

# MULTIMODAL MEDIATION IN EAL TEACHER EDUCATION

## Scaffolding meaning-making and interaction through picture books

ELISA BERTOLDI  
UNIVERSITÀ DEGLI STUDI DI PADOVA

**Abstract** - This article explores the multimodal mediation strategies employed by student teachers during picturebook read-aloud performances in English as an Additional Language (EAL). The data for the present study was gathered during a postgraduate course at the University of Udine (Italy). Using Multimodal Interaction Analysis as an analytical framework, the study focuses on two self-video recordings of student teachers' read-aloud performances and employs an innovative approach to data visualization and analysis through TMA software to generate graphs for the visual representations of modal configurations. The article discusses the co-deployment of various semiotic resources, including spoken language, gestures, facial expressions, gaze direction, orchestrated by student teachers to scaffold meaning-making and engage an imagined audience of children. The findings reveal the central role of spoken language and prosodic features in student teachers' mediation strategies, these are complemented by the strategic use of visual modes in specific sequences. The results emphasize the need for developing student teachers' multimodal awareness and reflective practices for picturebook mediation. The study highlights also the value of Multimodal Interaction Analysis as a tool for understanding and assessing multimodal practices and offering teacher educators insights into designing activities that promote effective multimodal mediation strategies.

**Keywords:** Mediation; Picturebook read-alouds; EAL teacher education; Multimodal Interaction Analysis; Modal Configuration

## 1. Introduction

Picturebooks are a form of children's literature that rely on the interdependence of words and images across double-spread pages, where verbal and visual elements work together to construct and convey meaning (Nikolajeva and Scott 2006; Nodelman 1988). The importance of picturebooks in additional language education contexts is well-supported by research studies. For learners of English as an Additional Language<sup>1</sup> (from now on EAL), these multimodal texts have been valued as powerful tools for fostering interaction and promote communicative development (see, among others, Ellis 2018; Ellis and Brewster 2014; Ellis and Mourão 2021; Fleta 2019; Masoni 2019). They can provide opportunities to infer meaning, develop comprehension strategies, and expand communicative competence in context.

The read-aloud performance is the context in which a picturebook is shared by adults with an audience of children and it requires a process of mediation. Mediation involves the adult presenting the verbal text, the illustrations, and the peritextual features of the picturebook, facilitating children's access to and interaction through the contents of

<sup>1</sup> The term English as an Additional Language is used in this study to identify English as non-native language for the student teachers involved and to acknowledge the plural identity of English in their linguistic repertoire.

the picturebook transforming the read-aloud into a collaborative construction of meaning. Recent research has emphasized the fundamental role of the adult (i.e., teachers and educators) as a mediator during read-aloud performances (Ellis and Mourão 2021; 2022). As Masoni (2019) highlights, the teacher as mediator aims “to make the children understand the story, enjoy it, have fun, and take the message home,” while collaboratively constructing meaning with the audience (p. 83). A teacher-mediator should design, plan, and manage inclusive and interactive read-aloud performances, tailoring the experience to the audience's developmental and linguistic needs. Teacher-mediators should guide children through the pages of the picturebook by asking questions, encouraging critical thinking, and fostering participation through verbal and non-verbal means. They should scaffold children's comprehension and language production by reformulating or extending their responses and drawing attention to linguistic and visual features of the picturebook.

Mediation is a crucial skill for additional language teachers. As Bland (2019) points out, student teachers require in-depth guidance in extending their own literary competence, visual literacy, and critical literacy to fully leverage the affordances of picturebooks. These skills are essential for supporting young learners' responses to picturebooks. However, read-aloud performances are inherently multimodal events, where the teacher's role as mediator extends beyond merely reading the verbal text aloud. To engage children in meaningful communicative interaction, mediators should be able to orchestrate various modes, such as gestures, facial expressions, and body movements.

Despite the recognized importance of picturebooks read-aloud sessions in fostering communicative development in EAL, university teacher education courses often provide limited opportunities for student teachers to develop multimodal mediation skills. As a result, little research has been conducted on teacher education practices for promoting the development of multimodal mediation skills. The present article addresses this issue by focusing on a practical experience that allowed a group of student teachers at the University of Udine (Italy) to engage in picturebook read-aloud performances to develop their multimodal mediation skills.

Specifically, the following research questions are addressed:

“What modal configurations do student teachers orchestrate in their EAL read-aloud performances to scaffold meaning-making and engage an imagined audience of children?”

“What tools does Multimodal Interaction Analysis offer to understand and assess the co-deployment of modes in EAL read-aloud performances?”

The context of this study is the Primary Education degree program at the University of Udine (Italy), which prepares students for teaching in nursery and primary schools. Specifically, the research was conducted during English Laboratory V<sup>2</sup>, a 10-hour course designed for student teachers in the final year of their postgraduate degree. The course provided student teachers with an opportunity to put theoretical knowledge into practice, developing their skills in picturebook multimodal mediation for EAL education. The article discusses the multimodal mediation strategies student teachers employ to create engaging and interactive read-aloud sessions for EAL education with children. It also provides teacher educators with ideas for designing activities that highlight the

<sup>2</sup> The course at the core of this study, *English Laboratory V*, was held in November 2022 and taught by Prof. Maria Bortoluzzi. The author conducted the study during this course as part of a larger PhD project, using it as a context to investigate multimodal mediation interactions in picturebook mediation for EAL.

multimodal nature of read-aloud performances and support student teachers in developing effective mediation strategies.

Section 2 of the article outlines the existing literature on the use of picturebooks in teacher education programs. Section 3 describes the specific context of the study, and the data collection process. Section 4 introduces Multimodal Interaction Analysis (from now on MIA) (Norris 2004, 2019, 2020) as the analytical framework used to examine modal configurations in student teachers' EAL read-aloud performances. Section 5 presents the analysis of modal configurations in two specific video recordings, focusing on the co-deployment of various semiotic modes of scaffolding meaning-making and engaging their imagined audience. Section 6 discusses the multimodal mediation strategies employed by student teachers and the implications of the findings for teacher education.

## 2. Picturebooks and their mediation in teacher education courses

The body of work exploring how children's picturebooks can be used in teacher education courses is small but growing. Daly and Blakeney-Williams (2015) documented the use of picturebooks in teacher education in Australia across a range of curriculum areas including science, visual arts, drama, mathematics, and literacy. Their findings showed the potential of picturebooks to increase teachers' content knowledge, developing their visual analysis skills, and fostering discussion on social and cultural topics. Hartmann and Hélot (2021) conducted a study with French pre-service teachers to explore how a trilingual picturebook (i.e., French, German, and Alsace) could be used to develop awareness of biliteracy and multilingualism. Hoffmann's (2021) research work with German preservice teachers and picturebooks also showed how the reading aloud of picturebooks supported the teachers' awareness of how literacy develops. Daly and Short (2022) worked with nine USA preservice teachers and showed how working with bilingual picturebooks from around the world led to the development of a critical language awareness among the participants. These studies demonstrate that picturebooks are increasingly valued in teacher education around the world for their flexibility and potential to foster content learning, pedagogical knowledge, and cultural awareness.

Alongside this growing body of literature, the Fostering Dialogue<sup>3</sup> annual conference is an additional opportunity for educators, researchers, and practitioners to discuss innovative approaches to teaching children's literature at the university level (Campagnaro and Goga, 2022). The Fostering Dialogue conference aims at exploring the use of various forms of children's literature, including picturebooks, in tertiary education and emphasizes its importance for engaging learners, promoting dialogue, and developing multiliteracies across diverse educational settings.

In recent years, picturebooks have been increasingly introduced in EAL teacher education. Several projects and initiatives exemplify this shift, developing innovative approaches to integrating picturebooks into teacher training practices. For instance, the PEPELT<sup>4</sup> (Picturebooks in European Primary English Language Teaching) project provides resources for educators to explore the potential of picturebooks in fostering language acquisition, critical thinking, and cultural awareness among young learners. Through webinars, selected resources, and expert insights, PEPELT supports teachers in selecting and using picturebooks for teaching English in primary education. The European

<sup>3</sup> <https://www.hvl.no/en/collaboration/project/green-dialogues/>

<sup>4</sup> <https://pepelt21.com/>

Erasmus+ ICEPELL<sup>5</sup> project (Intercultural Citizenship Education through Picturebooks in Early English Language Learning) emphasizes the role of picturebooks in promoting intercultural understanding and citizenship education. The activities and resources created for the project are aimed at encouraging teachers to engage with selected picturebooks that can foster children's reflection on cultural diversity, promote the development of empathy, and the reflection on global issues. Finally, the Let's Tell a Tale and the TALEs<sup>6</sup> projects of the University of Udine (Italy), highlight the importance for pre-service teacher to engage in read-aloud performances for enhancing children's communicative development in EAL while addressing broader educational goals, such as ecoliteracy and environmental education (Bertoldi 2024; Bortoluzzi *et al.* 2022).

### 3. Research context and data

English Laboratory V of the Primary Education degree program at the University of Udine (Italy) combines theoretical exploration and practical activities to engage student teachers in reflecting on picturebook mediation as a communicative practice for EAL education with children. Lessons focus on the exploration of picturebooks as a form of children's literature, on the reflection on their potential use in additional language education, and on the design of read-aloud performances in EAL. Tasks, group discussions, hands-on activities and collaborative activities encourage student teachers to identify and reflect on the multimodal features of picturebooks and their implications for children's understanding and engagement in communicative interaction. A selection of picturebooks is made available, allowing student teachers to read them, analyse their features and choose one. By working in pairs or small groups, students identify and discuss multimodal elements in picturebooks and their potential use with young learners in EAL education contexts. The course also includes practical sessions where student teachers engage with the design of read-aloud performances. Through learning-by-doing activities, they practice using semiotic resources such as voice modulation, gestures, gaze, and book handling to enhance children's comprehension and engagement. At the end of the course, student teachers perform in pairs or small groups a read-aloud of a picturebook in front of the whole class. The decision to have student teachers work in pairs or small groups of three is related to one of the aims of English Laboratory V: to enhance the students' language proficiency as future teachers of English in Italian nursery and primary schools. Pair and group work offer participants opportunities to interact and communicate using the target language while carrying out their coursework. By engaging in joint read-aloud planning and discussion in English, students can use and develop their communicative skills, particularly in the context of language education, classroom management and interaction. The data for the present study were collected in November 2022. Seventy-eight students attended the course and all of them completed their coursework and gave final presentations in small groups. At the end of the course, students were invited to voluntarily share the video recording or their read-alouds for the research purposes of this study. Out of the entire cohort, 25 students, comprising 10 groups, volunteered to share their video recordings for research purposes. Each recording was produced by a distinct group of student teachers, with some working in pairs and others in groups of three. The

<sup>5</sup> <https://icepell.eu/index.php/about-2/>

<sup>6</sup> <https://yell.uniud.it/storytelling>

participants used their smartphones to record the videos and shared them via Google Drive.

By analysing the self-video recordings of two groups of students, the present study aims to examine the mediation strategies employed by the student teachers in their read-aloud performances. Two self-recorded videos were selected for this article. The selection was based on their similar duration, the equal number of participants (two in each video), and their consistent perspective: a front-facing, eye-level view of the student teachers framed as a medium shot (Bateman *et al.* 2017). The study design and data gathering follow the ethical guidelines of the University of Udine. Consent forms were submitted to the student teachers as they enrolled as volunteers sharing their video recordings and questionnaires. The first video (VR1), lasting 4 minutes and 11 seconds, was recorded by TE and SG, who read aloud *Same Same but Different* by Jenny Sue Kostecki-Shaw (2011), a picturebook about the exchange of letters between two pen pals, an Indian boy and an American boy, which explores cultural differences and similarities. TE and SG take turns voicing the characters, with TE interpreting the Indian boy and SG the American boy. In their performance, they hold the picturebook together and take turns reading aloud the written text in each double-spread page. The second video (VR2), lasting 4 minutes and 19 seconds, involves students BG and CA performing a read-aloud of *Aaaarrgghh, Spider!* by Lydia Monks (2013). The picturebook talks about a female spider who repeatedly tries to be accepted as a pet by a family, only to be rejected each time. BG and CA take turns reading aloud from a double-spread page, with each interpreting the spider as the main character and the family. They also use a spider soft toy to embody the character's actions, taking turns holding the picturebook and the toy.

Since student teachers were required to work together to design and perform their read-alouds, the self-video recorded read-aloud performances are analysed in this article as collaborative performances. Collaboration was integral to the process, with each participant sharing responsibilities for designing and performing the read-aloud. The student teachers were asked to read aloud approximately equal portions of the written text from the selected picturebooks, ensuring a balanced contribution to the performance.

As part of the study, student teachers completed an online self-reflection questionnaire following the recording of their read-aloud performances. The questionnaire consisted of ten questions aimed at prompting participants to reflect on their use of various performance features, including speech quality, prosodic features, gaze, gestures, picturebook handling, and the use of props. This activity was designed to encourage student teachers to reflect on the role of multimodal communication in facilitating meaning-making and fostering interaction during read-alouds with children. In this article, the student teacher's answers to the self-reflection questionnaires are not analysed in detail; however, they provide additional insights for discussing the mediation strategies used during the read-aloud performances.

#### 4. Multimodal Interaction Analysis: a framework for analysing modal configurations

In this study, MIA (Norris 2016; 2017) provides a holistic framework for analysing the complex ensembles of modes orchestrated by student teachers in interaction. In MIA, the unit of analysis is mediated action—that is, how social actors interact with or through semiotic resources to communicate (Norris 2004; 2019). Mediated actions are examined at two levels. At the lower level, each mode manifests through lower-level actions (Norris 2004). For example, a postural shift represents a concrete lower-level action, while the

broader mode is proxemics (Norris 2020). Coherent ensembles of multiple lower-level actions form higher-level actions, such as performing a character's speech, speculating about an event, or directing the audience's attention to an illustration. Every higher-level action unfolds through the orchestration of multiple modes (e.g., spoken language, gestures, gaze, facial expressions) and is realized through lower-level mediated actions (e.g., intonation shifts, gestures, gaze movements).

In this study, student teachers' self-video recordings are analysed using MIA to examine modal configuration in their read-aloud performances. Modal configuration refers to the hierarchical organization of lower-level actions that constitute a higher-level action and their co-deployment in meaning-making processes. To analyse modal configuration, Norris (2004 2019) introduces the concept of modal density, which helps investigate the complexity and intensity of modes and lower-level actions within a higher-level action. Modal complexity describes how multiple modes are orchestrated to perform an action, while modal intensity refers to the prominence or dominance of a particular communicative mode within a higher-level action and its indispensability in interaction. The analysis of modal configurations in student teacher's read-aloud performances focuses on two higher-level actions performed by student teachers in VR1 and VR2 to mediate the picturebooks (Bertoldi 2025):

- Higher-level actions of scaffolding meaning-making embeds the lower-level actions co-deployed by the student teachers to share the written text of the picturebook facilitating the imagined audience's comprehension.
- Higher-level actions of engaging the imagined audience embeds the lower-level actions co-deployed by the student teachers to invite the imagined audience to engage during the read-aloud performance and to share their verbal and non-verbal responses.

The analysis of student teachers' higher-level action in the self-recorded read-aloud videos involved the annotation and the transcription of lower-level actions time-aligned with video frames using ELAN<sup>7</sup> software. The annotation scheme used to systematically analyse the video recordings is based on a study on multimodal interactions between storytellers and children during read-aloud sessions in EAL (Bertoldi 2024, 2025). The communicative modes analysed in this study are spoken language and prosodic features, gesture, gaze, facial expressions and object handling (distinguishing between picturebooks and props).

Student teachers' spoken language is transcribed distinguishing between the reading aloud of the verbal text of the picturebook and the use of read-aloud talk (Ellis and Mourão 2022), that is spontaneous spoken language that goes beyond the verbal text of the picturebook. Prosodical features are annotated to understand how tone, intonation, pitch, volume and pacing are used to emphasize expressions and convey feelings and emotions (Lwin 2020). Gesture are annotated distinguishing between deictic gesture (i.e., pointing to specific parts of the book, such as illustrations or text, to direct the viewer's attention), iconic gesture (i.e., movements that visually represent elements of the story, such as miming characters' actions), beat gesture (i.e., small rhythmic hand movements that emphasize certain parts of the narrative or add dynamism to the delivery) and metaphoric gesture (i.e., hands or body movements that represent abstract concepts or ideas) (Norris 2004). Gaze shifts are annotated to analyse the way student teachers engage with a hypothetical audience and the way they addressed attention to the picturebook when reading or pointing to illustrations (McIntyre *et al.* 2017). Facial expressions are annotated

<sup>7</sup> <https://archive.mpi.nl/tla/elan>



to analyse how student teachers convey emotions or emphasize key sequences in the story (Kachel *et al.* 2021). Finally, picturebook and props handling are annotated to analyse how the student teachers interact with the objects and how objects contribute to the read-aloud enhancing the meaning-making process and multimodal interaction (Daniel 2012).

The ELAN annotations are displayed in transcription tables generated by TMA<sup>8</sup> (Tool for Multimodal Analysis) software (Bertoldi 2025) for further analysis. TMA is a newly developed software that elaborates annotation files generated through ELAN. It offers researchers functionalities that can contribute to identifying and visualizing co-occurrences and recurrent patterns in multimodal datasets (Bateman 2022; O'Halloran 2018). TMA-generated graphs (see Figure 1 and Figure 3) are used in this article to visualize the overall modal configuration of all lower-level actions in student teachers' performances. Graphs illustrate modal configurations by mapping the relationships and simultaneous occurrences between different communicative modes and transforming it into a network of nodes connected by edges (Barabási and Pósfai 2016; Lim 2023).

In a TMA-generated graph, each node represents a specific group of lower-level actions (i.e., spoken language, deictic gestures, happy facial expressions, etc.). The size of a node indicates the number of connections that specific group of lower-level actions has with other groups of lower-level actions: nodes with more connections are depicted larger, making it easier to identify the most connected lower-level actions in the modal configuration of the performance. Edges between nodes indicate the level of interconnections or co-occurrences between two groups of lower-level actions. The numerical values placed on the edges indicate the weight of edges, that is frequency of interconnections with other lower-level actions. Finally, in TMA-generated graphs, nodes are also color-coded to reflect the clusters to which they belong, with clusters denoting groups of lower-level actions that are frequently combined or performed simultaneously in the performance.

By using graphs as a data visualization technique, the modal intensity of a particular lower-level action is not determined solely by its individual prominence. Rather, it is also identified by level of interconnectedness with other lower-level actions within higher-level actions in the student teacher's performance. From this perspective, the groups of lower-level actions represented by larger nodes with more connections can be considered to have higher modal intensity within higher-level actions, as they are more deeply integrated and intertwined with other modes in the overall complexity of the performance. This data visualization approach provides an innovative way to understand and analyse modal configuration, complexity and intensity, moving beyond the traditional notion of dominance and hierarchy, and instead focusing on the intricate interplay and co-deployment of different communicative modes within a complex multimodal interaction.

In the following section, the analysis of VR1 and VR2 TMA-generated graphs focuses on the clusters with a higher number of interconnected nodes and larger node sizes. These clusters represent ensembles of lower-level actions playing a crucial role in scaffolding meaning-making and engaging the audience, thus having a higher modal intensity and complexity. Smaller clusters with fewer nodes and smaller sizes will not be discussed in detail, as the analysis revealed that, while they are relevant in isolated instances, they do not significantly impact the overall configuration of the performances.

<sup>8</sup> <https://www.tmanalysis.it/>

## 5. Modal configurations in student teachers' read-aloud performances

This section discusses the TMA-generated graphs (see Figure 1 and Figure 3) that display the modal configurations in VR1 and VR2. The presence of various clusters and the different size of nodes in the graphs suggest the complexity and the varying levels of intensity with which student teachers employed various modes to mediate the picturebooks in their performances. In both graphs, the role of spoken language in meaning-making is crucial, due to the dual function of spoken language in read-alouds, which is both the goal of interaction and one of the means through which interaction is instantiated. Thus, spoken language can be considered a “default mode” (Stoeckl and Messner 2021, p. 198) that occurs in all actions performed by student teachers in their performances. The graphs show the co-deployment of spoken language, prosodic features (i.e., the variations in voice tone, intonation, pitch, speed and volume which accompany the intonation units), gestures (i.e., all deictic, iconic and metaphoric gestures used during the performance), shifts in gaze orientation (i.e., looking at the images of the book and looking at the imagined audience of children), body movements (i.e., arm and hand movements), facial expressions (i.e., happy, curious, angry, surprised) and prop handling in student teachers' mediation of the picturebooks.

### 5.1. Modal configuration in VR1: orchestrating modes to convey intercultural connections

In Figure 1, five main clusters of nodes (i.e., red, orange, sand, green and purple) can be identified, these are groups of lower-level actions that are frequently combined or performed simultaneously by the student teachers in VR1.





Figure 1

TMA-generated graph representing the modal configuration of lower-level actions in VR1 (4'11").

The picturebook *Same Same but Different* is written from a first-person perspective, narrating the points of view of the two main characters, who are children exchanging letters with each other. The red cluster in the graph indicates that TE and SG co-deploy spoken language and some prosodic features with the highest intensity to convey the characters' emotions, feelings, and information. Throughout the read-aloud performance, the student teachers consistently employ a high-pitched tone to represent the two children in the story. This is reflected in the largest nodes in Figure 1, which correspond to annotations of spoken language combined with a high pitch of voice. This vocal choice can not only help the audience distinguish the children's voices from that of an adult narrator, but it can also enhance the authenticity of the characters. By using a high pitch, the student teachers make the children's perspectives more salient, potentially fostering identification and empathy among the imagined audience. In *Same, Same but Different*, each double-spread page introduces an aspect of the two children's lives as they share letters with their pen-pal. One child presents a scene from his daily life, such as his home, school, or cultural traditions, and the other child responds by showing a similar aspect from his own perspective. The student teachers give salience to key words mentioned by the characters by lowering their voice volume and slowing down the pace. For example, when the characters in the story mention the members of their families, the student teachers significantly slow the pace and reduce the volume of their voices – this strategy can draw the imagined audience's attention to each expression they utter. Finally, the student teachers change their voice intonation when they exclaim the repeated phrase "Same, same but different!". Despite the differences in their surroundings, traditions, and experiences, children's characters in the picturebook always conclude each exchange with the phrase, "Same, same but different!". This recurring phrase highlights that, while their

lives are distinct, the two characters share commonalities that connect them.

The sand cluster shows that gaze directed towards the images of the picturebook and deictic gestures take on high intensity in specific sequences and that they are frequently co-deployed alongside enthusiastic and expressive vocal tones. This interplay can be found, for example, when student teachers use pointing gestures to draw the imagined audience's attention to key elements in the illustrations (e.g., the house, the flat): the goal is to facilitate the comprehension of key expressions in the narrative.

The orange cluster highlights ensembles of lower-level actions including increased vocal volume and direct gaze, which take on high intensity and are co-deployed with surprised facial expressions and iconic gestures. These ensembles correspond to sequences in which the student teachers present natural elements, such as the trees, the river, and the sun, as mentioned by the two characters in the written text of the picturebook. This multimodal orchestration can contribute to achieving two goals: capturing the audience's attention through variation in voice prosody and scaffolding children's understanding of meaning by visually depicting the verbal expressions associated with specific characteristics of the natural elements. For example, the flowing movement of the river is mimed by one of the student teachers through an iconic hand gesture. The surprised facial expressions can further contribute to stimulating the audience's response, creating an emotional connection to the narrative. Additionally, the student teachers foster a connection with the imagined audience by looking directly at the camera, promoting engagement.

The purple cluster represents ensembles of metaphoric gestures, hands and arms movements and happy facial expressions, with gesture and facial expressions taking on high intensity. In general, these lower-level actions are rarely integrated in the student teacher's performance. However, a significant exception can be observed near the conclusion of the story, where the characters in the picturebook share distinct ways of saying "hello" in their respective countries. In this instance, the student teachers mirror the illustrations by collaboratively performing the greetings, embodying the characters' actions and emotions (Figure 2). In this configuration, gestures and movements assume a high modal intensity as they convey a central message of the picturebook: the importance of fostering interpersonal connections. This orchestration of lower-level actions can play a crucial role in scaffolding the audience's meaning-making process and engagement.



Figure 2

Student teachers performing different ways of saying "hello" in VR2: 03'22''–03'29''.

In the picturebook *Same Same but Different*, each child concludes his letter by asking a question related to an aspect of his life he has just shared with his/her pen-pal. In their performance, the student teachers use these questions as an opportunity to engage the imagined audience. More precisely, they ask questions by directly gazing at the camera, thereby simulating eye contact, and adopting curious facial expressions and a tone of voice that conveys curiosity. The green cluster in the graph highlights this ensemble of lower-level actions which is used as a strategy to stimulate the audience to participate in the narrative, fostering a sense of interaction and connection.

## 5.2. Modal configuration in VR2: embodying a spider's emotions

Figure 3 shows the configuration of lower-level actions in VR2. In this performance, too, language and prosodic features (i.e., increased voice volume, high tone and rising intonation) show a high modal intensity, but in addition also the intensity of use of gestures and facial expressions increases.

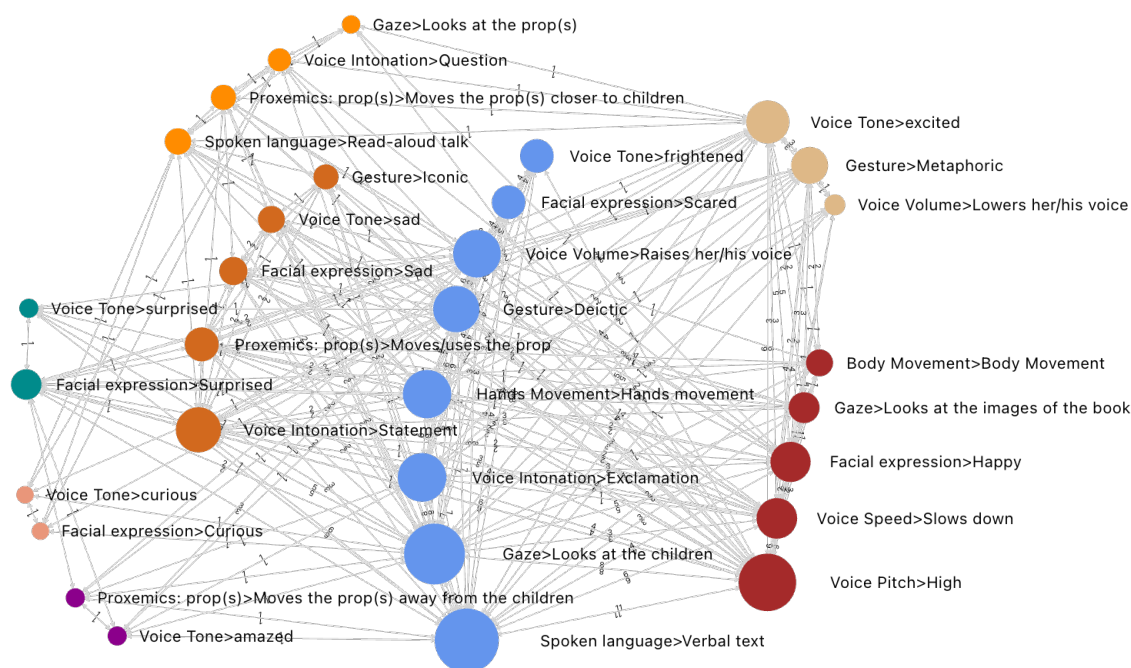


Figure 3

TMA-generated graph representing the modal configuration of lower-level actions in VR2 (4'19'').

Figure 3 reveals four main clusters of nodes (blue, brown, orange, and red), which are the most interconnected both internally and with one another. The most prominent cluster is the blue one, comprising the lower-level actions of spoken language, specifically the student teachers' intonation units (Chafe 1987) while reading aloud the verbal text of the picturebook. In their performance, student teachers BG and CA took turns reading aloud from a double-spread page, each interpreting the roles of the spider as the main character and various members of the family. Notably, within this cluster, we find ensembles of exclamations, increased voice volume, frightened tone of voice, scared facial expressions, and deictic gestures. This corresponds to a repeated pattern that the student teachers recurrently employed in their performance: at the end of each sequence of the story, they collaboratively voiced the family's exclamation, yelling "Aaaarrghh SPIDER! Out you go!" in unison. They coordinated their vocal delivery by raising the volume of their voices and performing a "go away" deictic gesture while using scared facial expressions (see Table 1). Through this multimodal ensemble, student teachers embodied the meaning of the written text, offering visual and aural depictions of spoken language and giving salience to the repeated chunk of language and to the recurring event in the narrative.


Time	Video frame	Lower-level actions - Transcription/Annotation
01':14" 01':21"		<p><b>[Spoken language]:</b> Verbal text of the picturebook</p> <p><b>[BG+CA-Transcription]:</b> Aaaar:::gghh spi:::der! Out you go!</p> <p><b>[Voice Intonation]:</b> Exclamation</p> <p><b>[Voice Tone]:</b> Frightened</p> <p><b>[Voice Volume]:</b> Raise their voices</p> <p><b>[Facial expression]:</b> Scared</p> <p><b>[Gaze]:</b> Look at the imagined audience of children</p> <p><b>[Hands Movement]:</b> Hands movement</p> <p><b>[Gesture]:</b> Deictic</p>

Table 1  
Excerpt from VR2 TMA transcription table.

The brown cluster in Figure 3 includes ensembles of statements, combined with sad facial expressions, a sad tone of voice, the use of the spider soft toy and iconic gestures. Multimodal ensembles embedded in this cluster correspond to sequences in the picturebook where the spider's attempts to please the family fail, leading to expressions of sadness and the formulation of a new idea for gaining acceptance. In their performance, the student teachers embody these recurrent sequences: whenever the spider's attempt is rejected, they use a sad tone of voice and sad facial expression. Concurrently, they utter "Oh Dear" further emphasizing the spider's disappointment through spoken language. They also use iconic gestures (i.e., cry gesture) and the manipulation of the spider soft toy (i.e., they move her legs pretending it is crying too). Immediately following these expressions of sadness, the student teachers transition into narrating the spider's new idea or plan. As the student teachers read sequences like "I know, I'll show them what a great dancer I am", they move the spider toy closer to the camera and slow down their pace, making their speech clearer. Additionally, they raise their pitch and show a happy facial expression in their vocal delivery, conveying the spider's positive emotional state as she finds a new way to impress the family members. The student teachers incorporate also body movements to visually represent the action verbs mentioned in the written text. For example, when reading the lines "None of their pets can dance like me! Look at me! Watch me dance!", they embody the spider's actions by performing dancing movements. In the sequences where the spider presents a new idea or plan, the student teachers' gaze is frequently oriented towards the images in the picturebook. This strategic use of gaze direction can contribute to giving salience to the expression "Look at me!" uttered by the spider character, and it can invite the imagined audience to address their attention to the illustrations, scaffolding the meaning-making process and promoting engagement. These multimodal ensembles are embedded in the red cluster in the TMA-generated graph.

In Figure 3, another cluster emerges: the orange cluster. This cluster includes the use of read-aloud talk, where the student teachers expand beyond the written text of the picturebook by incorporating invitations, questions, and greetings addressed directly to the imagined audience. Specifically, the student teachers utter phrases such as "Would you like to listen to my story?" and "Can you dance like me?" as invitations and questions posed to the imagined audience. They also express greetings like "Hello children!" and "Bye Bye!" to establish rapport with the audience. When delivering these read-aloud talk elements, the student teachers move the spider soft toy as if it were the spider character itself speaking to the audience. In addition, the student teachers' gaze directed towards the soft toy conveys the idea of the spider directly engaging with the student teachers and the audience.

## 6. Discussion and implications for teacher educators

In EAL read-aloud sessions, where children have limited exposure to the target language, most words and expressions in the written text of the picturebook may be unfamiliar when first heard. Thus, the teacher's multimodal mediation strategies during the performance are essential for children to access the meaning potential of the picturebook and participate in interaction. In EAL read-aloud performances, new expressions emerge only when conveyed through non-verbal communication modes (Wright 2013). Storytellers should scaffold communication using expressive techniques such as gestures, facial expressions, body positioning, gaze, and prosodic features of spoken language (Ellis and Mourão 2022). Similarly, Heathfield (2014) highlights the importance of voice, facial expressions, and body language in storytelling, while Masoni (2019) underscores the role of gaze, gestures, voice modulation, and body movement in meaning-making for children. Thus, in EAL contexts, multimodal performance features are fundamental for supporting children's meaning-making process and instantiating interaction, and the effectiveness of mediation strategies hinges on the way various modes are coherently co-deployed.

The analysis of VR1 and VR2 graphs highlights the complex interplay of various communicative modes, the varying levels of modal intensity and the intricate interconnections between lower-level actions employed by the student teachers to mediate the picturebooks. However, in the two read-aloud performances, while spoken language and prosodic features maintain consistently high modal intensity, the analysis revealed that other modes, such as gestures, facial expressions, gaze orientation, body movements, and prop handling (only in VR2), are used with varying levels of modal intensity, often rising in specific sequences or key moments during the performances.

In addition, the results from the self-reflection questionnaires suggest that, while student teachers recognize the importance of combining various semiotic resources, the complexities of multimodal orchestration posed challenges to them. The following excerpts are drawn from the student teachers' final comments in the self-reflection questionnaire, providing insights into their perspectives on the multimodal orchestration:

*"I think my performance is quite clear for children, because I tried to combine gestures, use of puppet and voice in different ways, but there are surely some aspects that can be improve [sic] in order to make the performance more complete" TE (VR1).*

*"I enjoyed doing this activity. I think I should work more on the body movements and Gestures while reading and on the use of the picturebook" SG (VR1).*

*"It was almost the first time doing it in English, so I think I need more practice, especially in front of other people" BG (VR2).*

*"I was quite self-confident but sometimes It was difficult for me reading the book and doing movements at the same time." CA (VR2).*

While all the student teachers orchestrated ensembles of gestures, gaze, voice features, and picturebook handling in their performances, the answers to the self-reflection questionnaires revealed a higher level of confidence in their use of spoken language and prosodic features than in their ability to orchestrate visual resources effectively. Most of the student teachers were satisfied with the quality of their speech and speech delivery. In their opinion, their speech sounded fluent or quite fluent and mainly comprehensible, and the speed of speech delivery was adequate. Answers to the questions about the use of the picturebook during the performance show that the students found it difficult to hold the picturebook and refer to the illustrations during the performance. Similarly, the student teachers' answers related to the use of gaze and gesture revealed varying levels of confidence.

The analysis of VR1 and VR2 revealed that, in their performances, the student teachers foregrounded the modes they felt more confident about, particularly spoken language and prosodic features. Meanwhile, modes they felt less confident about, such as gesture, gaze, and picturebook handling, were often used as supportive, background resources. In Figure 1 and Figure 2, verbal and aural resources emerge as highly intense lower-level actions. This suggests that the student teachers used verbal and vocal features as a primary means for scaffolding meaning-making and engaging the imagined audience. Visual resources such as gesture, gaze patterns, and facial expressions were co-deployed with spoken language less frequently and in specific sequences or key moments of the performances.

Findings suggest that English Laboratory V played a role in developing student teachers' awareness and use of mediation strategies. However, providing more targeted training opportunities could further empower them in effectively orchestrating the multimodal resources available for scaffolding meaning making and engaging children in interaction during EAL read-alouds. Activities should focus on:

- the use of gestures and body movements to provide concrete representations of meaning;
- the manipulation of the picturebook to draw children's attention and give salience to the illustrations;
- the use of gaze to guide children's attention and create a shared focus for contextualized interaction.

## 7. Concluding remarks

This article offered insights into the multimodal mediation strategies employed by two groups of student teachers during EAL read-aloud performances and the affordances of Multimodal Interaction Analysis (MIA) as an analytical framework.

The analysis of configurations of modes in student teachers' read-aloud performances showed that, while spoken language and prosodic features emerged as central modes, used consistently with high intensity, visual modes like gestures, facial expressions, gaze orientation, and prop manipulation were employed with varying levels of intensity, often rising to prominence in specific sequences or key moments. The student teachers recognized the importance of combining various semiotic resources, but they faced challenges in effectively orchestrating visual modes like gestures, gaze, and picturebook handling. The varying levels of intensity observed in the use of visual modes by student teachers during specific sequences or key moments indicate a need for focused training and practice in integrating these modes with spoken language.

As Bland suggests, EAL teachers need to “wear the storyteller’s hat” as frequently as possible to develop a rich repertoire of “ritualized language”, an effective use of “creative teacher talk” and a confident use of “varied paralanguage” (2015, p. 190). The results from this study align with Bland's recommendation. The results highlighted the need for teacher education programmes to provide more hands-on opportunities for student teachers to develop multimodal awareness, embodied communication and reflective practices, enabling them to refine their multimodal mediation skills.

Teacher educators should advocate for a paradigm shift at the institutional level, where multimodal communication and multiliteracies become a foundational component in mainstream EAL teacher education programs, especially for those preparing educators in nursery and primary schools. This will contribute to future teachers developing knowledge and skills to create inclusive and engaging experiences for EAL education with children.

Findings from the present study suggest that encouraging student teachers to



engage in reflective practices through self-video recordings and self-reflection questionnaires can contribute to developing their multimodal awareness. By critically examining their own performances and receiving constructive feedback, student teachers can gain insights into the relevance of effective co-deployment of verbal, aural and visual resources in teaching practices.

Finally, the results from the study highlight the value of incorporating MIA into teacher education practices. MIA offered a systematic framework for a fine-grained analysis of student teachers' strategies for interaction, communication, and embodied communication in educational contexts. Integrating graphs as data visualization techniques provided a powerful means to analyse modal configurations and the intricate interconnections between different modes co-deployed by student teachers for scaffolding meaning-making and promoting engagement during read-aloud performances. These data visualisation techniques offer teacher educators tools for promoting reflection, discussion and deeper understanding of multimodal practices in educational contexts.

The two self-video recordings analysed in this article are a representative sample of the read-aloud performances given by all student teachers at the end of the English Laboratory V in November 2022. However, a larger sample size could potentially reveal additional patterns or variations in modal configurations. Further research could focus also on the development of multimodal competencies in student teachers over time. Longitudinal studies could be conducted to track the progression of student teachers' multimodal competencies providing valuable insights into the development of mediation skills. By following a cohort of student teachers throughout their teacher education programme, researchers could investigate how their understanding and application of multimodal ensembles evolve over time. The findings from these longitudinal studies could inform the design and implementation of targeted training modules or workshops on the effective use of various semiotic resources (e.g., gestures, gaze, voice modulation, props) for mediating communication for additional language education.

**Bionote:** Elisa Bertoldi (PhD) is a postdoctoral researcher at the University of Padua (Italy). She is also a teacher educator for pre-service and in-service courses at the University of Udine (Italy) and a teacher of English as a foreign language in Italian primary schools. She co-authored the book *Let's Tell a Tale. Storytelling with Children in English L2* (Forum, 2019). Her recent research work and publications deal with multimodal interactions in picturebook mediation for English language education.

**Author's adress:** [elisa.bertoldi.1@unipd.it](mailto:elisa.bertoldi.1@unipd.it)



## References

- Barabási A.L. and Pósfai M. 2016, *Network Science*, Cambridge University Press, Cambridge.
- Bateman J.A. 2022, *Growing Theory for Practice: Empirical Multimodality beyond the Case Study*, in “Multimodal Communication” 11 [1], pp. 63-74.
- Bateman J.A., Wildfeuer J. and Hiippala T. 2017, *Multimodality. Foundations, Research And Analysis: A Problem-Oriented Introduction*, De Gruyter Mouton, Berlin.
- Bertoldi E. 2024, *Telling Tales about Nature in English L2: Selecting and Performing Picture Books for Children*, in Gatti C.M. and Hoffmann J. (eds.), *Storytelling as a Cultural Practice. Pedagogical and Linguistic Perspectives*, Peter Lang Publishing House, Berlin, pp. 147-166.
- Bertoldi E. 2025, *Storytelling Events in English L2: Multimodal Orchestration in Picturebook Read-Alouds*, in Mourão S. and Bennett K. (eds.), *Mediation in Multimodal Literature, Education, and Translation: Picturebooks and Graphic Narratives*, Routledge, London and New York, pp. 187-206.
- Bland J. 2015, *Oral Storytelling in the Primary English Classroom*, in Bland J. (ed.), *Teaching English to Young Learners. Critical Issues in Language Teaching with 3–12-Year-Olds*, Bloomsbury Academic, London and New York, pp. 183-198.
- Bland J. 2019, *Teaching English to Young Learners: More Teacher Education and More Children's Literature!*, in “Children's Literature in English Language Education” 7 [2], pp. 79-103.
- Bortoluzzi M., Bertoldi E. and Marenzi I. 2022, *Storytelling with Children in Informal Contexts. Learning to Narrate Across the Offline/Online Boundaries*, in Sindoni M.G. and Moschini I. (eds.), *Multimodal Literacies across Digital Learning Contexts*, Routledge, London and New York, pp. 72-89.
- Campagnaro M. and Goga N. 2022, *Material Green Entanglements: Research on Student Teachers' Aesthetic and Ecocritical Engagement with Picturebooks of their Own Choice*, in “Children's Literature” 15, pp. 308-322.
- Chafe W. 1987, *Cognitive Constraints on Information Flow*, in Russell S. and Tomlin S. (eds.), *Coherence and Grounding in Discourse*, John Benjamins, Amsterdam, pp. 21-51.
- Daly N. and Blakeney-Williams M.M. 2015, *Picturebooks in Teacher Education: Eight Teacher Educators Share their Practice*, in “Australian Journal of Teacher Education” 40 [3], pp. 89-101.
- Daly N. and Short K.G. 2022, *Preservice Teachers' Encounters with Dual Language Picturebooks*, in “Australian Journal of Teacher Education” 47 [7], pp. 76-89.
- Daniel A.K. 2012, *Storytelling across the Primary Curriculum*, Routledge, London and New York.
- Ellis G. 2018, *The Picturebook in Elementary ELT. Multiple Literacies with Bob Staae's BLUEBIRD*, in Bland J. (ed.), *Using Literature in English Language Education: Challenging Reading For 8-18 Years Old*, Bloomsbury, London and New York, pp. 83-104.
- Ellis G. and Brewster J. 2014, *Tell It Again. The Storytelling Handbook for Primary English Language Teachers*, British Council. [https://www.teachingenglish.org.uk/sites/teacheng/files/pub\\_D467/Storytelling\\_handbook\\_FINAL\\_web.pdf](https://www.teachingenglish.org.uk/sites/teacheng/files/pub_D467/Storytelling_handbook_FINAL_web.pdf). (21.01.2025).
- Ellis G. and Mourão S. 2021, *Demystifying the Read-Aloud*, in “English Teaching Professional” 36, pp. 22-25.
- Ellis G. and Mourão S. 2022, *Identifying Quality in Asynchronous, Pre-Recorded Picturebook Read-Alouds for Children Learning English as Another Language*, in Paran A. and Stadler-Heer S. (eds.), *Taking Literature Online: New Perspectives on Literature in Language Learning and Teaching*, Bloomsbury, London and New York, pp. 47-70.
- Fleta T. 2019, *Picturebooks: An Effective Tool to Encourage Children's English L2 Oral Production*, in “TEANGA” Special Issue 10, pp. 243-261.
- Hartmann E. and Helot C. 2021, *The Three Robbers in Three Languages: Exploring a Multilingual Picturebook with Bilingual Student Teachers*, in “Journal of Literary Education” 4, pp. 174-195.
- Heathfield D. 2014, *Storytelling with Our Students: Techniques for Telling Tales from Around the World*, DELTA Teacher Development Series, London.
- Hoffmann J. 2021, *Student Teachers and Kindergarten Children Talking about Picturebooks. Focusing School in Didactic Research Labs at University*, in “Journal of Literary Education” 4, pp. 69-95.
- Kachel G., Hardecker D. and Bohn M. 2021, *Young Children's Developing Ability to Integrate Gestural and Emotional Cues*, in “Journal of Experimental Child Psychology”, 201.
- Kostecki-Shaw J.S. 2011, *Same, Same but Different*, Henry Holt and Co., New York.
- Lim F. 2023, *Mapping the Lesson: Network Graphs and Microgenres*, in “Journal of Education for Teaching” 50, pp. 1-17.
- Lwin S. 2020, *A Multimodal Perspective on Applied Storytelling Performances. Narrativity in Context*, Routledge, London and New York.
- Masoni L. 2019, *Tale, Performance, and Culture in EFL Storytelling with Young Learners. Stories Meant to*

- Be Told*, Cambridge Scholars Publishing, Newcastle upon Tyne.
- McIntyre N.A., Mainhard T. and Klassen R. 2017, *Are You Looking to Teach? Cultural, Temporal and Dynamic Insights into Expert Teacher Gaze*, in “Learning and Instruction” 49, pp. 41-53.
- Monks L. 2013, *Aaaarrgghh, SPIDER!*, Pearson Education Limited, London.
- Nikolajeva M. and Scott C. 2006, *How Picturebooks Work*, Routledge, London and New York.
- Nodelman P. 1988, *Words about Pictures: The Narrative Art of Children’s Picture Books*, University of Georgia Press, Athens.
- Norris S. 2004, *Analyzing Multimodal Interaction. A Methodological Framework*, Routledge, London and New York.
- Norris S. 2016, *Interaction. Language and Modal Configurations in Multimodal Actions in the Art Classroom*, in Klug N. and Stoeckl H. (eds.), *Language in Multimodal Contexts*, De Gruyter Mouton, Berlin.
- Norris S. 2017, *Modal Density and Modal Configurations: Multimodal Actions*, in Jewitt C. (ed.), *The Routledge Handbook of Multimodal Analysis*, 2nd ed., Routledge, London and New York, pp. 86-99.
- Norris S. 2019, *Systematically Working with Multimodal Data. Research Methods in Multimodal Discourse Analysis*, Wiley, Hoboken.
- Norris S. 2020, *Multimodal Theory and Methodology for the Analysis of (Inter)Action and Identity*, Routledge, London and New York.
- O’Halloran K.L., Tan S., Pham D., Bateman J. and Vande-Moere A. 2018, *A Digital Mixed Methods Research Design: Integrating Multimodal Analysis with Data Mining and Information Visualization for Big Data Analytics*, in “Journal of Mixed Methods Research” 12 [1], pp. 11-30.
- Stoeckl H. and Messner M. 2021, *Tam Pam Pam Pam And Mi – Fa – Sol: Constituting Musical Instructions through Multimodal Interaction in Orchestra Rehearsals*, in “Multimodal Communication” 10 [3], pp. 193-209.
- Wright A. 2013, *Stories as Symphonies*, in Bland J. and Lüdege C. (eds.), *Children’s Literature in Second Language Education*, Bloomsbury, London and New York, pp. 205-215.