing gesture contribute to making the response format recognizable, whereas gaze reorientation specifies the temporal point at which that projected response becomes collectively actionable. What deserves further attention is the teacher’s subsequent ability to reclaim the conversational floor. By redirecting her gaze back toward the displayed material (line 4), she effectively suspends

the opportunity for students' entry and resumes sole speakership. This case shows that gaze oriented to and away from students serves as a finely calibrated interactional resource through which teachers regulate not only the initiation but also the suspension of whole-class participation. Taken Excerpts (3) and (4) together, we can see that gaze reorientation performs a distinctive interactional function in the mobilization of students' choral responses. Unlike materially anchored gestures which primarily delimit the projected response space, gaze organizes the temporal architecture of collective participation by transforming a projected response into an interactionally relevant next action for the class. These findings highlight that choral responses are accomplished not through any single multimodal resource and teachers' embodied conduct should not be understood as a simple collection of independent communicative cues. Instead, different resources are components of a laminated ensemble whose interactional significance arises from their temporal and sequential coordination.

4.3 Successive prompts as a repair to initial unsuccessful elicitation

The preceding subsections have shown how teachers mobilize students' choral responses through the coordinated deployment of material and embodied resources that render both the projected response and the moment of collective entry publicly recognizable. This subsection examines the mobilization of choral response in a broader sequential environment, shifting focus from how teachers project a specific response to how teachers progressively establish a publicly recognizable response format that can subsequently be mobilized by the class as a whole. Such successive prompts, frequently accompanied by demonstrations, confirmations, or reformulations, constitute a recurrent practice through which teachers repair an initially unsuccessful elicitation and progressively establish the conditions for collective participation, which can be illustrated by Excerpt (5) given below.

Excerpt 5. ((Prior to the excerpt, the teacher takes out two pieces of clothing, holding one in each hand. In this instance, the teacher asks the students to describe the color and price of the clothes in Chinese.))

01 T: 你 看 这 两 件 衣服.
 nǐ kàn zhè liǎng jiàn yīfu.
 2SG see this two CLF cloth
 'Here are two clothes.

02

这件 是:?

zhè jiàn shì:?

this CLF COP

This is:?’

----left hand raising the red cloth----



Figure 5a.

03

(1.0s)

04

Ss: 红色 的.

hóngsè de.

red NR

‘Red.’

05

T: 对 了.

duì le.

right PST

‘Right.’

06

→ 这件 是:?

zhè jiàn shì:?

this CLF COP

This is:?’

----right hand raising the white cloth----



Figure 5b.

- 07 Ss: → 白色 的.
báisè de.
 white NR
 ‘White.
- 08 T: 白色 的.
báisè de.
 white NR
 ‘White.
- 09 样子(.) 也 不 一样.
yàngzi(.) yě bù yíyàng.
 appearance also NEG same
 They also look different.
- 10 而且 °呢°,
érqiě ne,
 also PRT
 ‘And,
 ----turn the two clothes around-----
- 11 这 件:?
zhè jiàn:?
 this CLF
 ‘This: ?

-----raising the red cloth in the left hand-----



Figure 5c.

12 Ss: 价钱-

jiàqián-

price

'Price-'

13 T: 二 百 块 钱.

yì bǎi kuài qián.

one hundred CLF money

'One hundred yuan.

14 → 这 件:?

zhè jiàn:?

this CLF

This:?

----right hand raising the white cloth-----



Figure 5d.

15 Ss: → 两 百.
liǎng bǎi.
 two hundred
 'Two hundred.'

16 T: 二 百 块.
èr bǎi kuài.
 two hundred CLF.
 'Two hundred.'

17 °二 百 块°.
°èr bǎi kuài°.
 two hundred CLF.
 'Two hundred.'

In this excerpt, the class is practicing the useful expressions to describe clothing items, including color and price. At line 2, the teacher raises one item of clothing while producing the designedly incomplete utterance *zhe jian shi* 'this one is'. However, rather than responding immediately, the class remains silent for approximately one second before eventually supplying *hongse de* 'red' (lines 3-4). Following the delayed response from students, the teacher produces an explicit confirmation *duile* 'right' (line 5) and then repeats the same designedly incomplete utterance while displaying the second item of clothing (line 6). On this occasion, the class produces the projected response without delay (line 7, indicating that the preceding sequence has successfully established a recognizable response format. A comparable sequential development occurs at line 11, where the teacher raises another clothing item bearing a price tag while producing the

rising-intonation designedly incomplete prompt *zhe jian* ‘this one’, thereby projecting a further opportunity for collective response. Rather than supplying the exact price, the students initially respond with *jiaqian* ‘price’ (line 12), displaying an orientation to the semantic domain of the teacher’s question without yet identifying the projected response itself. Instead of treating this as a failed response, the teacher provides the target model *yi bai kuai qian* ‘one hundred yuan’ (line 13) and then recycle the same prompt with the second item of clothing (line 14). On this occasion, the class immediately produces the exact price *liang bai* ‘two hundred’ (line 15). From the two cases discussed so far, we can see how successive prompts can be employed by teachers as a sequential practice to mobilize expected choral response from students. The initial prompt typically serves as an interactional template to publicly establish a recognizable format of choral response, ensuring the subsequent prompt to receive desirable choral responses.

Unlike Excerpt (5) where the projected response format is established primarily through the teacher’s own demonstration and modeling, the interactional work in Excerpt (6) below is collaboratively accomplished by the teacher and an individual student.

Excerpt 6. ((Prior to the excerpt, the teacher introduces the new word *jile* ‘extremely’ to the students, explaining its meaning and how it is used. In this instance, the teacher prompts the students to produce phrases or sentences using the word *jile*.)

- 01 T: 极:了.
 jí:le.
 extremely
 Extremely.
 ((pointing to the word *jile* ‘extremely’ on the blackboard))
- 02 (1.1s)
- 03 °极了°.
 jile.
 extremely
 Extremely.

- 04 前边(.) 一般 用 形容词.
qiánbiān(.) yìbān yòng xíngróngcí.
 front generally use adjective
 Generally, we use an adjective before it.
- 05 °对°.
 °*duì*°.
 right
 Right.
- 06 °形容词°.
 °*xíngróngcí*°.
 adjective
 Adjective.
- 07 ((T turns around to write on the blackboard))
 (4.3s)
- 08 °意思° 是: 非常.
 °*yìsi*° *shì:* *fēicháng*.
 meaning COP very
 Its meaning is ‘very’.
- 09 .hhh 非常.
 .hhh *fēicháng*.
 very
 Very.
- 10 对 吗?
 duì ma?
 right PRT
 Right?

11 好::

hǎo::

okay

Okay.

12 上下班 的 时候:: .hhh

shàngxiàbān de shíhou:: .hhh

commute ATT time

During commuting hours,

13 (0.8s)

14 北京 街上 的 车::

běijīng jiēshàng de chē::

Beijing street ATT car

the cars on the street of Beijing:

15 怎么样?

zěnmeyang?

how

How?

---hand summoning---



Figure 6a.

16 (2.0s)

- 17 S1: 多 极了.
duō jǐle.
 many extremely
 ‘Extremely many.’
- 18 T: 对.
duì.
 right
 ‘Right.’
- 19 句子.
jùzi.
 sentence
 Sentence.
 ----hand pointing to a student---
- 20 上下班 的 时候::.
shàngxiàbān de shíhou::.
 commute ATT time
 During commuting hours.’
- 21 Ss: 上下班 的 时候(.) 呢:: 北京 的 车 多 极了.
shàngxiàbān de shíhou(.) e:: běijīng de chē duō jǐle.
 commute ATT time INJ Beijing ATT car many extremely
 ‘There are extremely many cars in Beijing during commuting hours.’
- 22 T: 多 极了.
duō jǐle.
 many extremely
 ‘Extremely many.’
- 23 刚才 我们 说 了.
gāngcái wǒmen shuō le.
 just.now 1PL say PST
 Just now we said that,

24

→ 哈尔滨 的 冬天:

hāěrbīn de dōngtiān:

Harbin ATT winter

Harbin's winter:'

----point to the picture of Harbin on the blackboard---

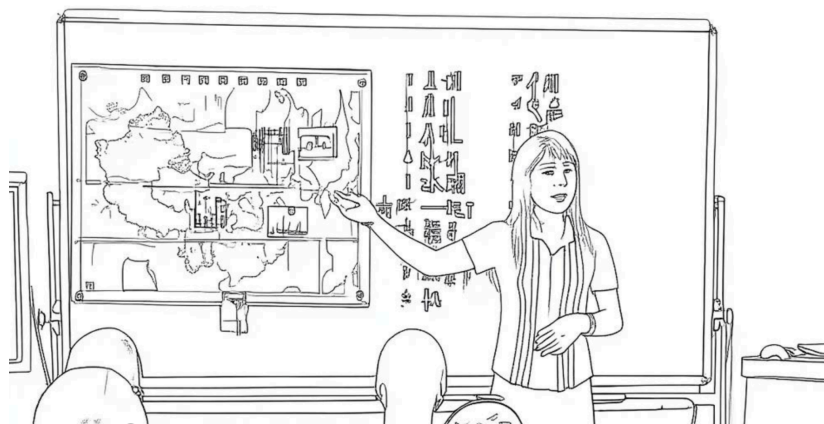


Figure 6b.

25 Ss: → 冷 极了.

lěng jile.

cold extremely

'Extremely cold.'

In this excerpt, the teacher first introduces the lexical item *jile* 'extremely' by explaining its meaning and grammatical usage (lines 1-10) and then contextualizes the target expression through an example concerning traffic conditions in Beijing (lines 11-14). Notably, the designedly incomplete utterance produced at line 14 does not by itself occasion an immediate collective response. Rather than treating the absence of uptake as interactionally problematic, the teacher incrementally increases the projectability of the anticipated response by reformulating the prompt with the interrogative *zenmeyang* 'how' (line 15) to make the semantic domain of the projected completion more explicit. Following a two-second silence, a student makes an attempt to supply a response *duo jile* 'extremely many' (line 17), which is immediately ratified by the teacher with an explicit confirmation *dui* 'right' (line 18). After this, the teacher invites the whole class to produce a complete sentence and further recycles the target expression through subsequent instructional talk (lines 19-20). Building on this interactional foundation, the teacher then introduces the parallel prompt *Haerbin de dongtian* 'Harbin's winter' (line 24). Notably, unlike the earlier sequence, the class produces the projected choral response

immediately before the teacher supplementing the designedly incomplete utterance with the interrogative *zenmeyang* 'how' to specify the semantic domain of the projected completion. The interactional achievement of this sequence again lies in the role played by the initial prompt. What deserves particular attention in this case is that, unlike Expect (5), the teacher's first prompt does not culminate in a successful choral response but only receives a response from a single student. Its key function here is to provide an interactional environment in which the teacher can confirm or reformulate students' initial attempt and thereby make the relationship between the prompt and the expected format of the choral response more recognizable to students. By the time the teacher launches the subsequent parallel prompt, the response format has already been interactionally constituted, enabling the class to orient to it as an established participation framework for collective production and to produce the projected response without further interactional scaffolding.

Taken together, the cases examined in this subsection indicate that, rather than relying on a single elicitation move, teachers use successive prompting to progressively establish the recognizability and projectability of choral response formats. Initial prompts need not themselves achieve immediate collective uptake but rather serve as sequential resources through which recognizable response templates are progressively assembled and subsequently presupposed by later prompts. In this sense, successive prompting can be understood as a sequential repair practice through which teachers transform initially unsuccessful elicitation into coordinated whole-class participation. These findings also demonstrate that the mobilization of choral responses does not merely depend on teachers' demonstrations but is collaboratively accomplished through the interactional co-construction of response templates by teachers and students.

5. Discussion and Conclusion

The present study examined how teachers mobilize students' choral responses in Chinese L2 classrooms through the coordinated deployment of multimodal interactional resources. Drawing on video-recorded classroom interaction and multimodal Conversation Analysis, this study examines materially anchored pointing gestures, gaze reorientation, and successive prompting as three key resources through which teachers progressively establish the projectability and temporality of collective participation. Findings highlight that the mobilization of choral response constitutes collaboratively accomplished interactional achievements emerging from the dynamic coordination of verbal, embodied, material, and sequential resources which requires the collaboration between teachers and students. This study has theoretical and pedagogical contributions in the following three aspects.

The first theoretical contribution of this study concerns the relationship between projectability and temporality in multimodal interaction. While previous research on choral production has emphasized the role of syntactic and embodied resources in projecting recognizable responses (Lerner, 2002) and that social action unfolds through the temporal coordination of talk, embodiment, and materiality (Deppermann & Streeck, 2018; Mondada, 2018), the present findings extend this line of inquiry by showing that the projection of choral responses depends on the temporal organization of multiple modalities and different modalities contribute distinct forms of projectability. Material artifacts and pointing gestures establish what response is relevant; gaze reorientation specifies when collective entry becomes interactionally appropriate; and successive prompting progressively establishes how a recognizable response format is to be produced. These observations contribute to ongoing discussions of *multimodal temporality* (Mondada et al., 2025) by demonstrating that the mobilization of collective participation is achieved through the strategic coordination of multiple interactional resources operating on temporal trajectories.

A second contribution concerns the interactional constitution of the whole class as a collective next speaker in educational settings. While previous studies have shown that choral productions in everyday interaction are made possible through participants' orientation to syntactic projection and sequential organization (De Stefani & Mondada, 2014; Lerner, 2002; Tekin, 2023) and that teachers routinely retain primary rights over turn allocation within institutionally asymmetrical participation frameworks (McHoul, 1978; Mehan, 1979; Seedhouse, 2004), the present study brings these two strands of research together by demonstrating how teachers transform a collection of individual students into a ratified collective next speaker through the systematic coordination of multimodal interactional resources. Rather than being institutionally given, our findings show that constituting the class as the collective next speaker requires teachers to render both the projected content of the next action and the appropriate moment for collective entry publicly recognizable and progressively establish a shared orientation to the projected response. These findings extend previous research on *collective speakership* by showing that whole-class participation in educational settings emerges as a situated respecification of turn allocation in everyday interaction.

The present study also contributes to the growing literature on Classroom Interactional Competence (CIC; Walsh, 2012, 2014; Sert, 2015). While existing discussions of CIC have emphasized teachers' ability to manage turn-taking, repair, and participation in ways that facilitate learning opportunities, the present study suggests that CIC should be understood as teachers' capacity to assemble multiple semiotic resources into coherent participation frameworks that enable learners to recognize both the content and the timing of pedagogically relevant actions. From this perspective, the mobilization of choral responses is considerably more than the implementation of a classroom drill. Teachers display interactional competence by continuously calibrating verbal design, embodied

conduct, visual materials, and sequential organization to learners' evolving interactional understandings. Particularly revealing are cases involving successive prompting, in which teachers transform initially unsuccessful elicitation attempts into opportunities for progressively establishing publicly recognizable response formats. Learning opportunities therefore emerge not simply because teachers ask questions but because they systematically reshape the interactional environment until participation becomes jointly achievable. This perspective also invites a reconceptualization of choral response within L2 pedagogy. Rather than viewing choral responses as mechanical repetition or teacher-controlled recitation, the present study demonstrates that they constitute collaborative interactional achievements requiring teachers and learners to establish shared orientations to projectability, temporality, and multimodal coordination. In other words, the pedagogical value of choral response lies not only in repeated language production but also in learners' developing ability to coordinate their participation with the evolving interactional organization of classroom activities.

Taken together, the present study advances current understandings of classroom interaction by demonstrating that the mobilization of students' choral responses is fundamentally a multimodal and interactionally organized accomplishment. By integrating insights from Conversation Analysis, multimodal interaction research, and Classroom Interactional Competence, it shows that materially anchored gestures, gaze reorientation, and successive prompting perform analytically distinct yet interactionally complementary functions in organizing collective participation. This study also highlights that whole-class participation should be understood as an emergent interactional achievement accomplished through participants' shared orientation to the projectability, temporality, and multimodal organization of social action.

References

- Balaman, U. (2018). Embodied Resources in a Repetition Activity in a Preschool L2. *Novitas-Research on Youth and Language*, 12(1), 27-51.
- Candela, A. (1999). Students' power in classroom discourse. *Linguistics and Education*, 10(2), 139-163.
- Clayman, S. E. (1993). Booring: The Anatomy of a Disaffiliative Response. *American Sociological Review*, 58(1), 110-130.
- Cummins, F. (2018). *The Ground from Which We Speak: Joint Speech and the Collective Subject*. Cambridge Scholars.
- Cummins, F., & Longo, L. (2023). The empirical discovery of domains of assembly and communion. *Language Sciences*, 100, 101586.
- De Stefani, E., & Mondada, L. (2014). Reorganizing mobile formations: When 'guided' participants initiate reorientations in guided tours. *Space and Culture*, 17(2), 157-175.
- Deppermann, A., & Streeck, J. (2018). *Time in Embodied Interaction. Synchronicity and Sequentiality of Multimodal Resources*. Benjamins.
- Du Bois J. W., Schuetze-Coburn, S., Cumming, S., & Paoline, D. (1993). Outline of discourse transcription. In J. A. Edwards & M. D. Lampert (Eds.), *Talking Data: Transcription and Coding in Discourse Research* (pp. 45-90). Lawrence Erlbaum Associates.
- Duncan, S. (1972). Some signals and rules for taking speaking turns in conversation. *Journal of Personality and Social Psychology*, 23, 283-292.
- Glenn, P. J. (2003). *Laughter in Interaction*. Cambridge University Press.
- Goodwin, C. (1981). *Conversational Organisation: Interaction between Speakers and Hearers*. Academic Press.
- Hall, J. K., & Looney, S. D. (2019). *The Embodied Work of Teaching*. Multilingual Matters.
- Hazel, S., & Mortensen, K. (2019). Designedly incomplete objects as elicitation tools in classroom interaction. In D. Day & J. Wagner (Eds.), *Objects, Bodies and Work Practice* (pp. 216-249). Multilingual Matters.

- Hosoda, Y., & Aline, D. (2013). Two preferences in question-answer sequences in language classroom context. *Classroom Discourse*, 4(1), 63-88.
- Jefferson, G. (1973). A case of precision timing in ordinary conversation: overlapped tag-positioned address terms in closing sequences. *Semiotica*, 9, 47-96.
- Jefferson, G. (1979). A technique for inviting laughter and its subsequent acceptance declination. In G. Psathas (Ed.), *Everyday Language* (pp. 79-96). Irvington.
- Jefferson, G. (1986). Notes on 'latency' in overlap on-set. *Human Studies*, 9, 153-183.
- Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. H. Lerner (Ed.), *Conversation Analysis: Studies from the First Generation* (pp. 13-31). John Benjamins.
- Kendon, A. (1973). The role of visible behavior in the organization of social interaction. In M. von Cranach & I. Vine (Eds.), *Social Communication and Movement: Studies of Interaction and Expression in Man and Chimpanzee* (pp. 29-74). Academic Press.
- 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 B*, 378, 20210473.
- Kerrison, A. T. (2018). *We're All behind You: the Co-construction of Turns and Sequences-At-Cheering*. Unpublished Doctoral Dissertation. Ulster University.
- 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.
- Lee, Y.-A. (2009). Yes-No Questions in the Third-Turn Position: Pedagogical Discourse Processes. *Discourse Processes*, 45(3), 237-262.
- Lerner, G. H. (2002). Turn-sharing: The choral coproduction of talk-in-interaction. In C. Ford, B. Fox, and S. A. Thompson (Eds.), *The Language of Turn and Sequence* (pp. 225-256). Oxford University Press.
- McHoul, A.W. (1978). The organization of turns at formal talk in the classroom. *Language in Society*, 7, 183-213.
- Mehan, H. (1979). *Learning Lessons: Social Organisation in the Classroom*. Harvard University Press.

- Mondada, L. (2018). Multiple temporalities of language and body in interaction: challenges for transcribing multimodality. *Research on Language and Social Interaction*, 51, 85-106.
- Mondada, L. (2024). Multimodal transcription conventions. In J. D. Robinson, R. Clift, K. H. Kendrick, & C. W. Raymond (Eds.), *The Cambridge Handbook of Methods in Conversation Analysis* (pp. 990-998). Cambridge University Press.
- Mondada, L., & Tekin, B. S., & Koda, M. (2025). A sequential approach to simultaneity in social interaction: The emergent organization of choral actions. *Language & Communication*, 102, 1-14.
- Pfänder, S., & Couper-Kuhlen, E. (2019). Turn-sharing revisited: an exploration of simultaneous speech in interactions between couples. *Journal of Pragmatics*, 147, 22-48.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696-735.
- Seedhouse, P. (2004). *The Interactional Architecture of the Language Classroom: A Conversation Analysis Perspective*. Oxford: Blackwell.
- Sari, C. C. (2020). Conversation Analysis: Turn-Taking Mechanism and Power Relation in Classroom Setting. *Celtic: A Journal of Culture, English Language Teaching, Literature and Linguistics*, 7(2), 118-136.
- Sert, O. (2015). *Social Interaction and L2 Classroom Discourse*. Edinburgh University Press.
- Tekin, B. S. (2023). Cheering together: the interactional organization of choral vocalizations. *Language Communication*, 88, 73-89.
- Walker, G. (2016). Phonetic variation and interactional contingencies in simultaneous responses. *Discourse Processes*, 53, 298-324.
- Walsh, S. (2012). Conceptualising classroom interactional competence. *Novitas-ROYAL*, 6(1), 1-14.
- Walsh, S. (2014). Developing classroom interactional competence. *Language Issues*, 25(1), 4-8.
- Wang, X. (2023). The interplay of designedly incomplete utterances and representational gestures in Chinese as a second language classrooms. *At the Crossroads*

Between Grammar and Multimodality: Resources for Interaction, *Bulletin suisse de linguistique appliquée*, 117, 81-104.

Wang, X. (2024). Multimodal resources for managing student participation after unsuccessful designedly incomplete utterances in CSL classroom. In H. Tao & Y. Zhou (Eds.), *Teacher-Student Engagement Patterns in CSL Classrooms and Beyond: Video-Based Multimodal Analysis* (pp. 279-307). John Benjamin.

Waring, H. Z. (2011). Learner initiatives and learning opportunities. *Classroom Discourse*, 2(2), 201-218.

Yagi, J. (2022). Achieving (a)synchrony through choral chanting: Co-operative corrections in taiko ensemble rehearsals. *Journal of Pragmatics*, 195, 48-68.

Zhou, Y. (2024). What's next after a correct answer: Maximizing engagement by shifting participation frameworks in the Chinese as a second language classroom. In H. Tao & Y. Zhou (Eds.), *Teacher-Student Engagement Patterns in CSL Classrooms and Beyond: Video-Based Multimodal Analysis* (pp. 251-278). John Benjamin.

Acknowledgements of Generative AI Use

This article was written with the assistance of ChatGPT (GPT-5, OpenAI, 2025). The use of ChatGPT was limited to polishing language and correcting grammatical mistakes. All conceptual ideas, analyses, and interpretations are the author's own. All outputs were critically reviewed, edited, and verified by the authors prior to submission.

Appendix A. Abbreviations used in the gloss

1SG	First person singular
2SG	Second person singular
3SG	Third person singular
ATT	Attributive marker <i>de</i>
CLF	Classifier
COP	Copula
CONJ	Conjunction
DE	Complementizer <i>de</i>
DET	Determiner
EXP	Experiential aspect
INJ	Interjection
NEG	Negation
NOM	Nominalizer
PRT	Utterance-final particle
PST	Past tense marker

Appendix B. Transcription conventions

.	Final falling intonation
,	Slight rising intonation
?	Sharp rising intonation
→	Target utterance
-	Abrupt cut-off
(.)	Short pause
(0.8)	Periods of silence, in tenths of a second
<u>He</u> says	Syllables or words louder than surrounding speech by the same speaker
°word°	Syllables or words distinctly quieter than surrounding speech by the same speaker
hhh	Outbreath, proportional to the number of letters
.hhh	Inbreath, proportional to the number of letters
::	Lengthening of the sound just preceding them, proportional to the number of colons
[]	Speech overlap
(())	Transcribers' comment or non-verbal actions
*	Gesture
+	Gaze