PPLING
International Applied Linguistics Research Conference
https://doi.org/10.29140/9781763711686-4

Multimodal Sequential Analysis of Word Searches in L2 Interactions: The Use of Pointing Gesture, Gaze, and Embodied Resources for Resolving Trouble-in-Talk in English as a Medium of Instruction (EMI) Classrooms

Shin Jung Ho
Fo Guang University, Taiwan

Abstract

This study explores how L2 students signal and resolve trouble-in-talk in English-medium instruction (EMI) classrooms. It examines two types of word search sequences, joint and solitary, using approximately 12 hours of video-recorded lessons from EMI courses at two universities in Taiwan. Using multimodal sequential analysis, the study shows that in the first extract, a student responds to the teacher's referential and open-ended question by laughing, saying "yeah," and then producing "uhm," which clearly signals repair. While saying "uhm," he points at his smartphone. A half-second pause follows, during which he holds the gesture and redirects his gaze to the teacher, creating a transition-relevance place for the teacher to supply the missing word. In the second extract, from a cooking class, a student answers the teacher's display question about ingredients with the deictic "this" and a long, stretched "um," while gazing at and pointing to the dried mushroom with a "thinking face." She thus frames the search as solitary, vocally sustaining progressivity while visibly displaying the search. Other participants withhold entry until she shifts to a quick pointing gesture with a wh-question asking for the ingredient's name. Her expression simultaneously changes from a thinking face to a smile, which builds rapport with co-participants after the

E-mail: sjho@mail.fgu.edu.tw

In Y. Wang, J. Matthews, & B. Zou (Eds.), *Proceedings of the 2nd International Applied Linguistics Research Conference, 2026* (pp. 26–34). Castledown.

solitary search. Overall, word searches in EMI classrooms are collaboratively accomplished: learners display lexical difficulty through pointing gesture, gaze, and multimodal resources, and teachers and peers respond with minimal support that keeps interaction moving, prioritizing subject-content learning over grammatical accuracy.

Keywords: Word search, pointing gesture, gaze, embodied conduct, multimodal sequential analysis

Background

This study investigates how second-language learners flag lexical problems and work through troubles-in-talk in English as a medium of instruction (EMI) classrooms—specifically, cases of joint and solitary word searches (Dressel, 2020) where the problematic item has not yet been produced. EMI refers to content courses delivered in English to students for whom English is largely not a native language. Prior work by Duran et al. (2022) examined word searches in Turkish EMI settings, finding that these searches are jointly managed through verbal and embodied conduct that makes the lexical trouble visible to participants. Their study also found that EMI instructors tend to prioritize content delivery over linguistic accuracy, often letting lexical searches go unresolved so that the flow of the lesson and disciplinary content take precedence. Examining multimodal word searches, the current study extends this line of inquiry to EMI classrooms in Taiwan. It focuses on the role of embodied resources and instructors' content-first orientation, while adding a new dimension: it shows how physical objects present in the classroom and learners' bilingual repertoires also help drive collaborative word retrieval.

Methodology

The dataset consists of 12 hours of video-recorded classroom interaction from two Taiwanese university EMI courses—one on media literacy and one on vegetable-based cooking. Transcripts were produced using conversation analysis conventions, supplemented with additional notation to capture embodied conduct. From this corpus, a collection of word search sequences was compiled and analysed to answer two questions: (1) which verbal, embodied, and material resources participants draw on to launch and resolve word searches, and (2) what interactional characteristics define word searches

specifically within the EMI setting. The analysis focuses on how participants deploy multimodal resources as lexical retrieval unfolds.

Data Analysis

Joint Word Search

Prior to the extract, the teacher initiates classroom participation by asking about the impact media have in students' lives. Following extended silences and the absence of student bids to respond, the teacher incrementally constructs the question (lines 8–12), allowing extended time for students to formulate a response before selecting Bowen through gaze, thereby accomplishing embodied turn allocation. Bowen first initiates repair with a clarification question (line 13), followed by the teacher's explanation sequence (lines 14–16). Then Bowen subsequently displays ongoing response preparation through repetitive desk tapping with alternating index and middle fingers (line 17). Rather than treating the silence as disengagement, the teacher's "no idea?" (line 18) serves as an epistemic check, assessing Bowen's epistemic access before offering further guidance (lines 19–20). After Bowen responds, "I can see uh (.) many countries news," he shifts his gaze towards the teacher, thereby displaying completion of his turn and making relevant the teacher's next action (line 22). The teacher reciprocates with a smile and a nod, which publicly acknowledge Bowen's contribution as acceptable before she produces the next question. Three lines are omitted in which Bowen identifies "Yahoo" as his news application. In line 26, the teacher uses the turn-initial particle "so" as an upshot marker to launch a follow-up question. In the next line, Bowen confirms that he reads news every day and laughs. Then he continues with an agreement token "yeah" followed by "uhm," which straightforwardly indicates repair (Schegloff, 2010). Accompanying "uhm," he points at his smartphone. Then there is a 0.5-second pause, which involves a gesture hold and a re-orientation of gaze towards the teacher (line 30). In this way, the student establishes a transition-relevance place for the teacher to provide the sought-for word.

Bowen begins his next turn with the stretched turn-initial sound "e::" and the hesitation marker "en" before producing "everyday message" (line 31). Simultaneously, Bowen maintains his gaze towards the teacher. Rather than continuing to search independently, he draws on gaze, pointing gestures, and the smartphone resting on the desk as multimodal resources to recruit the teacher's participation by identifying the location of the sought-after lexical item. The sustained pointing does not merely identify a physical object; it publicly locates the semantic domain from which the missing expression is to

be recovered, thereby transforming the word search into a collaborative activity. The teacher subsequently offers the candidate solution, supplying the missing prepositional phrase and the intended referent (line 32). The teacher's solution therefore provides the minimal grammatical resource necessary to complete the emerging sentence construction, but not the full clause.

Extract 1

```
08   TEA:      what impact does media have in your life
09             (1.0)
10   TEA:      or you can say what good impact does media have
11             (0.5)
12   TEA:      on your *everyday life
     tea              *looks at Bowen -- >
13   Bowen:    uh Δgood impact or bad impact
     bowen       Δ looks towards teacher -- >
14   TEA:      I *think it doesn't matt*er
     tea        *looks towards screen projection* looks at Bowen-- >
15             because it doesn't say it has to be bad
16             or good impact
17             Δ©(3.0)©
     bowen     Δlooks down ©lightly taps desk with alternating index and middle fingers©
18   TEA:      no idea?
19             maybe you can say good impact first
20             on you
21             (0.8)
22   Bowen:    I can see uh(.)Δ many countries ¤news¤
     bowen     Δ looks at teacher -- >
     tea                                    ¤shows smile, nods head¤
               ((three lines omitted))
26   TEA:      so how often you read news on yahoo
27   Bowen:    everyday .hh hhh [hahaΔ
     bowen                          Δ looks down -- >
28   Tea:      [everyday what news
29   Bowen:    ©Yeah uhm
     bowen     ©index finger points at smart phone on desk, holds-- >
30        →    Δ(0.5)
     bowen     Δlooks towards teacher -- >
31   Bowen:    e::en everyday ©message©
     bowen                   ©retracts gesture©
32   TEA:      from smart phone
33   Bowen:    Yeah
```

Interactionally, the sequence shows word search accomplished through coordinated talk, gaze, gesture, and material objects. Even before his verbal turn begins, Bowen's hesitation, sustained gaze, and sustained pointing at lines 29–30 project the lexical trouble, making it publicly visible and recruiting the teacher as the more knowledgeable participant in resolving it. The smartphone serves as a material anchor for the search, giving the teacher a

shared object through which to identify what Bowen means. Rather than reformulating the whole utterance or correcting its grammar, the teacher supplies the missing preposition "from" and the sought-for item "smart phone" to complete the noun phrase. This minimal intervention reflects a broader EMI communication pattern: teachers favour content communication and conversational progressivity over explicit grammatical correction. The word search is thus jointly accomplished through embodied conduct, material resources, and the teacher's pedagogical scaffolding.

Solitary Word Search

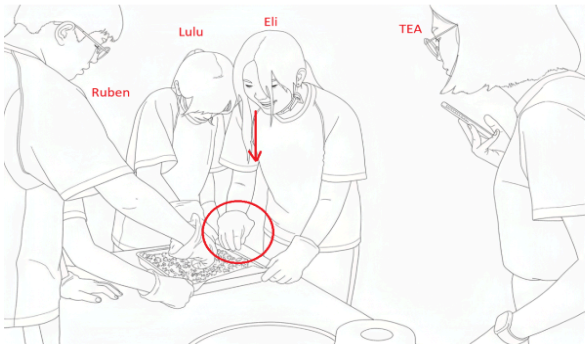
Prior to the extract, the teacher establishes a restaurant role-play in which students were asked to introduce the dish they are making to foreign customers. Standing beside the working station, she consults the questions displayed on her smartphone. While looking down at the smartphone, the teacher asks students to identify the ingredients of the dish (line 1), thereby projecting a response from the class. Students concurrently orient their gaze towards the vegetable rice cake placed on the desk. Lulu first proposes "vegetable meat," followed by Ruben's confirmation of "meat." The teacher's continuer ("hmm mm") accompanied by head nodding acknowledges the contributions while withholding sequence closure, thereby making further responses relevant. Eli subsequently contributes "rice" while directing her gaze towards the dish. The teacher immediately upgrades the formulation to "sticky rice" before pursuing additional ingredients with the follow-up question, "and what else" (line 7), thereby sustaining the ongoing question–answer and role-play sequence.

In line 8, Eli attempts to identify another ingredient. She begins with the hesitation marker "uh" and the deictic expression "this," while simultaneously looking at and pointing towards a mushroom on the rice cake. The deictic term, together with the pointing gesture, establishes a "prospective indexical" reference to a visually available object without yet providing its lexical label (Goodwin, 1996, p. 384). Eli then produces another hesitation marker "u:::m" while displaying a thinking face, maintaining the floor and projecting that lexical retrieval is still underway. Throughout this search phase, she looks at the rice cake and continues searching independently, using embodied conduct to preserve turn progressivity while orienting to the mushroom as the intended referent. Other participants also temporarily stop and withhold visible or vocal entry into the turn-taking. Only after this extended solitary search does Eli explicitly request assistance by code-switching into Chinese with the formulaic search marker, "*zhè shì shénme*" ("what's this," line 11). Unlike the initial hesitation and deictic reference, this wh-question publicly transforms the search from an individual retrieval effort into a collab-

orative activity by explicitly inviting others to supply the missing lexical item. Lulu immediately responds with "mushroom," which Eli repeats before the teacher ratifies the solution with "yes" (line 14).

Extract 2

01	TEA		*what are the ingredients for this dish
	tea		*looks down at smart phone -->
02	Lulu:		Mm + vegetable me[at
	lulu		+looks at rice cake on desk-- >
03	Ruben:		[meat
04	TEA		¤hmm mm¤
			¤nods head¤
05	Eli:		Δrice
	eli		Δlooks at rice cake on the desk-- >
06	TEA		sticky rice
07			and what else
08	Eli:	→	uh ©#this
	eli		©#points at mushroom -- >
09	Eli:		©u:::m
	eli		©displays thinking face -- >
10			mmm
11			©*zhè shì shénme=*
			what's this
	eli		©smile -- >
12	Lulu		mush[room
13	Eli:		[mushroom
14	Tea:		yes



#Fig1 solitary word search

The sequence illustrates a progression from individual lexical search to collaboratively accomplished word retrieval, in which Eli initially exhausts her own interactional resources—including gaze, pointing, deictic reference, and hesitation—before overtly recruiting peer help through a code-switching strategy. Also, Eli's question occurs with a change of facial expression from a thinking face to a smile. The smile indicates that the speaker is increasing the rapport among co-participants after reserving the entitlement to speak during the solitary word search sequence. The teacher's acceptance of the bilingual search marker further demonstrates that the temporary use of the native language is treated as a legitimate interactional resource for maintaining participation and accomplishing lexical retrieval.

Discussion

In EMI classrooms, learners signal lexical difficulty and progressively recruit teachers or classmates for assistance through pointing gestures and other multimodal resources. Consistent with previous findings, teachers offer minimal linguistic support to keep interaction moving and get disciplinary content across. Their acceptance of bilingual resources and light grammatical scaffolding indicates that sustaining participation and content comprehension matter more than explicitly correcting language form (Duran et al., 2022).

Conclusion

This study shows that word searches are collaboratively accomplished. Pedagogically, allowing and recognizing multimodal and bilingual resources can support L2 students in resolving lexical difficulties while maintaining participation and access to subject-matter content.

Appendix. Transcription Notation

The video-recorded data were transcribed following Jefferson's (2004) conventions for conversation analysis. To represent participants' embodied conduct and the temporary trajectory of embodied actions (e.g., gestures and gaze), the transcripts also incorporate Mondada's (2018) conventions for multimodal interaction.

[ ]	Overlapping utterances
=	Latching: when there is no interval between adjacent utterances

(1.0)	Timed silence within or between utterances in tenths of a second
-	A glottal stop, or abrupt cutting off of sound
( )	Uncertain hearing
(( ))	Transcriber's remark
:	Elongated sound
.	Falling intonation, e.g., sentence final
,	Continuing intonation, e.g., phrase final
?	Rising intonation
underline	Stressed sound
↑↓	Marked change in pitch: upward or downward
hhh	Audible outbreath
.hh	Audible inbreath
#	Refers to onset of the figure
* *	Descriptions of the teacher's gaze are delimited between
¤ ¤	Descriptions of the teacher's gestures and embodied actions are delimited between
Δ Δ	Descriptions of the student's gaze are delimited between
© ©	Descriptions of the student's gestures and embodied actions are delimited between
+ +	Descriptions of the other participant's gaze are delimited between
--->	The action described continues across subsequent lines
(.)	Micropause

References

Dressel, D. (2020). Multimodal word searches in collaborative storytelling: On the local mobilization and negotiation of coparticipation. *Journal of Pragmatics, 170*, 37–54. https://doi.org/10.1016/j.pragma.2020.08.010

Duran, D., Kurhila, S., & Sert, O. (2022). Word search sequences in teacher-student interaction in an English as medium of instruction context. *International Journal of Bilingual Education and Bilingualism, 25*(2), 502–521. https://doi.org/10.1080/13670050.2019.1703896

Goodwin, C. (1996). Transparent vision. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Eds.), *Interaction and grammar* (pp. 370–404). Cambridge University Press. https://doi.org/10.1017/cbo9780511620874.008

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. https://doi.org/10.1075/pbns.125.02jef

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

Schegloff, E. A. (2010). Some other "uh(m)"s. *Discourse Processes, 47*(2), 130–174. https://doi.org/10.1080/01638530903223380