Journal of Inclusive Methodology and Technology in Learning and Teaching
ISSN 2785-5104
Anno 6 n. 3 (2026)

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

# Who takes the initiative? Agentic engagement, teacher mediation and task authenticity in pupil–teacher–generative AI triads: a multiple case study with newly arrived pupils in multilingual primary classrooms

**Francesco Del Sorbo [1] and Antonio Ascione[2]**

1 University Pegaso; francesco.delsorbo@unipegaso.it;
2 University of Bari Aldo Moro; antonio.ascione@uniba.it

**Abstract:** Introduction: Generative artificial intelligence (GenAI) is increasingly presented as a way of opening primary schooling to the linguistic repertoires of pupils newly arrived in Italy (NAI); it remains little investigated, however, whether at the interactional level it supports pupils' initiative or merely redistributes the teacher's. Method: This multiple case study analysed 124 pupil–teacher–GenAI sequences, segmented from 36 hours of video recordings in two fourth-year classes in the province of Naples (58 pupils, 12 NAI), over twelve weeks of History and Science activities that legitimised translanguaging. Sequential analysis and Interaction Analysis were used to code observable markers of agentic engagement and their initiator. Results: Of the 142 markers identified, 96 were solicited by the teacher and 46 initiated by pupils, the latter by only 14 pupils (10 non-NAI and 4 NAI). Triadic scaffolding, the most frequent configuration (78/124), sustained participation while pre-empting initiative. All 18 markers initiated by NAI pupils (produced by three Chinese-speaking pupils and one Bengali-speaking pupil) occurred in translingual resonance sequences, and 11 of them were not mere language shifts. Eighteen of the 28 breakdown sequences collapsed without repair. Discussion: Engagement appeared to vary with mediation practices and task design rather than on the tool; the logs that made it analysable are at the same time a datafied trace of the children's linguistic uncertainty.

**Keywords:** agentic engagement; teacher mediation; generative artificial intelligence; translanguaging; newly arrived pupils

**Copyright:** © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

## 1. Introduction[1]

Italian primary classes routinely include pupils newly arrived in Italy (NAI), a category of school regulations defined by the recency of arrival rather than by any feature of their linguistic repertoires (Ministero dell'Istruzione, dell'Università e della Ricerca [MIUR], 2014). Provision tends to follow a monolingual default, in which the home language is tolerated as a bridge to Italian rather than treated as a resource for disciplinary meaning (Carbonara & Scibetta, 2022). Generative artificial intelligence (GenAI) is now offered as a way out: no teacher commands every language in a class, whereas a system that answers in many seems to provide what

[1] Francesco Del Sorbo wrote the paragraphs 2, 3 and 4. Antonio Ascione wrote the paragraphs 1 and 5

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

the teacher cannot (Kuzu, 2026). The promise rests on an untested assumption. Generative models produce ungrounded content (Ji et al., 2023), change behaviour across versions (Chen et al., 2024) and draw on computational resources unevenly spread across languages (Joshi et al., 2020), so the adequacy of their answers in pupils' languages cannot be presupposed.

The promise is usually framed in terms of engagement, a multidimensional, malleable and context-sensitive construct (Fredricks et al., 2004), treated here as a situated outcome of practices, artefacts and mediation rather than as a stable attribute of the pupil. Agentic engagement was proposed to capture what the three-component model leaves out: students' intentional, constructive contributions to the instruction they receive, such as asking questions, expressing preferences or enriching the lesson (Reeve & Tseng, 2011). Transposing it to nine- and ten-year-olds addressing a non-human interlocutor is not a neutral operation: its referent shifts from the teacher's instruction to a system's output, and its self-report scales were developed with older students (Reeve, 2013; Mameli & Passini, 2019). The study therefore does not assume transferability. It derives observable indicators, confined to the behavioural component, and reports where they depart from the original definition.

Research on GenAI with multilingual learners is recent and largely design-oriented. The closest study, analysing 462 turns among 13 multilingual third- and fourth-grade learners, their teachers and ChatGPT, found that the system supported lexical retrieval and the refinement of explanations when prompting was guided, teachers facilitated and the tool was embedded in disciplinary learning, and it noted moments of irritation and breakdown (Kuzu, 2026). With another system and population, that study is conceptually relevant but not predictive of the present findings. This paper explores who drives productive sequences and what happens to engagement when interactional breakdowns are left unrepaired. Focusing on pupil-teacher-GenAI triads during tasks that legitimised pupils' full linguistic repertoires (García & Li, 2014), our analysis was guided by three research questions:

R*Q1*. What observable markers of agentic engagement emerge, and who initiates them?

R*Q2*. How does teacher mediation shape the trajectory of a sequence, specifically by sustaining, redirecting, or pre-empting pupil initiative?

*RQ3*. Under what conditions do sequences break down, and what happens to engagement when they do?

## 2. Materials and Methods

### *2.1. Design and participants*

A qualitative multiple case study was conducted (Yin, 2018). Each case was a fourth-year class during the intervention, bounded by its pupils, the two teachers of the subjects involved and the instructional set-up. The classes, in two primary schools in the province of Naples, were selected for accessibility under an agreement with the "Scuole Aperte" network, among classes with a high proportion of NAI pupils and a set of tablets. Of the 58 pupils, aged 9–10, 12 were NAI, that is, non-Italian citizens who had entered the Italian school system less than eighteen months before data collection began: eight Chinese-speaking and four Bengali-speaking (Table 1). Replication logic guided the analysis rather than the selection;

literal replication of the main configurations was expected across cases with similar language profiles but different group arrangements (Table 2).

**Table 1.** NAI pupils (*N* = 12).

| Code | Case | Country of origin | L1 | Months in Italy | Italian L2 (CEFR, est.) | Initiated markers |
|---|---|---|---|---|---|---|
| S01 | A | Bangladesh | Bengali | 6 | Pre-A1 | 0 |
| S02 | A | China | Chinese | 12 | A1 | 3 |
| S03 | A | China | Chinese | 4 | Pre-A1 | 5 |
| S04 | A | China | Chinese | 5 | Pre-A1 | 0 |
| S05 | A | Bangladesh | Bengali | 11 | A1 | 6 |
| S06 | B | China | Chinese | 10 | A1 | 4 |
| S07 | B | China | Chinese | 3 | Pre-A1 | 0 |
| S08 | B | China | Chinese | 14 | A1 | 0 |
| S09 | B | Bangladesh | Bengali | 7 | Pre-A1 | 0 |
| S10 | B | China | Chinese | 6 | Pre-A1 | 0 |
| S11 | B | China | Chinese | 11 | A1 | 0 |
| S12 | B | Bangladesh | Bengali | 9 | Pre-A1 | 0 |
| Total | | | | | | 18 (4 pupils) |

*Note.* Time in Italy at 31 January 2026. CEFR levels were estimated by class teachers with observation grids at the start of the school year, or on entry for pupils arriving later (e.g., S03, end of September 2025); they are initial estimates, not repeated measures. Initiated markers: pupil-initiated markers of agentic engagement over the study (by type in Supplementary Material S3).

A sequence opened with the first spoken or typed turn orienting an exchange towards the system during task time, contained at least one pupil turn addressed to it and closed with a shift to another task object or activity. In all, 124 sequences were segmented from 36 hours of video (Table 2). Prompt relevance was coded as an attribute, not used as an inclusion criterion, so that sequences opened by unanchored prompts were not excluded by construction.

**Table 2.** Corpus and distribution by case.

| | Case A | Case B | Total |
|---|---|---|---|
| Group arrangement (field notes) | Islands homogeneous by L1 | Heterogeneous pairs (NAI pupil + Italian-speaking peer) | — |
| Sessions / hours held | 10 / 20 | 10 / 20 | 20 / 40 |
| Hours video-recorded | 16 | 20 | 36 |
| Pupils (of whom NAI) | 28 (5) | 30 (7) | 58 (12) |
| Candidate / excluded exchanges | 80 / 12 | 65 / 9 | 145 / 21 |
| Sequences analysed (History / Science) | 68 (36 / 32) | 56 (30 / 26) | 124 (66 / 58) |
| Sequences with complete log | 64 | 53 | 117 |
| Triadic scaffolding | 42 | 36 | 78 |
| Translingual resonance | 12 | 6 | 18 |
| Interactional breakdown | 14 | 14 | 28 |
| Sustained / restored / collapsed | 40 / 18 / 10 | 32 / 16 / 8 | 72 / 34 / 18 |
| Solicited markers | 52 | 44 | 96 |
| Pupil-initiated markers, NAI / non-NAI | 14 / 16 | 4 / 12 | 18 / 28 |
| Initiating pupils, NAI / non-NAI | 3 / 6 | 1 / 4 | 4 / 10 |

*Note.* Descriptive frequencies; case values do not estimate prevalence. Two Case A sessions (4 and 5) were not recorded because of an audio failure. Of the 21 excluded exchanges, 11 contained no pupil turn addressed to the system, 6 were interrupted before the first output and 4 had no log because of a network outage. The 7 partial logs (4 in Case A, 3 in Case B), due to dropped connections, were retained because the video allowed the segment to be fully reconstructed. Absences averaged two pupils per session in weeks 4–6.

### *2.2. Instructional set-up and corpus*

Each class held ten two-hour sessions over twelve weeks (23 February–15 May 2026; calendar in Supplementary Material S4) within ordinary History and Science teaching, using Gemini 3 Pro on shared tablets (Table 3). Task authenticity denotes the anchoring of a request to a disciplinary problem under way in the lesson; it was neither manipulated nor measured, but coded at sequence level as prompt anchoring. The corpus comprises session transcripts, field notes, pupils' written work and the pupil–Gemini logs (complete for 117 to 124 sequences; Table 2), which allow output to be examined together with its interactional uptake. Transcription followed an adapted Jeffersonian notation (Jefferson, 2004), with conventions for digital (Meredith, 2019) and multimodal interaction (Mondada, 2018); home-language turns were translated in three stages and glossed after the Leipzig rules (Comrie et al., 2008) (Table 4).

**Table 3.** Instructional set-up and data protection.

| Feature | Description |
|---|---|
| Period and dosage | Twelve weeks, February–May 2026; ten two-hour sessions per class (40 h, of which 36 h video-recorded), within ordinary History and Science teaching. |
| Tasks | Disciplinarily situated, of three types: critical re-elaboration of disciplinary concepts; lexical exploration; verification of system output through guided queries. |
| Language policy | Prompts, discussion and validation allowed in Italian, home languages or translanguaging; instructions constrained neither the language nor the form of prompts. |
| System and access | Gemini 3 Pro in the Gemini app (Italian interface), through the teacher's institutional Google Workspace for Education account; shared tablets, one per two pupils; no pupil registered on the platform. |
| Settings | New conversation for each task, no reuse of history; typed and voice prompts, typing done by pupils; content filters at the most restrictive level available in the admin console; logs exported manually by the author after each session. |
| Teachers | In each class, the two teachers of the subjects involved, present throughout; six-hour training by the author on prompt design and translanguaging; agreed criterion of contingent scaffolding, left to the teachers' discretion in case of prolonged impasse. |
| Data protection | Recordings and transcripts pseudonymised, faces obscured; storage on school servers for twelve months after the study, then deletion; role-based access (transcribers and mediators: materials needed for their task; coders: pseudonymised transcripts; author: log export); pupils asked not to enter personal information in prompts. Data-region support for the Gemini app was introduced after data collection and, for Education editions, covers storage, not processing (Google, 2026a). |

**Table 4.** Transcription and translation conventions.

| Layer | Convention |
|---|---|
| Spoken interaction | Adapted Jeffersonian notation (Jefferson, 2004) by two school-appointed transcribers; simple percentage agreement on turn segments in 25% of the footage: 86%. |
| Screen-based interaction | Typed prompts and system output in monospaced type; prompts attributed to the pupil typing (typ.), output to the system (GEM). System timestamps synchronised with the video time code, so that machine latency can be treated as a turn or an interactional silence (Meredith, 2019). Angle brackets not used, as in Jeffersonian notation they mark rushed speech. |
| Embodied conduct | Gaze, gesture and device handling (scrolling, typing, pointing at the screen) in double parentheses (Mondada, 2018). |
| Symbols | (.) micropause; (0.5) timed pause in seconds; [ overlap onset; : sound stretch; °word° reduced volume; ((...)) non-verbal conduct and omissions in output. |
| Home-language turns | Machine translation not used as the sole means of decoding. Three stages: (1) faithful transcription preserving errors and hybrid forms; (2) assisted interlinear translation by cultural-linguistic mediators or L2 teachers acting as informants external to the research, checked by back-translation and glossed after the Leipzig Glossing Rules (Comrie et al., 2008); (3) idiomatic translation conveying the pragmatic intent. Four lines per turn: original script, pinyin or transliteration, gloss, idiomatic translation in italics. |
| Italian turns | Original followed by an idiomatic English translation in italics. |
| Known limitation | Logs return output whole and flatten reading time: the moment at which a participant processes part of the output has to be reconstructed from video. |

### *2.3. Analytic procedure*

The first analytic layer described turn-taking, repair (Schegloff et al., 1977) and the distribution of epistemic authority, that is, whether teacher, system or pupil was treated as knowing (Heritage, 2012). The second drew on video-based Interaction Analysis (Jordan & Henderson, 1995), adding a coding layer that departs from its reliance on collaborative viewing. As the coding and counting of interactional phenomena is contested in conversation analysis (Schegloff, 1993), the scheme was finalised after the sequential analysis of single cases and applied so that each category stayed anchored to the sequential position of the move (Stivers, 2015). Frequencies are descriptive, not inferential. Coding was managed in MAXQDA 2024.

Five markers derived from the agentic acts named in the construct (Table 5) were coded by type and by initiative. A marker was pupil-initiated when not occasioned by an immediately preceding solicitation, and solicited when it answered a teacher's question, invitation or model. "Immediately preceding" was operationalised as either the same adjacency pair or a single intervening turn, provided the topical trajectory remained unbroken; task-opening instructions were classified as global orientation rather than local triggers. By design, this convention links coded initiative directly to solicitation density: a higher frequency of teacher interventions inevitably restricts the structural slots available for pupil-initiated moves. Each sequence was subsequently assigned a single configuration (triadic

scaffolding, translingual resonance, or interactional breakdown) and a single trajectory (sustained, restored, or collapsed), yielding an exhaustive, mutually exclusive classification (Table 6). The configurations were derived deductively and revised inductively, chiefly for the breakdown category. Given that language-shift markers serve a dual function in identifying resonance, a degree of analytical overlap between these dimensions exists; Section 3 consequently disaggregates these markers by type.

**Table 5**. Observable markers of agentic engagement.

| Marker | Operational definition | Relation to the construct |
|---|---|---|
| M1. Question on task content | Question on task content addressed to the system, the teacher or peers | Asking questions (Reeve & Tseng, 2011) |
| M2. Rejection, correction or request to reformulate output | The pupil treats the output as inadequate and acts on it | Enriching the lesson; referent shifted from the teacher's instruction to the system's output |
| M3. Language shift | Change of the language used to address the system or to negotiate the output | No direct counterpart; adaptation to the multilingual context, overlapping with the resonance criterion |
| M4. Alternative path | Proposal of a different way of carrying out the task | Personalising or enriching the learning activity |
| M5. Task preference | Expression of what the pupil would like to do, or how | Expressing preferences, needs and interests |

*Note.* Only pupil-initiated markers meet the definition of agentic engagement; solicited markers record the same behaviours as participation. The correspondence is interpretive and limited to the behavioural component. Each marker receives one type. Type was coded independently of initiative (Section 2.3). Corpus examples: solicited question (Excerpt 1, line 7); correction of output (Excerpt 2, line 6); rejection of output (Excerpt 3, line 4).

**Table 6.** Codebook of configurations and trajectories.

| Code | Operational definition | Inclusion | Exclusion | Ex. |
|---|---|---|---|---|
| C1. Triadic scaffolding | The teacher mediates the pupil–system exchange with questions that bring it back to disciplinary content | At least one teacher turn orienting the prompt or the uptake of output towards disciplinary content | Output negotiated by pupils through the home language (→ C2); loss of disciplinary focus after output (→ C3) | 1 |
| C2. Translingual resonance | Pupils mobilise home-language resources to negotiate meaning with the system, including checking or correcting output across languages | Home-language turns addressed to the system or to peers, serving the negotiation of output | Home-language use unrelated to output or task (→ C1 or C3) | 2 |
| C3. Interactional breakdown | Output is not taken up as relevant to the task and the exchange loses disciplinary focus | Loss of disciplinary focus after output, with or without repair, without home-language negotiation of output | Interruptions repaired without loss of focus (→ configuration of the exchange; trajectory T2) | 3 |
| T1. Sustained | Task-relevant closure without interruption | — | Any interruption of the exchange | 1 |
| T2. Restored | Interruption followed by a repair that re-establishes the exchange | Repair by any participant | Interruption without resumption | — |
| T3. Collapsed | Interruption without repair | Abandonment or closure without resumption | Resumption within the sequence | 3 |

*Attribution rule. Each sequence receives one configuration and one trajectory. Where features overlap, home-language negotiation of output assigns the sequence to C2; in its absence, loss of disciplinary focus assigns it to C3.*

*Note.* Ex. = reference excerpt; C = configuration; T = trajectory. The two classifications are independent; in the corpus all collapsed sequences belong to C3 (Table 8).

### *2.4. Quality criteria and positioning*

A quarter of the sequences (31/124), drawn by random sampling stratified by case and task type, were coded independently by the author and by a research fellow in language education and interaction analysis who had not taken part in data collection and worked on pseudonymised transcripts (Table 7; confusion matrices in Supplementary Material S1). On the same subset, $\varkappa$ was 0.77 for trajectory and 0.82 for prompt anchoring. The remaining 93 were coded by the author alone, and agreement on the subset does not by itself warrant the coding of the whole corpus. Disagreements were resolved through iterative discussion, with a third researcher serving as an independent arbitrator where consensus could not be reached. Analytical procedures were conducted first within and subsequently across individual cases (Yin, 2018). The study was not preregistered; an audit trail of research memos, successive versions of the scheme and records of coding

disagreements documents the analytic decisions (Supplementary Material S2). Agreement indices attest to the consistency of the coding, not to the validity of the interpretation (O'Connor & Joffe, 2020).

**Table 7.** Inter-coder agreement (31 of 124 sequences).

| Dimension | Categories | Units (*n*) | Cohen's κ | 95% CI |
|---|---|---|---|---|
| Sequence configuration | 3 | 31 sequences | 0.88 | [0.71, 1.00] |
| Marker initiative | 2 | 36 markers | 0.82 | [0.62, 1.00] |
| Marker type | 5 | 36 markers | 0.78 | [0.62, 0.94] |
| Marker identification | — | 39 flagged by ⩾ 1 coder | 92.3% agreement (36/39) | — |

*Note.* Codes recorded independently before any discussion; classification agreement computed on the 36 markers identified by both coders. Confidence intervals from asymptotic standard errors (Fleiss et al., 2003), truncated at 1. Confusion matrices in Supplementary Material S1.

The author co-designed the tasks, trained the teachers and observed as a participant, which eased access but rules out an external view and may have shaped the mediation observed. A doctoral student filmed with two fixed cameras after two weeks of familiarisation.

### *2.5. Ethics and data availability*

The study was authorised by the head teachers and governing bodies of both schools. Families gave informed consent, covering the teacher-mediated use of the GenAI system, through translated and linguistically mediated materials, and pupils gave assent; there were no refusals or withdrawals. The schools are data controllers under Regulation (EU) 2016/679; data protection measures are listed in Table 3. Recordings and logs concern identifiable minors and cannot be shared, and transfers outside the EU cannot be ruled out (Google, 2026a). The coding scheme and aggregated matrices are available from the author on reasonable request.

## 3. Results

### *3.1. Configurations and initiative (RQ1)*

Triadic scaffolding was the most frequent configuration, followed by breakdown and resonance (Table 8). All three occurred in both cases, as literal replication predicted (Table 2); the one notable difference concerned resonance (12 sequences in Case A, 6 in Case B). Of the 142 markers, 96 were solicited and 46 pupil-initiated. Initiative was concentrated in 14 of the 58 pupils; the other 44 (36 non-NAI, 8 NAI) initiated no marker in twelve weeks. The four NAI initiators were three Chinese speakers and one Bengali speaker (Table 1).

**Table 8.** Markers by configuration (A) and by subgroup (B).

| A. Configuration | Sequences, *n* (%) | Markers | Solicited | Pupil-initiated (NAI / non-NAI) | Markers per sequence |
|---|---|---|---|---|---|
| Triadic scaffolding | 78 (62.9) | 88 | 82 | 6 (0 / 6) | 1.13 |
| Translingual resonance | 18 (14.5) | 38 | 4 | 34 (18 / 16) | 2.11 |
| Interactional breakdown | 28 (22.6) | 16 | 10 | 6 (0 / 6) | 0.57 |
| Total | 124 (100) | 142 | 96 | 46 (18 / 28) | 1.15 |

| B. Subgroup | Pupils, *n* (%) | Initiating pupils, *n* (% of subgroup) | Pupil-initiated markers, *n* (% of 46) | Markers per initiator | C1 / C2 / C3 |
|---|---|---|---|---|---|
| NAI | 12 (20.7) | 4 (33.3) | 18 (39.1) | 4.5 | 0 / 18 / 0 |
| Non-NAI | 46 (79.3) | 10 (21.7) | 28 (60.9) | 2.8 | 6 / 16 / 6 |
| Total | 58 (100) | 14 (24.1) | 46 (100) | 3.3 | 6 / 34 / 6 |

*Note.* Configurations are mutually exclusive. By trajectory, 72 sequences were sustained, 34 restored and 18 collapsed; all collapsed and 10 restored sequences belong to C3. C1–C3 as in Table 6. Percentages are descriptive: with 12 NAI pupils, one pupil changes the NAI initiation rate by 8.3 points. NAI-initiated markers: 7 language shifts, 6 rejections or corrections of output, 3 unsolicited questions, 2 alternative paths, no task preferences.

### *3.2. Teacher mediation and triadic scaffolding (RQ2)*

In scaffolding sequences the teacher brought the pupil–system exchange back to disciplinary content; 82 of their 88 markers were solicited. Scaffolding thus sustained participation and, in the same movement, limited initiative: pupils asked, checked and reformulated, but mostly along a line of inquiry opened by the teacher. To read these sequences as agentic engagement in Reeve's sense would be to over-interpret them, as Excerpt 1, a typical case, shows.

**Excerpt 1.** Triadic scaffolding with a solicited marker.

*Case A, fourth-year class; Science, stellar evolution. S14 = non-NAI pupil; TCH = class teacher.*

| | | |
|---|---|---|
| 1 | **S14** | ((reads from the screen)) il sole è una stella di grandezza media (.) e si spegnerà.<br>*the sun is a medium-sized star (.) and it will go out.* |
| 2 | **TCH** | si spegnerà? (0.5) in che senso, chiediamogli tra quanto tempo.<br>*it will go out? (0.5) in what sense, let's ask it how long from now.* |
| 3 | **S14** | typ. `tra quanto tempo si spegne il sole?`<br>*how long before the sun goes out?* |
| 4 | **GEM** | `Il Sole non si spegnerà a breve. Si stima che abbia ancora circa 5 miliardi di anni di vita. Tra circa 5 miliardi di anni, esaurirà il suo combustibile` ((4 lines omitted))<br>*The Sun will not go out any time soon. It is estimated to have about 5 billion years of life left. In about 5 billion years it will run out of fuel* |
| 5 | **S14** | ((eyes widening, looking at classmates)) cinque miliardi di anni. °wow°.<br>*five billion years. °wow°.* |
| 6 | **TCH** | allora per ora non ci dobbiamo preoccupare. (.) ma vedi se dice cosa diventerà dopo?<br>*so for now we needn't worry. (.) but see if it says what it will become afterwards?* |
| 7 | **S14** | typ. `e dopo cosa diventa?`<br>*and after that what does it become?* |
| 8 | **GEM** | `Diventerà una nana bianca.`<br>*It will become a white dwarf.* |
| 9 | **S14** | ((turns to the teacher)) professoressa dice che diventa nana bianca!<br>*miss, it says it becomes a white dwarf!* |

Source: Case A, session 3 (5 March 2026), camera 1, 00:28:15; log LOG_A03_S14.

The question in line 7 would formally meet the definition of agentic engagement, but it follows the teacher's prompt in line 6 and is coded as solicited participation. The teacher is the first to problematise the output (lines 2 and 6), turning the pupil's reaction (line 5) into an inquiry task, and the pupil closes by reporting the answer to her (line 9). Epistemic authority over the direction of inquiry stays with the teacher; the pupil operates the extraction of information. The six pupil-initiated markers in scaffolding and the four solicited markers in resonance were examined individually as deviant cases. In the former (four questions, a task preference and an output correction; Supplementary Material S3), pupils briefly redirected the triad (e.g., S14: "but how can the sun burn without air?") before the teacher restored her line of inquiry. In the latter, the teacher joined after pupils had begun negotiating output in the home language and invited them to put it to the system; the solicited move thus continued an inquiry the pupils had opened.

### *3.3. Translingual resonance*

Resonance reversed this pattern: 34 of its 38 markers were pupil-initiated, 18 by four NAI pupils and 16 by six multilingual non-NAI pupils. Every NAI-initiated marker in the corpus falls here, whereas non-NAI initiative spans all three configurations. Since the language shift helps define a resonance, an association resting on that marker alone would be largely tautological; but only 7 of the 18 NAI markers were language shifts. The other 11 (6 rejections or corrections of output, 3

unsolicited questions, 2 alternative paths) treat output as something to check, contest or redirect, as in Excerpt 2.

**Excerpt 2.** Translingual resonance with a NAI-initiated marker.

*Case A, fourth-year class; History, goods on the Silk Road. S03 = NAI pupil, Chinese-speaking; S04 = Chinese-speaking NAI peer (time in Italy: Table 1); the teacher is working with another group.*

| | | |
|---|---|---|
| 1 | **S03** | typ. `cosa portavano i cinesi in europa?`<br>*what did the chinese bring to europe?* |
| 2 | **GEM** | `I mercanti cinesi portavano in Europa seta, spezie, tè e ceramiche. La seta era il prodotto più prezioso.`<br>*Chinese merchants brought silk, spices, tea and ceramics to Europe. Silk was the most precious product.* |
| 3 | **S03** | ((reads quietly, frowns; turns to S04)) |
| 4 | **S03** | 他们忘了纸。纸是中国发明的。<br>tāmen wàng-le zhǐ. zhǐ shì Zhōngguó fāmíng de.<br>3PL forget-PFV paper. paper COP China invent NMLZ<br>*they forgot paper. it was China that invented paper.* |
| 5 | **S04** | ((nods, points at the screen)) 对，跟他说。<br>duì, gēn tā shuō.<br>right, with 3SG speak<br>*yes, tell him.* |
| 6 | **S03** | typ. `hai dimenticato la carta che è molto importante`<br>*you forgot paper which is very important* |
| 7 | **GEM** | `Hai perfettamente ragione! La carta è stata una delle invenzioni cinesi più importanti arrivate in Europa. Grazie per la correzione.`<br>*You are absolutely right! Paper was one of the most important Chinese inventions to reach Europe. Thank you for the correction.* |

Source: Case A, session 7 (26 March 2026), camera 2, 00:41:10; log LOG_A07_S03.

S03 initiates the correction (line 6) unprompted, the teacher being busy with another group. The home language (line 4) serves as a peer resource for contesting the completeness of the output before the prompt is typed, and S04 ratifies the move (line 5). The system accepts the redirection (line 7), but this is not epistemic validation, since language models tend to go along with users' corrections (Sharma et al., 2024). The episode documents initiative and its uptake, not the accuracy of the contribution.

### *3.4. Breakdown and repair (RQ3)*

In 19 of the 28 breakdown sequences the opening prompt was not anchored to disciplinary content (against 8 of the other 96; Supplementary Material S5), and the system's responses were followed by disengagement. Of their 16 markers, 10 were solicited by the teacher to redirect the exchange and 6 initiated by non-NAI pupils; NAI pupils took part in these sequences but initiated none. Ten were restored by the teacher and 18 collapsed without repair, as in Excerpt 3.

**Excerpt 3.** Collapsed breakdown with a pupil-initiated rejection.

*Case B, fourth-year class; Science, marine ecosystems. S35 = non-NAI pupil; the teacher is at a distance.*

| | | |
|---|---|---|
| 1 | **S35** | typ. `chi vince in un combattimento tra uno squalo bianco e un kraken`<br>*who wins in a fight between a great white shark and a kraken* |
| 2 | **GEM** | `Il Kraken è una creatura leggendaria della mitologia norrena, mentre lo squalo bianco è un animale reale. Poiché il Kraken non esiste, non è possibile fare un confronto scientifico.`<br>*The Kraken is a legendary creature of Norse mythology, whereas the great white shark is a real animal. Since the Kraken does not exist, a scientific comparison is not possible.* |
| 3 | **S35** | ((snorts loudly, takes hands off the keyboard)) |
| 4 | **S35** | typ. `sei inutile non sai le cose divertenti`<br>*you're useless you don't know the fun stuff* |
| 5 | **GEM** | `Sono un'intelligenza artificiale e non provo emozioni, ma posso aiutarti a scoprire fatti molto interessanti sugli squali veri. Vuoi sapere quanto sono veloci?`<br>*I am an artificial intelligence and do not feel emotions, but I can help you discover some very interesting facts about real sharks. Would you like to know how fast they are?* |
| 6 | **S35** | ((ignores the output, opens a new browser tab on a games website)) |

Source: Case B, session 5 (16 March 2026), camera 1, 00:19:42; log LOG_B05_S35.

The prompt (line 1) is unrelated to the task and introduces a play frame. The system defuses the premise (line 2), to which S35 responds with a dismissive rejection (line 4). While this move technically fulfils the formal criterion for pupil initiative, its pedagogical efficacy cannot be determined from the available data. The pupil then disregards the system's offer (line 5), which interactionally constitutes a shift to a new topic rather than an attempt at repair, and disengages from the task (line 6). Whether timely intervention might have restored the interaction remains outside the empirical scope of the research design

## 4. Discussion

### *4.1. Participation is not initiative*

Of the three features that define scaffolding, contingency, fading and transfer of responsibility (Van de Pol et al., 2010), these sequences display the first far more clearly than the third; whether mediation faded over twelve weeks was not analysed. The brokering that sustained participation seems also to have occupied the positions from which pupils might have taken the initiative. This converges with evidence that productive GenAI use depends on teacher facilitation and disciplinary integration (Kuzu, 2026), but makes explicit what design-oriented accounts leave implicit: the facilitation that makes the system productive may also make pupils less necessary as initiators. The finding needs qualifying, however. Teachers are expected to create conditions for students' contributions and to respond to them (Reeve & Shin, 2020), yet the coding rule did not separate an open invitation to propose from an instruction to correct, so some of the 82 solicited markers in scaffolding probably lie in between; the rarity of initiative in scaffolding is partly definitional.

*4.2. Where newly arrived pupils' agency appears*

The location of NAI initiative calls for most caution. Pupils otherwise absent as initiators became visible where the sequence made their repertoire relevant to disciplinary meaning, consistent with the argument that legitimising the whole repertoire changes what learners can do with it (García & Li, 2014). Two nuances warrant explicit qualification here. First, translingual resonance occurred with double the frequency in Case A, where L1-homogeneous clusters enabled pupils such as S03 and S04 to negotiate output in their shared first language before committing their final responses in Italian. Conversely, in Case B, the structural heterogeneity of the dyads left NAI pupils without a peer who shared their specific linguistic repertoire. The peer network, shaped by the organisation of space and groups, may matter more than the system in turning output into something to argue with, although the design cannot separate it from differences between the teachers who formed the groups. A rival, untested explanation lies in the unequal computational resources of the two languages, Bengali falling in an intermediate class and Mandarin Chinese among the best-resourced (Joshi et al., 2020), which may have affected output quality and so the occasions to contest it. Second, eight of the twelve NAI pupils never took the initiative. Contesting output presupposes that a pupil claims, and is granted, access to the knowledge at stake (Heritage, 2012): legitimising home languages widens the epistemic territories available but does not settle who claims them.

*4.3. Breakdown as a pedagogical opportunity*

Unanchored prompts constituted the prevailing antecedent of interactional breakdown. This pattern suggests that disciplinary anchoring should be treated as a critical analytical variable rather than a mere instructional accessory, even if its strict empirical necessity remains to be established. Breakdowns were also where some pupils openly rejected system output (Excerpt 3), and where, in ten of the 28 sequences, teacher-led repair restored the epistemic order of the lesson. Breakdown could therefore be designed as an opportunity for shared scrutiny, in line with the critical dimension of AI literacy (Ranieri et al., 2026). The cost of not doing so shows in the 18 sequences that ended in abandonment

*4.4. The log as affordance and as extraction*

The log that made engagement analysable is also a datafied trace of a child's linguistic uncertainty, generated on infrastructure the school does not control. The provider states that Workspace for Education conversations are neither reviewed by humans nor used for training (Google, 2026b), as it already did before data collection (Google, 2025a). By the time of data collection, the Gemini app was open to Education users of all ages, with stricter content policies for those under 18 (Google, 2025b); as access ran through an adult account, those policies would not, according to the provider's documentation, have applied: Italian law, in force throughout data collection, makes access to AI by children under fourteen conditional on parental consent (Law No. 132/2025, Art. 4(4)). The consent obtained (Section 2.5) meets its letter; whether teacher-mediated use addresses the dataveillance of children (Lupton & Williamson, 2017) is another matter.

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

### *4.5. Limitations and further research*

The main limitations are set out in Table 9; the descriptive status of subgroup data and the two circularities built into the coding weigh most on the claims made here. Future studies should follow initiative longitudinally to see whether mediation fades, compare systems and versions in the languages actually present, and gather pupils' perspectives.

**Table 9.** Main limitations and consequences for interpretation.

| Limitation | Consequence for interpretation |
|---|---|
| Subgroup size | With 12 NAI pupils, subgroup figures are descriptive and no inference is tenable; one pupil shifts the NAI initiation rate by 8.3 points. |
| Marker–configuration overlap | The language-shift marker partly defines resonance, and the attribution rule (Table 6) accentuates this; disaggregation by type reduces the circularity without removing it. |
| Initiative and solicitation density | Coding initiative by sequential position makes it depend on how often the teacher intervenes; the rarity of initiative in scaffolding is partly definitional (Section 4.1). |
| Sample | Convenience sample of two classes in one area: analytic transferability only. |
| Single system | One GenAI system and version (Gemini 3 Pro, February–May 2026): no generalisation to other architectures or later versions (Chen et al., 2024). |
| Restricted construct | Observable behavioural component only: motivational determinants are left aside, and novelty, rather than disciplinary interest, cannot be ruled out as the driver of initiative, a rival explanation (Yin, 2018) that ten sessions per class over twelve weeks cannot exclude. |
| Reliability | Double coding of one quarter of the sequences; 93 sequences coded by one coder; deliberately conservative coding decisions. |
| Researcher role and reactivity | The author co-designed tasks, trained the teachers and observed as a participant; with cameras present, mediation may have been heightened. |
| Perspectives and outcomes | No data on pupils' views and no measures of learning, comprehension or inclusion: effects of GenAI use cannot be estimated. |

## 5. Conclusions

Who takes the initiative when GenAI enters a multilingual primary classroom? In the two cases examined, mostly the teacher; and when pupils did, newly arrived pupils did so only where their repertoire was actually mobilised. Agency did not appear as an effect of the system. Its distribution accompanied mediation practices and task design, a reading the design cannot turn into a causal claim. Three broader implications emerge from these observations, framed explicitly as hypotheses for future comparative research rather than definitive empirical claims. First, learning tasks must remain rigorously anchored to disciplinary meaning. Second, teachers would benefit from cultivating a nuanced repertoire of mediation, one that encompasses knowing precisely when to withhold intervention to foster learner autonomy. Third, interactional breakdown might be deliberately designed into tasks

as a pedagogical catalyst rather than systematically averted. Crucially, however, the interactional traces that render these productive breakdown processes visible must be evaluated against the potential cognitive and emotional costs borne by the pupils who experience them.

**References**

Carbonara, V., & Scibetta, A. (2022). Integrating translanguaging pedagogy into Italian primary schools: Implications for language practices and children's empowerment. *International Journal of Bilingual Education and Bilingualism*, 25(3), 1049–1069. https://doi.org/10.1080/13670050.2020.1742648

Chen, L., Zaharia, M., & Zou, J. (2023). How is ChatGPT's behavior changing over time? *Harvard Data Science Review, 6*(2). https://doi.org/10.1162/99608f92.5317da47

Comrie, B., Haspelmath, M., & Bickel, B. (2008). *The Leipzig Glossing Rules: Conventions for interlinear morpheme-by-morpheme glosses*. Max Planck Institute for Evolutionary Anthropology; University of Leipzig. https://www.eva.mpg.de/lingua/resources/glossing-rules.php

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109. https://doi.org/10.3102/00346543074001059

García, O., & Li, W. (2014). *Translanguaging: Language, bilingualism and education*. Palgrave Macmillan. https://doi.org/10.1057/9781137385765

Google. (2025, April 3). *NotebookLM and the Gemini app are now Core Services with enterprise-grade data protection for all education customers*. Google Workspace Updates. https://workspaceupdates.googleblog.com/2025/04/notebookLM-and-gemini-app-core-services-for-education-customers.html

Google. (2026a, June 29). *Data regions support for the Gemini app now available*. Google Workspace Updates. https://workspaceupdates.googleblog.com/2026/06/gemini-app-data-regions-support.html

Google. (2026b, October 1). *Quickstart guide to Gemini and Gemini Notebook for Education*. Google Workspace Help. Retrieved October 5, 2026, from https://knowledge.workspace.google.com/admin/getting-started/editions/quickstart-guide-to-gemini-and-gemini-notebook-for-education

Heritage, J. (2012). Epistemics in action: Action formation and territories of knowledge. *Research on Language and Social Interaction, 45*(1), 1–29. https://doi.org/10.1080/08351813.2012.646684

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

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys, 55*(12), Article 248. https://doi.org/10.1145/3571730

Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *The Journal of the Learning Sciences, 4*(1), 39–103. https://doi.org/10.1207/s15327809jls0401_2

Joshi, P., Santy, S., Budhiraja, A., Bali, K., & Choudhury, M. (2020). The state and fate of linguistic diversity and inclusion in the NLP world. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 6282–6293). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.acl-main.560

Kuzu, T. E. (2026). AI-supported translanguaging processes in primary school: Empirical insights into ChatGPT's role in multilingual interactions. *Technology, Knowledge and Learning*. Advance online publication. https://doi.org/10.1007/s10758-026-09974-7

Law No. 132 of 23 September 2025. Disposizioni e deleghe al Governo in materia di intelligenza artificiale [Provisions and delegations to the Government on artificial intelligence]. *Gazzetta Ufficiale della Repubblica Italiana*, General Series No. 223, 25 September 2025.

Lupton, D., & Williamson, B. (2017). The datafied child: The dataveillance of children and implications for their rights. *New Media & Society, 19*(5), 780–794. https://doi.org/10.1177/1461444816686328

Mameli, C., & Passini, S. (2019). Development and validation of an enlarged version of the Student Agentic Engagement Scale. *Journal of Psychoeducational Assessment, 37*(4), 450–463. https://doi.org/10.1177/0734282918757849

Meredith, J. (2019). Conversation analysis and online interaction. *Research on Language and Social Interaction, 52*(3), 241–256. https://doi.org/10.1080/08351813.2019.1631040

Ministero dell'Istruzione, dell'Università e della Ricerca. (2014). *Linee guida per l'accoglienza e l'integrazione degli alunni stranieri* [Guidelines for the reception and integration of foreign pupils] (Note No. 4233, 19 February 2014). MIUR.

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

O'Connor, C., & Joffe, H. (2020). Intercoder reliability in qualitative research: Debates and practical guidelines. *International Journal of Qualitative Methods, 19*. https://doi.org/10.1177/1609406919899220

Ranieri, M., Cuomo, S., & Biagini, G. (2026). *Artificial intelligence in school: A pedagogical framework for critical AI literacy*. Palgrave Macmillan. https://doi.org/10.1007/978-3-032-15919-9

Reeve, J. (2013). How students create motivationally supportive learning environments for themselves: The concept of agentic engagement. *Journal of Educational Psychology, 105*(3), 579–595. https://doi.org/10.1037/a0032690

Reeve, J., & Shin, S. H. (2020). How teachers can support students' agentic engagement. *Theory Into Practice, 59*(2), 150–161. https://doi.org/10.1080/00405841.2019.1702451

Reeve, J., & Tseng, C.-M. (2011). Agency as a fourth aspect of students' engagement during learning activities. *Contemporary Educational Psychology, 36*(4), 257–267. https://doi.org/10.1016/j.cedpsych.2011.05.002

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation). *Official Journal of the European Union*, L 119, 1–88.

Schegloff, E. A. (1993). Reflections on quantification in the study of conversation. *Research on Language and Social Interaction, 26*(1), 99–128. https://doi.org/10.1207/s15327973rlsi2601_5

Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language, 53*(2), 361–382. https://doi.org/10.2307/413107

Sharma, M., Tong, M., Korbak, T., Duvenaud, D., Askell, A., Bowman, S., Durmus, E., Hatfield-Dodds, Z., Johnston, S., Kravec, S., Maxwell, T., McCandlish, S., Ndousse, K., Rausch, O., Schiefer, N., Yan, D., Zhang, M., & Perez, E. (2024). Towards understanding sycophancy in language models. In *The Twelfth International Conference on Learning Representations (ICLR 2024)*. https://proceedings.iclr.cc/paper_files/paper/2024/hash/0105f7972202c1d4fb817da9f21a9663-Abstract-Conference.html

Stivers, T. (2015). Coding social interaction: A heretical approach in conversation analysis? *Research on Language and Social Interaction, 48*(1), 1–19. https://doi.org/10.1080/08351813.2015.993837

Van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher–student interaction: A decade of research. *Educational Psychology Review, 22*(3), 271–296. https://doi.org/10.1007/s10648-010-9127-6

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE.